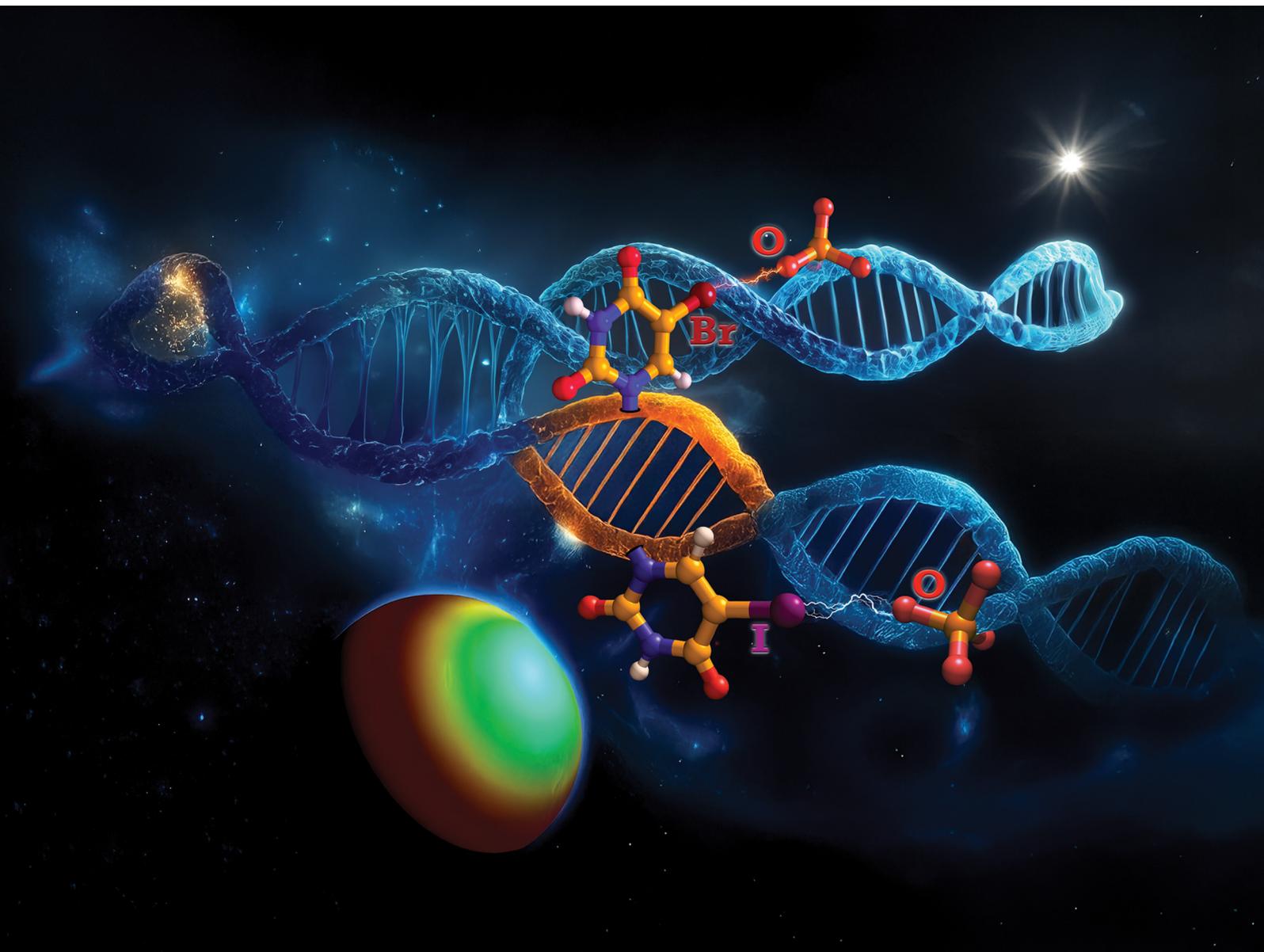


RSC Chemical Biology

rsc.li/rsc-chembio



ISSN 2633-0679

REVIEW ARTICLE

Jorge González-García, Antonio Bauzá *et al.*
Halogenation of nucleic acid structures: from chemical
biology to supramolecular chemistry

Cite this: *RSC Chem. Biol.*, 2025,
6, 1007

Halogenation of nucleic acid structures: from chemical biology to supramolecular chemistry

Catalina Nicolau,^a Julia Requena-Ramírez,^{id}^b Jorge González-García^{id}^{*b} and Antonio Bauzá^{id}^{*a}

Halogenation of organic molecules is a widespread resource used in the fields of chemical biology and rational drug design to improve the binding affinity or solubility of a given compound. Interestingly, the incorporation of Br and I in DNA/RNA bases has been routinely carried out during decades to facilitate the structural determination of nucleic acids, without accounting for its impact in the DNA/RNA structure and molecular recognition events involving other biological entities (e.g. proteins and peptides). This is of critical importance, since halogens undergo non-covalent binding (specially Br and I) through the formation of halogen bonding interactions, thus structurally influencing peptide/protein–DNA/RNA binding poses or altering the supramolecular architecture of isolated nucleic acid structures. In this review, the physical nature of halogen bonds involving nucleobases as well as their implications in (i) the formation of protein–DNA/RNA complexes and (ii) the stabilization of non-canonical DNA/RNA structures will be discussed, focusing on the role of this non-covalent interaction as a promising tool in nucleic acid chemistry.

Received 28th March 2025,
Accepted 22nd May 2025

DOI: 10.1039/d5cb00077g

rsc.li/rsc-chembio

Introduction

Chemical modification of nucleic acids

Chemically modifying DNA and RNA bases is a crucial process in the regulation of gene expression,^{1,2} being essential for understanding the fundamental mechanisms of epigenetics.¹ These modifications affect both the secondary nucleic acid structure and the protein–DNA/RNA binding mechanisms, thereby regulating processes related to cellular survival.³ During the last few stages of the past century, there was a significant increase in research related to the covalent modification of DNA/RNA, primarily due to its promising applications in clinical medicine, such as the use of biomarkers for disease prevention and diagnosis, which has led to a better understanding of the role of epigenetics in various diseases.^{4–7}

The chemical modification of DNA/RNA bases has been detected across all domains of life, including eukaryotic and prokaryotic organisms as well as viruses.⁸ More than 100 types of chemical modifications have been identified in RNA. Their functions are not yet clear, and they are mostly involved in tuning the translational process, controlling their stability and proper folding and determining the identity of the corresponding

RNA.^{9,10} The plethora of non-canonical RNA nucleobases includes simple modifications such as methylation or acetylation in addition to hypermodified ribonucleosides containing different heterocycles, such as queuosine and wyosine.^{11,12} In contrast, there is a small fraction of DNA modifications in comparison with RNA. One notable example is 5-methylcytosine (m⁵C), discovered over 50 years ago, which remains one of the most extensively studied epigenetic marks in mammals to date.¹³ In addition to m⁵C, other canonical DNA modifications include 5-hydroxymethylcytosine (hm⁵C), 5-formylcytosine (f⁵C), 5-carboxylcytosine (ca⁵C), N4-methylcytosine (m⁴C), N6-methyladenine (m⁶A), and N6-hydroxymethyladenine (hm⁶A), among others (see Fig. 1).

These DNA modifications affect not only its conformation and stability but also its protein-binding mechanism. In fact, the malfunction of the epigenetic regulatory mechanisms has deep implications for human health, particularly in diseases related to (1) cancer,^{14–17} (2) obesity¹⁸ and diabetes,^{19,20} and (3) neurodegenerative processes.^{21–24}

On the other hand, the non-canonical modification of nucleic acids encompasses various transformations, such as the incorporation of methyl (–Me), halogen (–X), or aryl (–Ar) groups likely at the C8 position of purine rings^{25–27} and the C5 position of pyrimidine rings (Fig. 2),^{28–30} although N4 and N6 positions are also used. Interestingly, these modifications play a key role in the formation of alternative DNA and RNA structures.^{31,32} For instance, modifications at the C8 position of the purine ring are often used as inducing agents of Z-conformation and guanine quadruplex formation in both DNA and RNA. These modifications

^a Ctra. de Valldemossa, km. 7.5, Universitat de les Illes Balears, 07122, Palma de Mallorca, Islas Baleares, Spain. E-mail: antonio.bauza@uib.es

^b Department of Inorganic Chemistry, Institute of Molecular Science (ICMol), University of Valencia, Catedrático José Beltrán 2, 46980 Paterna, Spain. E-mail: jorge.gonzalez@uv.es



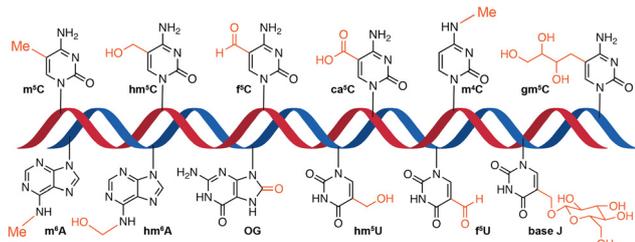


Fig. 1 Schematic representation of modified A and C bases present in DNA. m⁵C, 5-methylcytosine; hm⁵C, 5-hydroxymethylcytosine; f⁵C, 5-formylcytosine; ca⁵C, 5-carboxylcytosine; m⁴C, N4-methylcytosine; gm⁵C, 5-glycerylmethylcytosine; m⁶A, N6-methyladenine; hm⁶A, N6-hydroxymethyladenine; OG, 8-oxo-7,8-dihydroguanine; hm⁵U, 5-hydroxymethyluracil; f⁵U, 5-formyluracil; and base J, β-D-glucosyl-5-hydroxymethyluracil.

