



**HAL**  
open science

## Interactions of small molecules with DNA junctions

Kane T Mcquaid, Angélique Pipier, Christine J Cardin, David Monchaud

► **To cite this version:**

Kane T Mcquaid, Angélique Pipier, Christine J Cardin, David Monchaud. Interactions of small molecules with DNA junctions. *Nucleic Acids Research*, 2022, 10.1093/nar/gkac1043. hal-03855128v1

**HAL Id: hal-03855128**

**<https://hal.science/hal-03855128v1>**

Submitted on 16 Nov 2022 (v1), last revised 17 Jan 2023 (v2)

**HAL** is a multi-disciplinary open access archive for the deposit and dissemination of scientific research documents, whether they are published or not. The documents may come from teaching and research institutions in France or abroad, or from public or private research centers.

L'archive ouverte pluridisciplinaire **HAL**, est destinée au dépôt et à la diffusion de documents scientifiques de niveau recherche, publiés ou non, émanant des établissements d'enseignement et de recherche français ou étrangers, des laboratoires publics ou privés.

# Critical Reviews and Perspectives

## Interactions of small molecules with DNA junctions

Kane T. McQuaid<sup>1</sup>, Angélique Pipier<sup>2</sup>, Christine J. Cardin<sup>1,\*</sup> and David Monchaud<sup>1,2,\*</sup>

<sup>1</sup>Department of Chemistry, University of Reading, Whiteknights, Reading RG6 6AD, UK and <sup>2</sup>Institut de Chimie Moléculaire de l'Université de Bourgogne (ICMUB), CNRS UMR 6302, UBFC Dijon, 21078 Dijon, France

Received August 08, 2022; Revised October 13, 2022; Editorial Decision October 15, 2022; Accepted October 23, 2022

### ABSTRACT

The four natural DNA bases (A, T, G and C) associate in base pairs (A = C and G ≡ C), allowing the attached DNA strands to assemble into the canonical double helix of DNA (or duplex-DNA, also known as B-DNA). The intrinsic supramolecular properties of nucleobases make other associations possible (such as base triplets or quartets), which thus translates into a diversity of DNA structures beyond B-DNA. To date, the alphabet of DNA structures is ripe with approximately 20 letters (from A- to Z-DNA); however, only a few of them are being considered as key players in cell biology and, by extension, valuable targets for chemical biology intervention. In the present review, we summarise what is known about alternative DNA structures (what are they? When, where and how do they fold?) and proceed to discuss further about those considered nowadays as valuable therapeutic targets. We discuss in more detail the molecular tools (ligands) that have been recently developed to target these structures, particularly the three- and four-way DNA junctions, in order to intervene in the biological processes where they are involved. This new and stimulating chemical biology playground allows for devising innovative strategies to fight against genetic diseases.

### INTRODUCTION