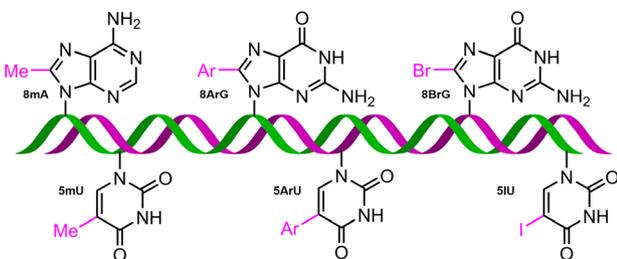


Fig. 2 Schematic representation of non-canonically modified A, G, and U bases. 8mA, 8-methyladenine; 8ArG, 8-arylguanine; 8BrG, 8-bromoguanine; 5mU, 5-methyluracil; 5ArU, 5-aryluracil; and 5IU, 5-iodouracil.

have also been employed to study alternative structures, including H-DNA and DNA triple helices,³³ such as methylation in tRNA,³⁴ which led to the formation of Levitt base pairs instead of the canonical Watson–Crick base pairs.

The conformation of ribose rings (in both DNA and RNA) is also crucial in determining the adoption of a particular nucleic acid structure (see Fig. 3). An example of this can be observed in purines, where conformational equilibrium exists between the *anti* and *syn* conformations of the ribose rings. Under

physiological conditions, the typical duplex DNA conformation is dominated by the B-form, where A · · · T and G · · · C base pairs³⁵ preferentially adopt the *anti* conformation. However, several studies have demonstrated that the introduction of bulky substituents at the C8 position of the purine ring shifts is enough to shift the equilibrium towards the *syn* conformation,^{36,37} thereby promoting the transition to alternative DNA structures such as Z-DNA^{38–40} or guanine quadruplexes.^{41–43} Additionally, the incorporation of halogens at the C2' position of the ribose ring has been studied for its role in stabilizing *i*-motif DNA structures. This phenomenon is correlated with genes that predispose individuals to certain diseases and plays a vital role in DNA's biological functions, as Z-DNA sequences are often located near transcription factor binding sites.^{44,45}

Despite the growing number of studies in this field, the role and conformational properties of non-canonical DNA structures are not yet fully understood. Advancing our knowledge of the effects of such nucleotide and oligonucleotide modifications is therefore crucial for gaining a deeper understanding of their structural implications for nucleic acid stability and protein–DNA binding mechanisms. To achieve this, both experimental techniques^{46,47} (such as mass spectrometry^{48,49} and DNA sequencing studies)⁵⁰ and computational tools (primarily molecular dynamics simulations and docking studies)⁵¹ are employed.

In this context, ongoing efforts to develop novel theoretical tools for studying and characterizing DNA and other biomolecules have established molecular dynamics simulations as a highly valuable tool for nucleic acid analysis. Furthermore, the combination of molecular dynamics and quantum mechanics calculations enables precise insights into the structure and dynamics of large biological systems (such as protein–DNA complexes), which is crucial for deciphering, at an atomistic level, the molecular mechanisms governing nucleic acid modification and recognition.

Nucleic acid halogenation and DNA damage

The presence of mutations in DNA is associated with genomic instability and, consequently, an increased predisposition to

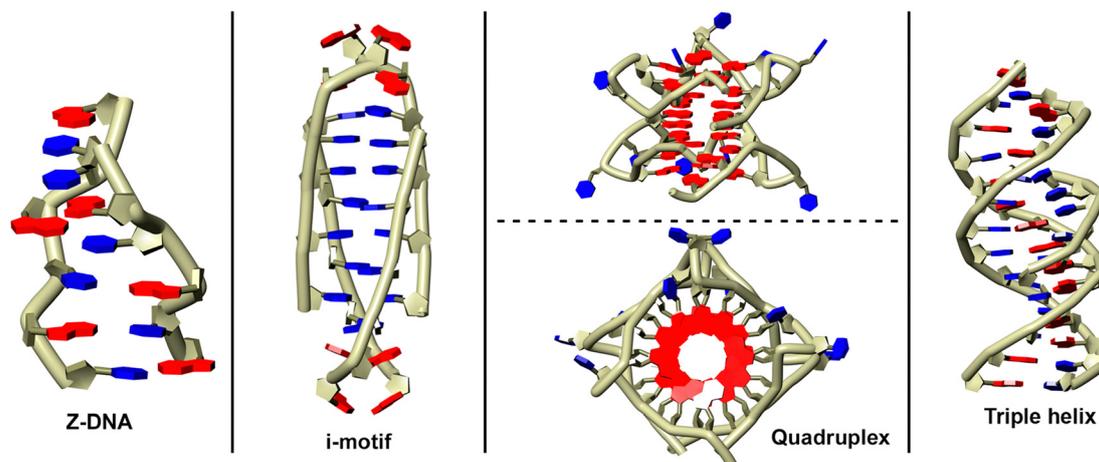


Fig. 3 Schematic representation of several non-canonical DNA structures: Z-DNA (a), *i*-motif (b), G-quadruplex (c), and triple helix (d).



cancer.^{52–55} Concretely, the accumulation of DNA damage over time through these mutations acts as a trigger for uncontrolled cell proliferation and dysfunctional cellular migration mechanisms, both of which are hallmarks of cancer cells.^{56,57}

There is a well-established correlation between mutations and DNA damage, with inflammatory processes being a major cause of mutagenic damage to DNA. A key feature of inflammatory responses is the production of chemical species designed to neutralize pathogenic agents. While these reactive species are essential for protecting the body against infection, they can also cause DNA damage. As a result, DNA repair mechanisms play a critical role in preventing mutations in the genetic code. However, these mechanisms may fail or become overwhelmed by excessive damage, allowing mutations to persist. Since cancer develops through the accumulation of mutations, it is concluded that unrepaired DNA damage due to inflammation contributes to cancer progression by increasing mutagenesis.

Inflammation and genomic instability are closely interconnected (Fig. 4, top), as inflammation promotes mutagenesis through the production of reactive oxygen and nitrogen species (RONS), which can directly damage DNA. In a parallel way, DNA damage can further exacerbate inflammation, creating a self-perpetuating feedback loop being regulated by a plethora of DNA repair mechanisms, transcription factors, and cellular signaling pathways. Due to the complexity of the molecular mechanisms involved in inflammation, DNA damage, and repair, these processes can become dysregulated relatively easily, ultimately leading to cancer.

In this context, various molecules have been classified as RONS, including nitric oxide (NO), which serves as a crucial cellular signaling molecule.^{58–62} Additionally, neutrophils and macrophages produce superoxide anions (O_2^-), while other enzymatic pathways contribute to a cascade of chemical reactions leading to the generation of a diverse array of RONS

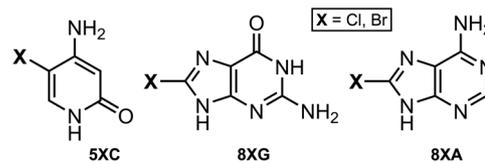


Fig. 5 Schematic representation of the main DNA halogenation products.

(Fig. 4, bottom). These include radicals (hydroxyl radical OH^\bullet and nitrogen dioxide radical NO_2^\bullet), anions (peroxynitrite $ONOO^-$ and nitrosoperoxocarbonate $ONOOCO_2^-$), anhydrides (nitrous anhydride N_2O_3), hypohalogenous acids (hypochlorous acid HOCl and hypobromous acid HOBr), and hydrogen peroxide (H_2O_2).^{63–65} Finally, studies have demonstrated that, in addition to RONS produced by immune cells, pro-inflammatory cytokines can also stimulate intracellular RONS production.^{66–68}

The mutagenic character of halogenated DNA is associated with the generation of hypohalous acids by inflammatory cells. Specifically, neutrophils secrete the enzyme myeloperoxidase, which in turn produces hypochlorous acid (HOCl).^{70–72} On the other hand, eosinophils secrete eosinophil peroxidase, generating hypobromous acid (HOBr).^{73,74} Both molecules react with DNA during inflammatory processes, leading to the formation of halogenated products such as 5-halocytosine (5-haloC), 8-haloguanine (8-haloG), and 8-halo-adenine (8-haloA) (Fig. 5).^{75–78}

Among these halogenated products, the most abundant halogenated base is 5-chlorocytosine (5ClC),^{79,80} which has been detected at significantly higher levels than other DNA lesions resulting from oxidation, deamination, or peroxidation, as demonstrated in *in vivo* studies.^{81,82} Due to its substantial and persistent accumulation in tissues under inflammatory stress—possibly due to inefficient repair mechanisms—as well as in cancers associated with inflammatory pathways,⁸³ 5ClC has been designated as a biomarker of chronic inflammation.⁸⁴

Incorporation of halogenated bases into DNA and RNA

Several studies have analyzed the effect of halogenation on the stability of base pairs in DNA and RNA, both theoretically and experimentally,^{85–100} partly due to its application as radiosensitizing agents in cancer therapies.^{101,102} In the following sections of this review, several selected examples are discussed in more detail.

Evaluating the stability of halogenated base pairs

The first selected example involves the study by Koseki and collaborators,⁸⁵ where they performed quantum mechanics calculations and molecular dynamics simulations of nucleic acids modified with 5-trifluorothymine (FTD) and 5-fluorouracil (5FU). These modified nucleotides, particularly FTD, had previously been used as antitumor agents,^{103–105} however, their structural effects on a DNA strand had not been thoroughly investigated. To address this, theoretical calculations (MP2/6-31++G**) were performed to analyze the stability of $A \cdots FTD$ and $A \cdots 5FU$ base pairs (Fig. 6a).

The results were compared with the canonical, non-halogenated base pairs $A \cdots T$ and $A \cdots U$, revealing a slight increase in stability

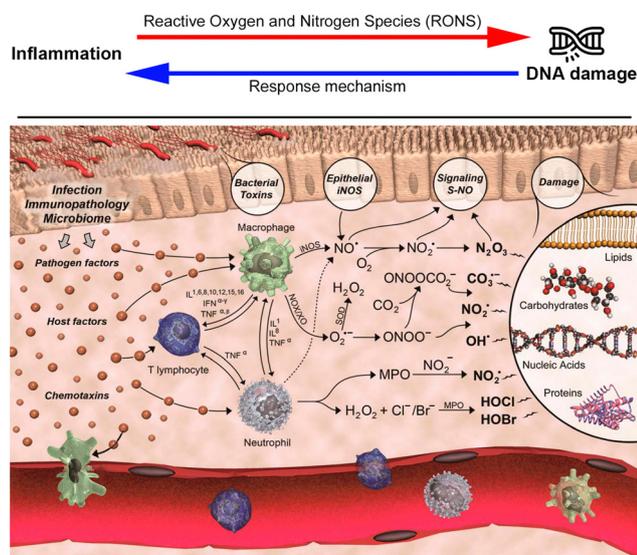


Fig. 4 Schematic representation of the relationship between inflammation and DNA damage, as well as its contribution to cancer (top). The biological chemistry of cancer, adapted from Mangerich and collaborators (bottom).⁶⁹



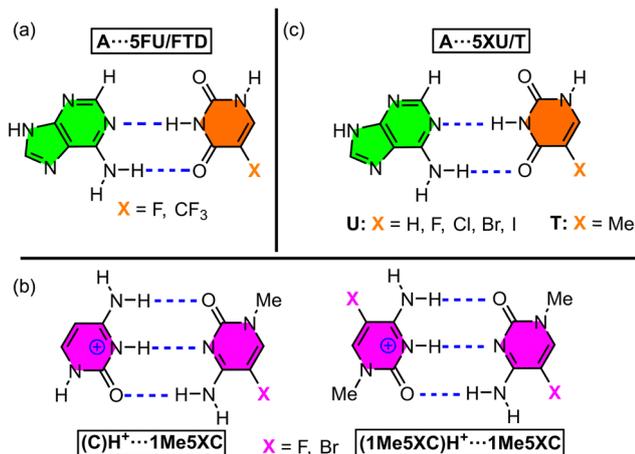


Fig. 6 Schematic representation of the base pairs used in the studies by Koseki⁸⁵ (a), Yang⁸⁶ (b), and Heshmati⁸⁷ (c).

due to the electron-accepting effects of the -F and -CF₃ substituents. Subsequently, molecular dynamics simulations were carried out on a DNA fragment to assess the stability and structural impact of incorporating halogenated nitrogenous bases. The authors concluded that the antitumor activity exhibited by 5-trifluoromethylthymine incorporation was due to the destabilizing effect of the -CF₃ group, which was involved in the formation of intra-strand hydrogen and halogen bonds. Finally, experimental analyses were conducted using UV/visible spectroscopy, along with *in vivo* assays, to complement and corroborate the data obtained from the simulations.

Another relevant study was conducted by Yang and coworkers,⁸⁶ where they evaluated the stability of cytosine (C)H⁺...C base pairs incorporating fluorine (F) and bromine (Br) atoms, and these modifications were involved in the transition from B-DNA to the *i*-motif structure. The authors employed the threshold collision induced dissociation technique using tandem ion beam mass spectrometry to study the formation of cytosine homodimers and heterodimers (Fig. 6b).

The authors concluded that halogenation at the C5 position reduced the strength of the interactions between the studied base pairs.^{106,107} This effect was significantly more pronounced in the stability of base pairs that incorporated a methylated cytosine at the C1 position. The authors suggested that the *i*-motif formation could be destabilized through C5 halogenation; however, these modified base pairs still exhibited greater stability than their neutral nitrogenous base counterparts, thus suggesting that halogenation is not sufficient to prevent *i*-motif formation but can alter the number of nucleotides required to induce its transition from a canonical Watson-Crick helix.

Another example is the work of Heshmati and collaborators,⁸⁷ who conducted computational analyses (B3LYP/6-311G* level of theory) to assess the stability of the A...5FT base pair (Fig. 6c). Their aim was to explore the potential use of halogenated thymine derivatives as DNA radiosensitizers in anticancer therapy¹⁰⁸⁻¹¹⁰ and observed that halogen atom incorporation led to substantial alterations in both the geometry and charge distribution of the thymine molecule, consequently affecting the stability of the A...5FT base pair and eventually provoking base pair opening.

Other interesting studies encompass those conducted by Guerra⁹¹ and Parker⁹² and collaborators, where the stability of A...T and G...C base pairs incorporating halogenated nucleotides was analyzed (Fig. 7a and b). Concretely, Guerra and collaborators⁹¹ carried out theoretical evaluations (BP86/TZVP level of theory) to understand the stability of the base pairs 8XA...6XT and 8XG...6XC (X = H, F, Cl, and Br), in which the hydrogen atom at positions 8 (A, G) and 6 (T, C) was replaced by halogens (Fig. 7a). The authors concluded that the effects varied significantly depending on both the halogenation position and the base pair under study. They observed contrasting effects, such as an enhancement in hydrogen bond donor capacity (N-H) alongside a reduction in hydrogen bond acceptor capacity (O, N). In general, fluorine incorporation into the pyrimidine base reinforced base pair stability, whereas its incorporation into the purine base had the opposite effect. Additionally, the authors highlighted the significant role of orbital interactions as an

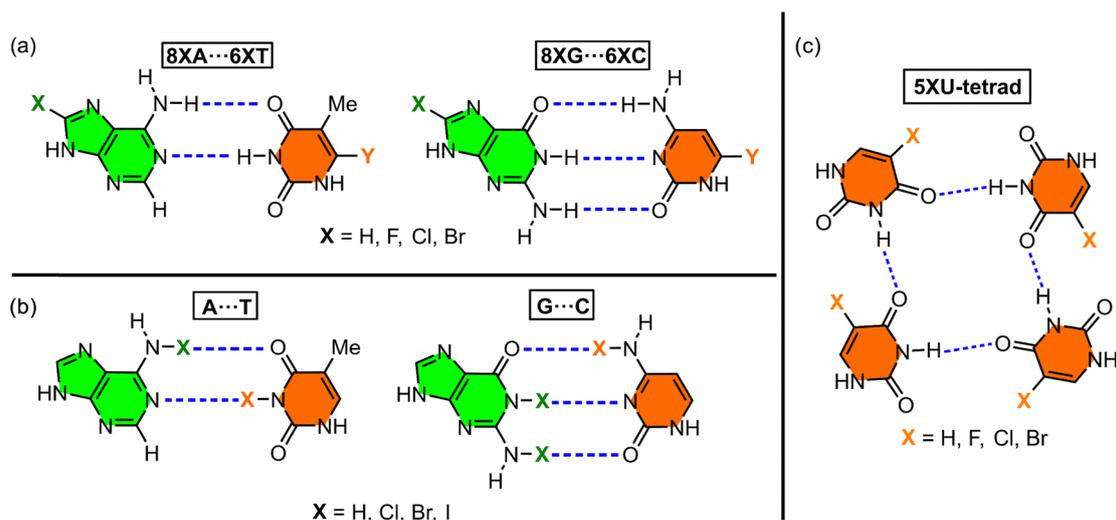


Fig. 7 Schematic representation of the base pairs used in the studies by Guerra⁹¹ (a), Parker⁹² (b), and Zhao⁸⁸ (c).



additional stabilization source beyond electrostatic contributions in these halogenated base pairs.

A final example worth mentioning is the study by Zhao and coworkers,⁸⁸ who investigated the stability of halogenated uracil tetrads (Fig. 7c) using DFT calculations (B3LYP/6-31+G** level of theory). Their results showed that stabilization energy of this ternary complex followed the order $H > F > Cl > Br$. Furthermore, electrostatic potential maps (MEP) revealed that the typical cation-binding site (Na^+/K^+) located at the center of the tetrad disappeared upon incorporation of the 5XU bases. Instead, an electron-deficient region emerged, favoring interactions with electron-rich species.

Beyond their role in modulating base pair stability through their electron-acceptor properties, halogen atoms have also been explored as substituents for conventional hydrogen bonding between bases. Specifically, Parker and collaborators⁹² performed computational studies (B3LYP/6-31G* level of theory) to assess the potential of halogen atoms as direct stabilizers of base pairs through halogen bond formation. They examined the stability of Watson-Crick base pairs $G \cdots C$ and $A \cdots T$, in which Cl, Br, and I replaced hydrogen atoms (Fig. 7b), and compared their stability with non-halogenated base pairs. The base pairs stabilized through halogen bonding exhibited greater stability, following the trend $Cl < Br < I$. Among these, bromine emerged as the ideal halogen bond donor, offering a higher polarizability than chlorine and a smaller atomic volume than iodine.

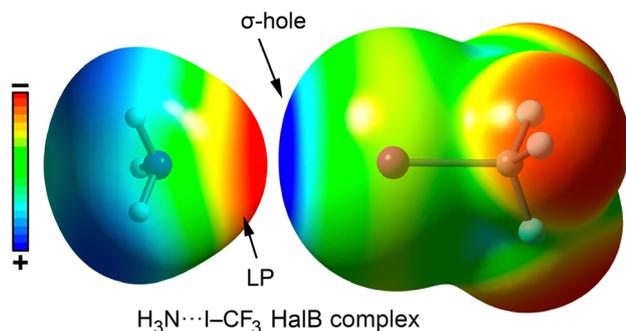


Fig. 8 Graphical representation of a HalB complex between NH_3 and CF_3I using their corresponding electrostatic potential surfaces. Electropositive regions are shown in blue color and electronegative in red color.

Halogen bonds in DNA and RNA supramolecular assemblies

Currently, halogen bonding (a non-covalent attractive interaction between a halogen and a Lewis base)^{111,112} has been established as a valuable molecular tool in various areas of modern chemistry, including solid-state chemistry, materials science, and supramolecular chemistry.¹¹³ In the field of chemical biology, halogen bonding has been mostly investigated in protein-ligand complex formation, becoming an essential tool in the design of novel therapeutic agents within pharmaceutical chemistry.^{114–118}

The stability of halogen bond (HalB) interactions arises from multiple contributions, including charge transfer,¹¹⁹ dispersion forces,¹²⁰ polarization,¹²¹ and electrostatics.^{122,123} Typically, the electrostatic component is based on the attraction between an electron-rich molecule (or region) and a localized region of positive electrostatic potential along the extension of the Y-X bond ($X = F, Cl, Br, I; Y = C$), known as the σ -hole.¹²⁴ This phenomenon is due to the polarization of halogen atoms when covalently bonded to another atom (*e.g.*, C), leading to an anisotropic electrostatic potential distribution around the halogen. Consequently, a σ -hole typically appears along the C-X bond direction, surrounded by a negatively charged σ -lump, associated with the lone electron pairs of the halogen (Fig. 8 and 9).

The σ -hole present on the halogen atom acts as an electrophilic region, facilitating favorable interactions with electron-rich species such as oxygen (O), nitrogen (N), sulfur (S), and π -systems. The polarizability of the halogen atom (which increases in the order $F < Cl < Br < I$) as well as the electron-withdrawing capacity of the group to which the halogen is bonded are key factors in determining the strength and directionality of the interaction. Additionally, the orbital term also contributes to the stabilization of HalBs, being usually based on the electronic donation from a lone pair or a π -system (such as a double/triple bond or an aromatic ring) belonging to the Lewis base to the σ -antibonding orbital* of the C-X bond, further strengthening the interaction.

Ho and collaborators^{125–129} have extensively studied HalB interactions in biological systems, establishing a series of criteria to consider a halogen bond as biologically relevant. These criteria are as follows:

1. It must affect the functionality of the biological system.

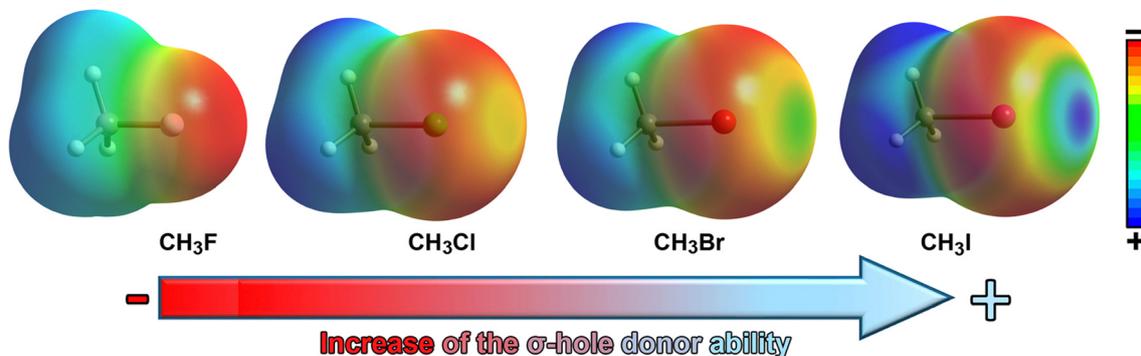


Fig. 9 Molecular electrostatic potential (MEP) surfaces of CH_3F , CH_3Cl , CH_3Br and CH_3I . Electropositive regions are shown in blue color and electronegative in red color.



- It must influence the structures and stability of biomolecules that define the system's function.
- It must possess sufficient stability to impact biological structures and their functions.
- It must exhibit strength comparable to that of a hydrogen bond.

Recently, our group proposed an additional criterion¹³⁰ based on Bader's "atoms in molecules" (QTAIM) theory,¹³¹ consisting of characterizing the HalB interaction through a bond critical point, as outlined in the definition of this interaction proposed by the International Union of Pure and Applied Chemistry (IUPAC).¹⁰⁸

Halogenated nucleotides are also incorporated into DNA or RNA to facilitate the resolution of their crystal structures using techniques such as single-wavelength anomalous dispersion (SAD) or multi-wavelength anomalous dispersion (MAD).^{132,133} These modifications can lead to minimal or localized effects as well as drastic and global structural changes,¹³⁴ which have been studied both experimentally and theoretically.^{135–140} The following sections describe various structures in which halogen bonding plays a fundamental role in stabilizing DNA and RNA architectures.

The first example is the X-ray structure of the sequence d(AC^{Br}UCGGA^{Br}UGA), which was used to resolve its non-brominated analog.¹⁴¹ Both sequences crystallized in a similar architecture (Fig. 10), exhibiting a duplex pattern stabilized by "head-to-head" interactions and intercalated base pairs of C···C, A···A, and G···G. As shown in the right side of the image, the bromine atom present in bromouridine (BrU) forms an inner-strand HalB with an oxygen atom from the phosphate group of an adjacent nucleotide, contributing to local duplex stabilization. Additionally, an inter-strand HalB is observed between two BrU nucleotides, where the bromouridine acts as both the HalB donor and the acceptor, which further stabilizes the supramolecular architecture (the left side of the image).

HalB interactions have also been employed to regulate the formation of a supramolecular biological complex known as the Holliday junction,¹⁴² which is an intermediate structure formed by four DNA strands during cellular recombination mechanisms.¹⁴³ To achieve this, Voth and collaborators¹⁴² replaced a thymine in the sequence d(CCAGTACTGG) with iodouridine (5IU), thereby facilitating the obtention of an X-ray structure (Fig. 11). The results showed that the halogenated structure corresponded to a Holliday junction, whereas the native sequence ultimately crystallized as a standard DNA duplex.¹⁴⁴

Crystallographic analysis of both structures revealed HalB between the 5IU base and an oxygen atom from a neighboring phosphate group, effectively replacing the hydrogen bond typically observed in such structures.^{145,146} Subsequent competitive studies involving other halouridines (5XU, X = F, Cl, Br) demonstrated that O···X HalB exhibited a more stable enthalpy of formation than the hydrogen bond. This finding was further corroborated through differential scanning calorimetry (DSC) and melting studies in solution.

HalBs have also been observed in RNA fragments, including the MS2–RNA complex,¹⁴⁷ which is considered a clear example of halogen bonding in nucleic acid structures,¹⁴⁸ as well as in X-ray structures of U2AF65–RNA complexes (Fig. 12).^{149,150} In the latter, 5BrU was used to accurately determine the position of the pyrimidine base (cytosine or uracil) through X-ray diffraction. More in detail, in the 3VAF structure, the incorporation of 5BrU led to the formation of an O···Br halogen bond with an oxygen atom from a neighboring phosphate group, which played a crucial role in stabilizing the conformation required for protein–RNA recognition.

Halogen bonds also play a role in protein–DNA/RNA recognition mechanisms. One example involves the Tdt protein, which belongs to the family of DNA polymerases responsible for nucleotide addition at the V(D)J union in T-cell receptors. In

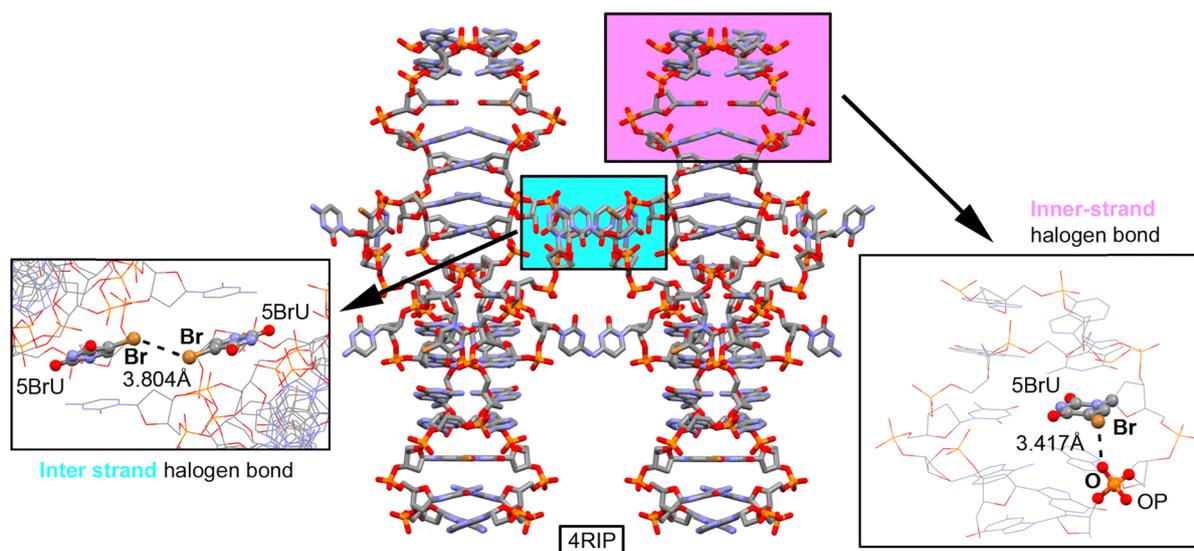


Fig. 10 X-ray structure of a duplex corresponding to the sequence d(AC^{Br}UCGGA^{Br}UGA). HalB interactions are highlighted in purple and cyan and magnified within the rectangular parts of the figure.¹⁴¹



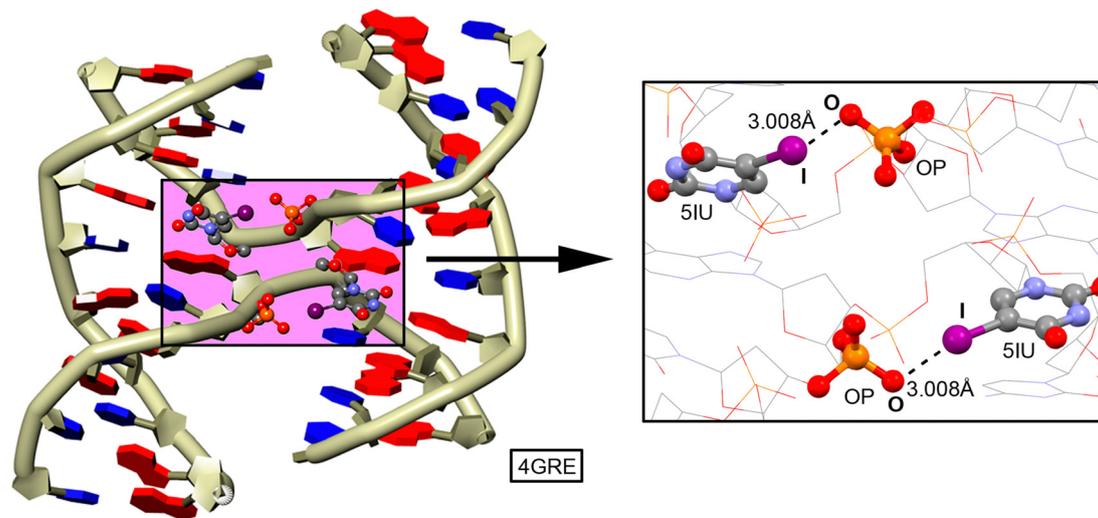


Fig. 11 Stabilizing effect of the HalB in the formation of a Holliday junction structure.¹⁴⁴ The interaction is magnified in the rectangular part of the figure.

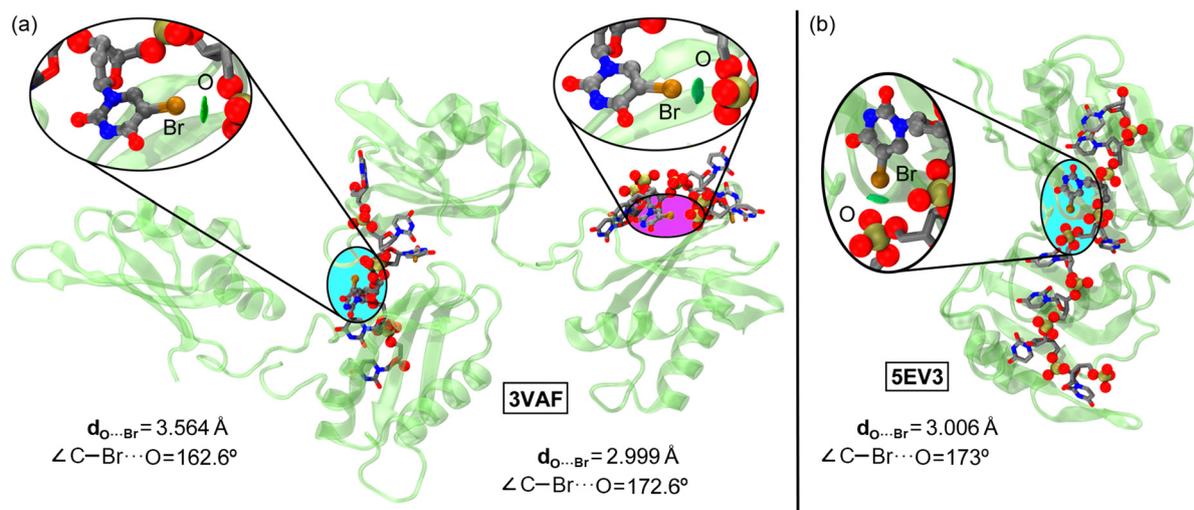


Fig. 12 X-ray structures of 3VAF (a) and 5EV3 (b). The HalB interaction is magnified in the circular parts of the figure. The O...Br distances and C-Br...O angles are also indicated. Visualization of the halogen bond interaction using the NCIplot analysis has been included in ref. 151. Reprinted (adapted) with permission from ref. 152. Copyright 2020 American Chemical Society.

the study by Gouge and collaborators,¹⁵³ a series of DNA oligomers with different lengths and compositions were described as analogs of the enzyme's pre-catalytic and post-catalytic states. To solve the 4I29 structure (Fig. 13a), an adenine was replaced by its brominated analog at the C8 position (8BrA), forming a halogen bond interaction with an oxygen atom from the carboxylate group of the ASP379 amino acid.

The second example involves the nucleotide excision repair (NER) enzyme, which belongs to the family of proteins responsible for DNA repair and serves as a primary defense against DNA damage. In this context, the study by Koch and collaborators¹⁵⁴ analyzed the mechanism by which the NER protein binds to damaged DNA. In the 5A3D structure, some bases were replaced with 5-iodouracil (5IU) to facilitate crystallographic structure resolution. Two of these bases were located in the protein-DNA

binding region and involved in O...I halogen bonds with an oxygen atom from the carboxylate group of the ASP240 amino acid (Fig. 13b), representing an additional non-covalent force contributing to protein-DNA recognition.

The final selected example involves the H/ACA RNP pseudouridine synthase enzyme, which catalyzes the isomerization of uridine to pseudouridine in RNA, a modification that plays a regulatory role in post-transcriptional RNA expression. The study by Zhou and coworkers¹⁵⁵ focuses on the catalytic mechanism of pseudouridylation through the conformational analysis of the glycosidic bond of uridine in a pre-reactive state. To achieve this, the authors utilized both native RNA and RNA containing 5BrU, both labeled with ³²P. Their findings revealed that 5BrU interacted with the protein in a way that stabilized the *anti* conformation of the nucleoside. This interaction was



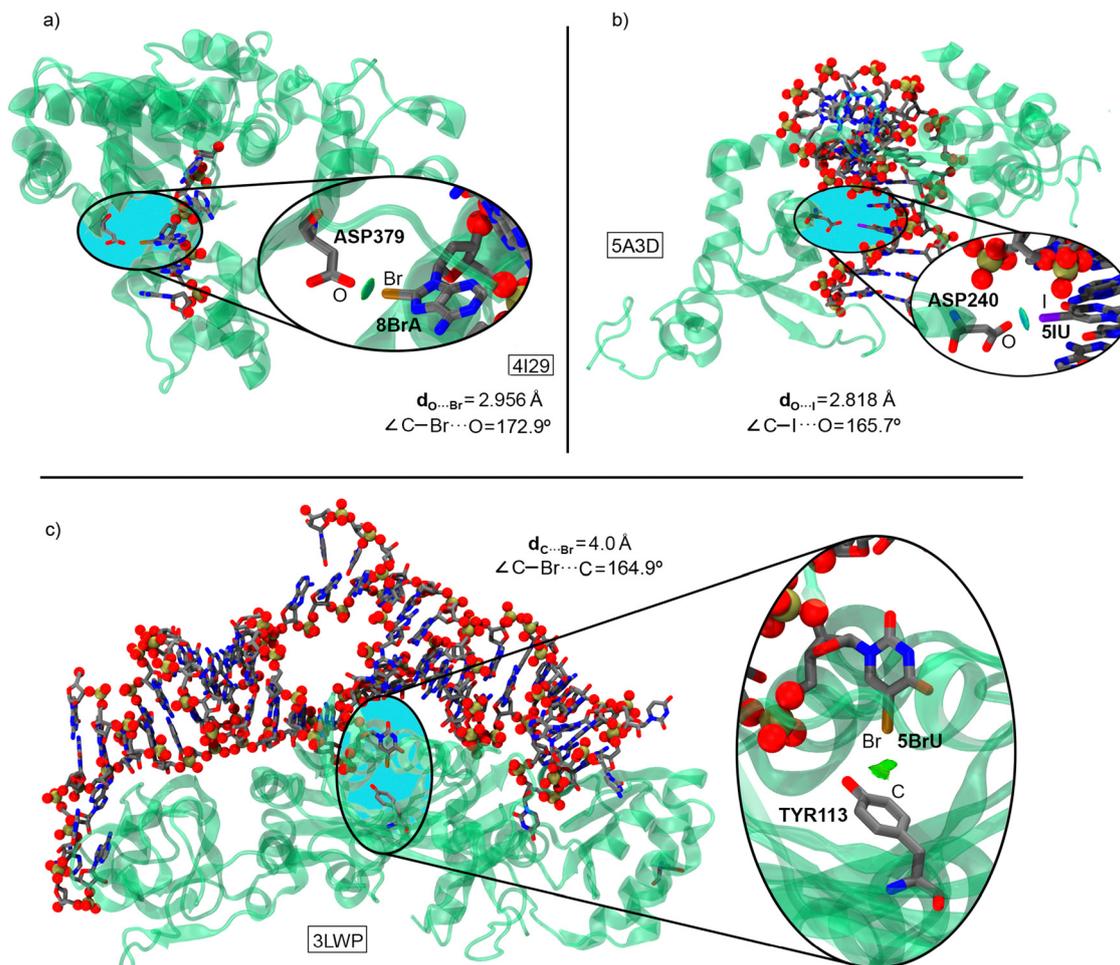


Fig. 13 X-ray structures of 4I29 (a),¹⁵³ 5A3D (b),¹⁵⁴ and 3LWP (c).¹⁵⁵ The HalB interaction is magnified in the circular parts of the figure. The O \cdots Br/I distances and C–Br/I \cdots O angles are also indicated. Visualization of the halogen bond interaction using the NCIPLOT analysis has been included in ref. 151. Reprinted (adapted) with permission from ref. 156. Copyright 2020 American Chemical Society.

identified as a C \cdots Br halogen bond between the bromine atom of uridine and the π -system of the aromatic amino acid TYR113 (Fig. 13c). By obtaining this non-reactive enzyme variant, the study advanced the understanding of its mechanism of action, providing an example of a biologically relevant halogen bond.

Future perspectives

The examples shown above highlighted the impact of nucleic acid halogenation in supramolecular biology. As a future perspective, we propose the use of this non-covalent force to guide the molecular recognition of specific nucleic acid sequences, since the selective recognition of target sites in single-stranded nucleic acids (ssDNA and ssRNA) as well as in more complex structures (*e.g.* G-quadruplexes) is of crucial importance. A routine tool used in the biological area for DNA/RNA detection and sensing is peptide nucleic acids (PNAs), a family of DNA analogs first reported by Nielsen's group in 1991.¹⁵⁷ In the PNA structure, the sugar-phosphate backbone of DNA is replaced by poly[N-(2-aminoethyl)glycine], while the nucleobases (A, G, C, and T) are attached to the central N atom in the monomeric unit through a –COCH₂– linking group (Fig. 14). PNAs exhibit

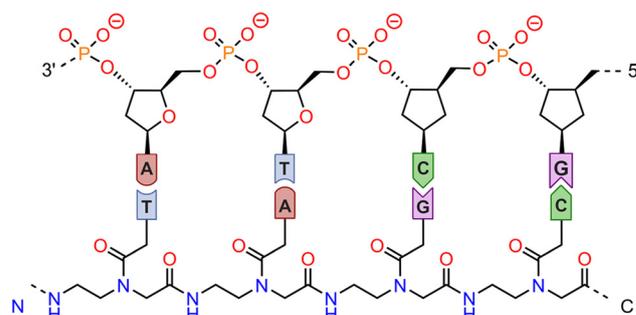


Fig. 14 Schematic representation of a PNA–nucleic acid complex through complementary A \cdots T and G \cdots C pairs.

remarkable ability to form highly stable duplexes with complementary ssDNA and ssRNA. We believe that the incorporation of brominated/iodinated nucleobases into PNAs might trigger a specific base to base recognition motifs and form striking supramolecular DNA assemblies based on HalBs, thus paving the way for new avenues in the field of supramolecular biology. One recent example is the study by Patil and collaborators¹⁵⁸



involving halogenated-U PNAs for dsRNA recognition; however, they rely on base-to-base hydrogen bonding interactions to achieve nucleic acid binding.

Conclusions

In this review, a literature survey related to chemical modifications of nucleic acids, focusing on halogenation, has been carried out. More in detail, the structural and biological consequences of nucleobase halogenation have been discussed alongside several examples that investigate the influence of this chemical modification on the stability of nucleic acid base pairs. Furthermore, we have extended our analysis to the use of halogen bonding interactions involving brominated and iodinated nucleobases as a stabilization source of (i) novel biological supramolecular assemblies, such as “Holliday junctions”, (ii) protein conformations and (iii) protein–nucleic acid complexes. Lastly, we propose the inclusion of halogenated nucleobases into PNAs capable of forming a base to base halogen bonds as new supramolecular tools for the recognition of specific nucleic acid sequences.

Author contributions

C. N. and J. R.-R. performed the data search and J. G.-G and A. B. designed the study and wrote the manuscript.

Data availability

No primary research results, software or code have been included and no new data were generated or analysed as part of this review.

Conflicts of interest

There are no conflicts to declare.

Acknowledgements

C. N. and A. B. thank the MICIU/AEI for financial support (PID2020-115637GB-I00 FEDER funds). J. M. R.-R. and J. G. G. thank the Spanish Ministry for Science and Innovation, The National Research Agency and FEDER funds from the EU (grant MFA/2022/014). J. G.-G. acknowledges the financial support from project CNS2024-154644 funded by MICIU/AEI/10.13039/501100011033. This study forms part of the Advanced Materials program and was supported by MCIN with funding from the European Union Next Generation EU (PRTR-C17.I1) and by Generalitat Valenciana.

Notes and references

- C. H. Waddington, The epigenotype, *Endeavour*, 1942, **1**, 18.
- C. H. Waddington, *Nature*, 1942, **150**, 563–565.

- S. Huang, M. Litt and C. Ann Blakey, *Epigenetic Gene Expression and Regulation*, Elsevier, New York, 2015.
- Y. Chen, T. Hong, S. Wang, J. Mo, T. Tian and X. Zhou, *Chem. Soc. Rev.*, 2017, **46**, 2844–2872.
- S. M. Freier and K. H. Altmann, *Nucleic Acids Res.*, 1997, **25**, 4429–4443.
- V. B. Pinheiro, A. I. Taylor, C. Cozens, M. Abramov, M. Renders, S. Zhang, J. C. Chaput, J. Wengel, S. Y. Peak-Chew, S. H. McLaughlin, P. Herdewijn and P. Holliger, *Science*, 2012, **336**, 341–344.
- S. Ni, H. Yao, L. Wang, J. Lu, F. Jiang, A. Lu and G. Zhang, *Int. J. Mol. Sci.*, 2017, **18**, 1683.
- Y. Motorin and M. Helm, *Wiley Interdiscip. Rev.: RNA*, 2011, **2**, 611–631.
- T. Carell, C. Brandmayr, A. Hienzsch, M. Müller, D. Pearson, V. Reiter, I. Thoma, P. Thumbs and M. Wagner, *Angew. Chem., Int. Ed.*, 2012, **51**, 7110–7131.
- I. A. Roundtree, M. E. Evans, T. Pan and C. He, *Cell*, 2017, **169**, 1187–1200.
- R. J. Ontiveros, J. Stoute and K. F. Liu, *Biochem. J.*, 2019, **476**, 1227–1245.
- W. A. Cantara, P. F. Crain, J. Rozenski, J. A. McCloskey, K. A. Harris, X. Zhang, F. A. P. Vendeix, D. Fabris and P. F. Agris, *Nucleic Acids Res.*, 2011, **39**, D195–D201.
- C. Luo, P. Hajkova and J. R. Ecker, *Science*, 2018, **361**, 1336–1340.
- K. Chen, J. Zhang, Z. Guo, Q. Ma, Z. Xu, Y. Zhou, Z. Xu, Z. Li, Y. Liu, X. Ye, X. Li, B. Yuan, Y. Ke, C. He, L. Zhou, J. Liu and W. Ci, *Cell Res.*, 2015, **26**, 103–118.
- Y. Tang, S. J. Zheng, C. B. Qi, Y. Q. Feng and B. F. Yuan, *Anal. Chem.*, 2015, **87**, 3445–3452.
- A. Seidel, P. Seidel, O. Manuwald and O. Herbarth, *Environ. Toxicol.*, 2015, **30**, 956–967.
- K. L. Mahon, W. Qu, J. Devaney, C. Paul, L. Castillo, R. J. Wykes, M. D. Chatfield, M. J. Boyer, M. R. Stockler, G. Marx, H. Gurney, G. Mallesara, P. L. Molloy, L. G. Horvath and S. J. Clark, *Br. J. Cancer*, 2014, **111**, 1802–1809.
- S. Pfeiffer, J. Krüger, A. Maierhofer, Y. Böttcher, N. Klötting, N. el Hajj, D. Schleinitz, M. R. Schön, A. Dietrich, M. Fasshauer, T. Lohmann, M. Dreßler, M. Stumvoll, T. Haaf, M. Blüher and P. Kovacs, *Sci. Rep.*, 2016, **6**, 27969.
- H. Kirchner, I. Sinha, H. Gao, M. A. Ruby, M. Schönke, J. M. Lindvall, R. Barrès, A. Krook, E. Näslund, K. Dahlman-Wright and J. R. Zierath, *Mol. Metab.*, 2016, **5**, 171–183.
- C. Ling and L. Groop, *Diabetes*, 2009, **58**, 2718–2725.
- R. B. Maccioni, J. P. Munoz and L. Barbeito, *Arch. Med. Res.*, 2001, **32**, 367–381.
- A. J. Phipps, J. C. Vickers, P. C. Taberlay and A. Woodhouse, *Neurobiol. Aging*, 2016, **45**, 30–42.
- L. de Boni, L. Riedel, I. Schmitt, T. F. Kraus, O. Kaut, D. Piston, S. Akbarian and U. Wullner, *Neurobiol. Aging*, 2015, **36**, 3334–3337.
- X. Su, Y. Chu, J. H. Kordower, B. Li, H. Cao, L. Huang, M. Nishida, L. Song, D. Wang and H. J. Federoff, *PLoS One*, 2015, **10**, e134087.
- J. Goodchild, *Bioconjugate Chem.*, 1990, **1**, 165–187.



- 26 O. Doluca, J. M. Withers and V. V. Filichev, *Chem. Rev.*, 2013, **113**, 3044–3083.
- 27 Y. Saito and R. H. E. Hudson, *J. Photochem. Photobiol., C*, 2018, **36**, 48–73.
- 28 L. K. McKenzie, R. El-Khoury, J. D. Thorpe, M. J. Damha and M. Hollenstein, *Chem. Soc. Rev.*, 2021, **50**, 5126–5164.
- 29 R. E. Cline, R. M. Fink and K. Fink, *J. Am. Chem. Soc.*, 1959, **81**, 2521–2527.
- 30 D. M. Frisch and D. W. Visser, *J. Am. Chem. Soc.*, 1959, **81**, 1756–1758.
- 31 R. A. Manderville and S. D. Wetmore, *Chem. Sci.*, 2016, **7**, 3482–3493.
- 32 R. J. Ontiveros, J. Stoute and K. F. Liu, *Biochem. J.*, 2019, **476**, 1227–1245.
- 33 L. E. Xodo, G. Manzini, F. Quadrifoglio, G. A. van der Marel and J. H. van Boom, *Nucleic Acids Res.*, 1991, **19**, 5625–5631.
- 34 V. Y. P. Väre, E. R. Eruysal, A. Narendran, K. L. Sarachan and P. F. Agris, *Biomolecules*, 2017, **7**, 29.
- 35 J. D. Watson and F. H. C. Crick, *Nature*, 1953, **171**, 737–738.
- 36 S. Uesugi and M. J. Ikehara, *Am. Chem. Soc.*, 1977, **99**, 3250–3253.
- 37 V. Vongsutilers, D. J. Phillips, B. C. Train, G. R. McKelvey, N. M. Thomsen, K. H. Shaughnessy, J. P. Lewis and P. M. Gannett, *Biophys. Chem.*, 2011, **154**, 41–48.
- 38 A. Rich and S. Zhang, *Nat. Rev. Genet.*, 2003, **4**, 566–572.
- 39 A. H. Wang, G. J. Quigley, F. J. Kolpak, J. L. Crawford, J. H. van Boom, G. van der Marel and A. Rich, *Nature*, 1979, **282**, 680–686.
- 40 S. C. Ha, K. Lowenhaupt, A. Rich, Y. G. Kim and K. K. Kim, *Nature*, 2005, **437**, 1183–1186.
- 41 J. Gros, F. Rosu, S. Amrane, A. De Cian, V. Gabelica, L. Lacroix and J. L. Mergny, *Nucleic Acids Res.*, 2007, **35**, 3064–3075.
- 42 J. T. Davis, *Angew. Chem., Int. Ed.*, 2004, **43**, 668–698.
- 43 Z. Y. Sun, X. N. Wang, S. Q. Cheng, X. X. Su and T. M. Ou, *Molecules*, 2019, **24**, 396; H. Abou Assi, R. W. Harkness, N. Martin-Pintado, C. J. Wilds, R. Campos-Olivas, A. K. Mittermaier, C. González and M. J. Damha, *Nucleic Acid Res.*, 2016, **44**, 4998–5009.
- 44 B. Wittig, S. Wölfl, T. Dorbic, W. Vahrson and A. Rich, *EMBO J.*, 1992, **11**, 4653–4663.
- 45 A. Maruyama, J. Mimura, N. Harada and K. Itoh, *Nucleic Acids Res.*, 2013, **41**, 5223–5234.
- 46 W. Y. Lai, J. Z. Mo, J. F. Yin, C. Lyu and H. L. Wang, *Trends Anal. Chem.*, 2019, **110**, 173–182.
- 47 M. Berney and J. F. McGouran, *Nat. Rev. Chem.*, 2018, **2**, 332–348.
- 48 J. Xiong, B. F. Yuan and Y. Q. Feng, *Chem. Res. Toxicol.*, 2019, **32**, 808–819.
- 49 M. D. Lan, B. F. Yuan and Y. Q. Feng, *Chin. Chem. Lett.*, 2019, **30**, 1–6.
- 50 M. Berney and J. F. McGouran, *Nat. Rev. Chem.*, 2018, **2**, 332–348.
- 51 M. Karplus and J. A. McCammon, *Nat. Struct. Biol.*, 2002, **9**, 646–652.
- 52 F. Colotta, P. Allavena, A. Sica, C. Garlanda and A. Mantovani, *Carcinogenesis*, 2009, **30**, 1073–1081.
- 53 L. Pikor, K. Thu, E. Vucic and W. Lam, *Cancer Metastasis Rev.*, 2013, **32**, 341–352.
- 54 S. Negrini, V. G. Gorgoulis and T. D. Halazonetis, *Nat. Rev. Mol. Cell Biol.*, 2010, **11**, 220–228.
- 55 D. Hanahan and R. A. Weinberg, *Cell*, 2011, **144**, 646–674.
- 56 M. R. Stratton, P. J. Campbell and P. A. Futreal, *Nature*, 2009, **458**, 719–724.
- 57 B. Vogelstein, N. Papadopoulos, V. E. Velculescu, S. Zhou, L. A. Diaz Jr. and K. W. Kinzler, *Science*, 2013, **339**, 1546–1558.
- 58 D. S. Bredt and S. H. Snyder, *Annu. Rev. Biochem.*, 1994, **63**, 175–195.
- 59 S. Moncada, R. M. Palmer and E. A. Higgs, *Pharmacol. Rev.*, 1991, **43**, 109–142.
- 60 D. D. Thomas, L. A. Ridnour, J. S. Isenberg, W. Flores-Santana, C. H. Switzer, S. Donzelli, P. Hussain, C. Vecoli, N. Paolucci, S. Ambs, C. A. Colton, C. C. Harris, D. D. Roberts and D. A. Wink, *Free Radical Bio. Med.*, 2008, **45**, 18–31.
- 61 R. S. Lewis, S. Tamir, S. R. Tannenbaum and W. M. Deen, *J. Biol. Chem.*, 1995, **270**, 29350–29355.
- 62 D. J. Stuehr and M. A. Marletta, *Cancer Res.*, 1987, **47**, 5590–5594.
- 63 R. Takahashi, K. Edashige, E. F. Sato, M. Inoue, T. Matsuno and K. Utsumi, *Arch. Biochem. Biophys.*, 1991, **285**, 325–330.
- 64 J. A. Badwey and M. L. Karnovsky, *Annu. Rev. Biochem.*, 1980, **49**, 695–726.
- 65 E. Van den Worm, C. J. Beukelman, A. J. Van den Berg, B. H. Kroes, R. P. Labadie and H. Van Dijk, *Eur. J. Pharmacol.*, 2001, **433**, 225–230.
- 66 I. M. Colin, S. Poncin, P. Leveque, B. Gallez and A. C. Gerard, *Thyroid*, 2014, **24**, 441–452.
- 67 M. Sundaresan, F. V. J. Yu ZX, D. J. Sulciner, J. S. Gutkind, K. Irani, P. J. Goldschmidt-Clermont and T. Finkel, *Biochem. J.*, 1996, **318**, 379–382.
- 68 D. Yang, S. G. Elner, Z. M. Bian, G. O. Till, H. R. Petty and V. M. Elner, *Exp. Eye Res.*, 2007, **85**, 462–472.
- 69 A. Mangerich, P. C. Dedon, J. G. Fox, S. R. Tannenbaum and G. N. Wogan, *Free Radic. Res.*, 2013, **47**, 958–986.
- 70 B. S. van der Veen, M. P. de Winther and P. Heeringa, *Antioxid. Redox Signaling*, 2009, **11**, 2899–2937.
- 71 C. C. Winterbourn, M. B. Hampton, J. H. Livesey and A. J. Kettle, *J. Biol. Chem.*, 2006, **281**, 39860–39869.
- 72 J. P. Henderson, J. Byun and J. W. Heinecke, *J. Biol. Chem.*, 1999, **274**, 33440–33448.
- 73 J. P. Henderson, J. Byun, M. V. Williams, M. L. McCormick, W. C. Parks, L. A. Ridnour and J. W. Heinecke, *Proc. Natl. Acad. Sci. U. S. A.*, 2001, **98**, 1631–1636.
- 74 H. Ohshima, M. Tatemichi and T. Sawa, *Arch. Biochem. Biophys.*, 2003, **417**, 3–11.
- 75 T. Asahi, H. Kondo, M. Masuda, H. Nishino, Y. Aratani, Y. Naito, T. Yoshikawa, S. Hisaka, Y. Kato and T. Osawa, *J. Biol. Chem.*, 2010, **285**, 9282–9291.
- 76 C. Badouard, M. Masuda, H. Nishino, J. Cadet, A. Favier and J. L. Ravanat, *J. Chromatogr. B: Anal. Technol. Biomed. Life Sci.*, 2005, **827**, 26–31.



- 77 M. Masuda, T. Suzuki, M. D. Friesen, J. L. Ravanat, J. Cadet, B. Pignatelli, H. Nishino and H. Ohshima, *J. Biol. Chem.*, 2001, **276**, 40486–40496.
- 78 Z. Shen, W. Wu and S. L. Hazen, *Biochemistry*, 2000, **39**, 5474–5482.
- 79 Y. Kawai, H. Morinaga, H. Kondo, N. Miyoshi, Y. Nakamura, K. Uchida and T. Osawa, *J. Bio. Chem.*, 2004, **279**, 51241–51249.
- 80 J. I. Kang Jr. and L. C. Sowers, *Chem. Res. Toxicol.*, 2008, **21**, 1211–1218.
- 81 A. Mangerich, C. G. Knutson, N. M. Parry, S. Muthupalani, W. Ye, E. Prestwich, L. Cui, J. L. McFaline, M. Mobley, Z. Ge, K. Taghizadeh, J. S. Wishnok, G. N. Wogan, J. G. Fox, S. R. Tannenbaum and P. C. Dedon, *Proc. Natl. Acad. Sci. U. S. A.*, 2012, **109**, E1820–1829.
- 82 B. I. Fedeles, B. D. Freudenthal, E. Yau, V. Singh, S. C. Chang, D. Li, J. C. Delaney, S. H. Wilson and J. M. Essigmann, *Proc. Natl. Acad. Sci. U. S. A.*, 2015, **112**, E4571–4580.
- 83 Y. Sato, S. Takahashi, Y. Kinouchi, M. Shiraki, K. Endo, Y. Matsumura, Y. Kakuta, M. Tosa, A. Motida, H. Abe, G. Imai, H. Yokoyama, E. Nomura, K. Negoro, S. Takagi, H. Aihara, K. I. Masumura, T. Nohmi and T. Shimosegawa, *Carcinogenesis*, 2006, **27**, 1068–1073.
- 84 C. G. Knutson, A. Mangerich, Y. Zeng, A. R. Raczynski, R. G. Liberman, P. Kang, W. Ye, E. G. Prestwich, K. Lu, J. S. Wishnok, J. R. Korzenik, G. N. Wogan, J. G. Fox, P. C. Dedon and S. R. Tannenbaum, *Proc. Natl. Acad. Sci. U. S. A.*, 2013, **110**, E2332–2341.
- 85 J. Koseki, M. Konno, A. Asai, N. Horie, K. Tsunekuni, K. Kawamoto, S. Obika, Y. Doki, M. Mori and H. Ishii, *Sci. Rep.*, 2020, **10**, 1138.
- 86 B. Yang, R. R. Wu and M. T. Rodgers, *J. Am. Soc. Mass Spectrom.*, 2015, **26**, 1469–1482.
- 87 E. Heshmati, P. Abdolmaleki, H. Mozdarani and A. S. Sarvestani, *Bioorg. Med. Chem. Lett.*, 2009, **19**, 5256–5260.
- 88 X. Zhao and F. Meng, *Theochem*, 2006, **770**, 157–162.
- 89 Z. Yang and M. T. Rodgers, *J. Am. Chem. Soc.*, 2004, **126**, 16217–16226.
- 90 M. Hanus, M. Kabelac, D. Nachtigallova and P. Hobza, *Biochemistry*, 2005, **44**, 1701–1707.
- 91 C. F. Guerra, T. van der Wijst and F. M. Bickelhaupt, *Struct. Chem.*, 2005, **16**, 211–221.
- 92 A. J. Parker, J. Stewart, K. J. Donald and C. A. Parish, *J. Am. Chem. Soc.*, 2012, **134**, 5165–5172.
- 93 R. Xu, Z. Hu, Z. Wang, Y. Liu, Z. Zhou, J. Xu, Z. Sun, H. Sun and J. Chen, *J. Phys. Chem. B*, 2020, **124**, 2560–2567.
- 94 V. Valinluck, W. Wu, P. Liu, J. W. Neidigh and L. C. Sowers, *Chem. Res. Toxicol.*, 2006, **19**, 556–562.
- 95 Z. Yang and M. T. Rodgers, *J. Am. Chem. Soc.*, 2004, **126**, 16217–16226.
- 96 F. Seela, X. Peng and H. Li, *J. Am. Chem. Soc.*, 2005, **127**, 7739–7751.
- 97 M. Hanus, M. Kabelác, D. Nachtigallová and P. Hobza, *Biochemistry*, 2005, **44**, 1701–1707.
- 98 E. Cubero, E. C. Sherer, F. J. Luque, M. Orozco and C. A. Laughton, *J. Am. Chem. Soc.*, 1999, **121**, 8653.
- 99 E. Cubero, C. A. Laughton, F. J. Luque and M. Orozco, *J. Am. Chem. Soc.*, 2000, **122**, 6891.
- 100 M. Orozco, B. Hernández and F. J. Luque, *J. Phys. Chem. B*, 1998, **102**, 5228.
- 101 C. J. McGinn, D. S. Shewach and T. S. Lawrence, *J. Natl. Cancer Inst.*, 1996, **88**, 1193.
- 102 A. K. Vallerga, D. A. Zarling and T. J. Kinsella, *Clin. Adv. Hematol. Oncol.*, 2004, **2**, 793.
- 103 C. Heidelberger, D. G. Parsons and D. C. Remy, *J. Med. Chem.*, 1964, **7**, 1–5.
- 104 E. P. Gillis, K. J. Eastman, M. D. Hill, D. J. Donnelly and N. A. Meanwell, *J. Med. Chem.*, 2015, **58**, 8315–8359.
- 105 H. Wójtowicz-Rajchel, *J. Fluorine Chem.*, 2012, **143**, 11–48.
- 106 B. Yang, R. R. Wu and M. T. Rodgers, *Anal. Chem.*, 2013, **85**, 11000–11006.
- 107 B. Yang and M. T. Rodgers, *J. Am. Chem. Soc.*, 2014, **136**, 282–290.
- 108 F. J. Sullivan, L. L. Herscher, J. A. Cook, J. Smith, S. M. Steinberg, A. H. Epstein, E. H. Oldfield, T. E. Goffman, T. J. Kinsella, J. B. Mitchell and E. Glatstein, *Int. J. Radiat. Oncol., Biol., Phys.*, 1994, **30**, 583–590.
- 109 R. C. Urtasun, T. J. Kinsella, N. Farnan, J. D. Delrowe, S. G. Lester and D. S. Fulton, *Int. J. Radiat. Oncol., Biol., Phys.*, 1996, **36**, 1163–1167.
- 110 J. A. Williams, X. Yuan, L. E. Dillehay, V. R. Shastri, H. Brem and J. R. Williams, *Int. J. Radiat. Oncol., Biol., Phys.*, 1998, **42**, 631–639.
- 111 G. R. Desiraju, P. S. Ho, L. Kloos, A. C. Legon, R. Marquardt, P. Metrangolo, P. Politzer, G. Resnati and K. Rissanen, *Pure Appl. Chem.*, 2013, **85**, 1711–1713.
- 112 P. Metrangolo, T. Pilati and G. Resnati, *CrystEngComm*, 2006, **8**, 946–947.
- 113 G. Cavallo, P. Metrangolo, R. Milani, T. Pilati, A. Priimagi, G. Resnati and G. Terraneo, *Chem. Rev.*, 2016, **116**, 2478–2601.
- 114 P. Y. S. Lam, C. G. Clark, A. M. Smallwood and R. S. Alexander, *Presented in part at The 238th ACS National Meeting*, Washington, DC, 2009.
- 115 Y. Lu, T. Shi, Y. Wang, H. Yang, X. Yan, X. Luo, H. Jiang and W. Zhu, *J. Med. Chem.*, 2009, **52**, 2854–2862.
- 116 Z. Xu, Z. Liu, T. Chen, T. Chen, Z. Wang, G. Tian, J. Shi, X. Wang, Y. Lu, X. Yan, G. Wang, H. Jiang, K. Chen, S. Wang, Y. Xu, J. Shen and W. Zhu, *J. Med. Chem.*, 2011, **54**, 5607–5611.
- 117 P. S. Ho, *Future Med. Chem.*, 2017, **9**, 637–640.
- 118 R. Wilcken, M. O. Zimmermann, A. Lange, A. C. Joerger and F. M. Boeckler, *J. Med. Chem.*, 2013, **56**, 1363–1388.
- 119 O. Hassel, *Presented in part at Nobel Lectures, Chemistry 1963–1970*, Amsterdam, 1972.
- 120 K. E. Riley, M. Vazquez, C. Umemura, C. Miller and K. A. Tran, *Chem. – Eur. J.*, 2016, **22**, 17690–17695.
- 121 T. Clark, J. S. Murray and P. Politzer, *Aust. J. Chem.*, 2013, **67**, 451–456.
- 122 K. Eskandari and H. Zariny, *Chem. Phys. Lett.*, 2010, **492**, 9–13.
- 123 H. J. Jahromi and K. Eskandari, *Struct. Chem.*, 2013, **24**, 1281–1287.



- 124 T. Clark, M. Hennemann, J. S. Murray and P. Politzer, *J. Mol. Model.*, 2007, **13**, 291–296.
- 125 P. Auffinger, F. A. Hays, E. Westhof and P. S. Ho, *Proc. Natl. Acad. Sci. U. S. A.*, 2004, **101**, 16789–16794.
- 126 P. S. Ho, in *Halogen Bonding I: Impact on Materials Chemistry and Life Sciences*, ed P. Metrangolo and G. Resnati, Springer, 2015, vol. 8, pp. 241–276.
- 127 P. S. Ho, *Top. Curr. Chem.*, 2015, **358**, 241–276.
- 128 M. C. Ford and P. S. Ho, *J. Med. Chem.*, 2016, **59**, 1655–1670.
- 129 R. S. Czarny, A. N. Ho and P. S. Ho, *Chem. Rec.*, 2021, **21**, 1240–1251.
- 130 A. Frontera and A. Bauzá, *Org. Biomol. Chem.*, 2021, **19**, 6858–6864.
- 131 R. F. W. Bader, *Chem. Rev.*, 1991, **91**, 893–928.
- 132 C. M. Ogata, W. A. Hendrickson, X. Gao and D. Patel, *J. Abstr. Cryst. Assoc. Mtg. Ser.*, 1989, **217**, 53.
- 133 Y. Xiong and M. Sundaralingam, *Nucleic Acids Res.*, 2000, **28**, 2171–2176.
- 134 E. Ennifar, S. Bernacchi, P. Wolff and P. Dumas, *RNA*, 2007, **13**, 1445–1452.
- 135 N. J. Thornton and T. van Mourik, *Int. J. Mol. Sci.*, 2020, **21**, 6571.
- 136 M. N. Piña, A. Frontera and A. Bauzá, *ACS Chem. Biol.*, 2020, **15**, 1942–1948.
- 137 D. Mitra, N. Bankoti, D. Michael, K. Sekar and T. N. Guru Row, *J. Chem. Sci.*, 2020, **132**, 1–11.
- 138 A. Frontera and A. Bauzá, *J. Chem. Theory Comput.*, 2020, **16**, 4744–4752.
- 139 M. Marín-Luna, I. Alkorta and J. Elguero, *Comput. Theor. Chem.*, 2016, **1076**, 101–108.
- 140 L. Xu, P. Sang, J. W. Zou, M. B. Xu, X. M. Li and Q. S. Yu, *Chem. Phys. Lett.*, 2011, **509**, 175–180.
- 141 S. Tripathi, D. Zhang and P. J. Paukstelis, *Nucleic Acids Res.*, 2015, **43**, 1937–1944.
- 142 A. R. Voth, F. A. Hays and P. S. Ho, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 6188–6193.
- 143 F. A. Hays, A. Teegarden, Z. J. Jones, M. Harms, D. Raup, J. Watson, E. Cavaliere and P. S. Ho, *Proc. Natl. Acad. Sci. U. S. A.*, 2015, **102**, 7157–7162.
- 144 F. A. Hays, J. M. Vargason and P. S. Ho, *Biochemistry*, 2003, **42**, 9586–9597.
- 145 B. F. Eichman, J. M. Vargason, B. H. M. Mooers and P. S. Ho, *Proc. Natl. Acad. Sci. U. S. A.*, 2000, **97**, 3971–3976.
- 146 B. F. Eichman, M. Ortiz-Lombardia, J. Aymami, M. Coll and P. S. Ho, *J. Mol. Biol.*, 2002, **320**, 1037–1051.
- 147 E. Grahm, T. Moss, C. Helgstrand, K. Fridborg, M. Sundaram, K. Tars, H. Lago, N. J. Stonehouse, D. R. Davis, P. G. Stockley and L. Liljas, *RNA*, 2001, **7**, 1616–1627.
- 148 M. H. Kolář and O. Tabarrini, *J. Med. Chem.*, 2017, **60**, 8681–8690.
- 149 J. L. Jenkins, A. A. Agrawal, A. Gupta, M. R. Green and C. L. Kielkopf, *Nucleic Acids Res.*, 2013, **41**, 3859–3873.
- 150 A. A. Agrawal, E. Salsi, R. Chatrikhi, S. Henderson, J. L. Jenkins, M. R. Green, D. N. Ermolenko and C. L. Kielkopf, *Nat. Commun.*, 2016, **7**, 10950.
- 151 J. Contreras-García, E. R. Johnson, S. Keinan, R. Chaudret, J. P. Piquemal, D. N. Beratan and W. Yang, *J. Chem. Theory Comput.*, 2011, **7**, 625–632.
- 152 M. N. Piña, A. Frontera and A. Bauzá, *ACS Chem. Biol.*, 2020, **15**, 1942–1948.
- 153 J. Gouge, S. Rosario, F. Romain, P. Beguin and M. Delarue, *J. Mol. Biol.*, 2013, **425**, 4334–4352.
- 154 S. C. Koch, J. Kuper, K. L. Gasteiger, N. Simon, R. Strasser, D. Eiser, S. Geiger, S. Schneider, C. Kisker and T. Carell, *Proc. Natl. Acad. Sci. U. S. A.*, 2015, **112**, 8272–8277.
- 155 J. Zhou, C. Lv, B. Liang, M. Chen, W. Yang and H. Li, *J. Mol. Biol.*, 2010, **401**, 690–695.
- 156 A. Frontera and A. Bauzá, *J. Chem. Theory Comput.*, 2020, **16**, 4744–4752.
- 157 P. E. Nielsen, M. Egholm, R. H. Berg and O. Buchardt, *Science*, 1991, **254**, 1497–1500.
- 158 K. M. Patil, D. F. K. Toh, Z. Yuan, Z. Meng, Z. Shu, H. Zhang, A. A. Lerk Ong, M. S. Krishna, L. Lu, Y. Lu and G. Chen, *Nucleic Acids Res.*, 2018, **46**, 7506–7521.

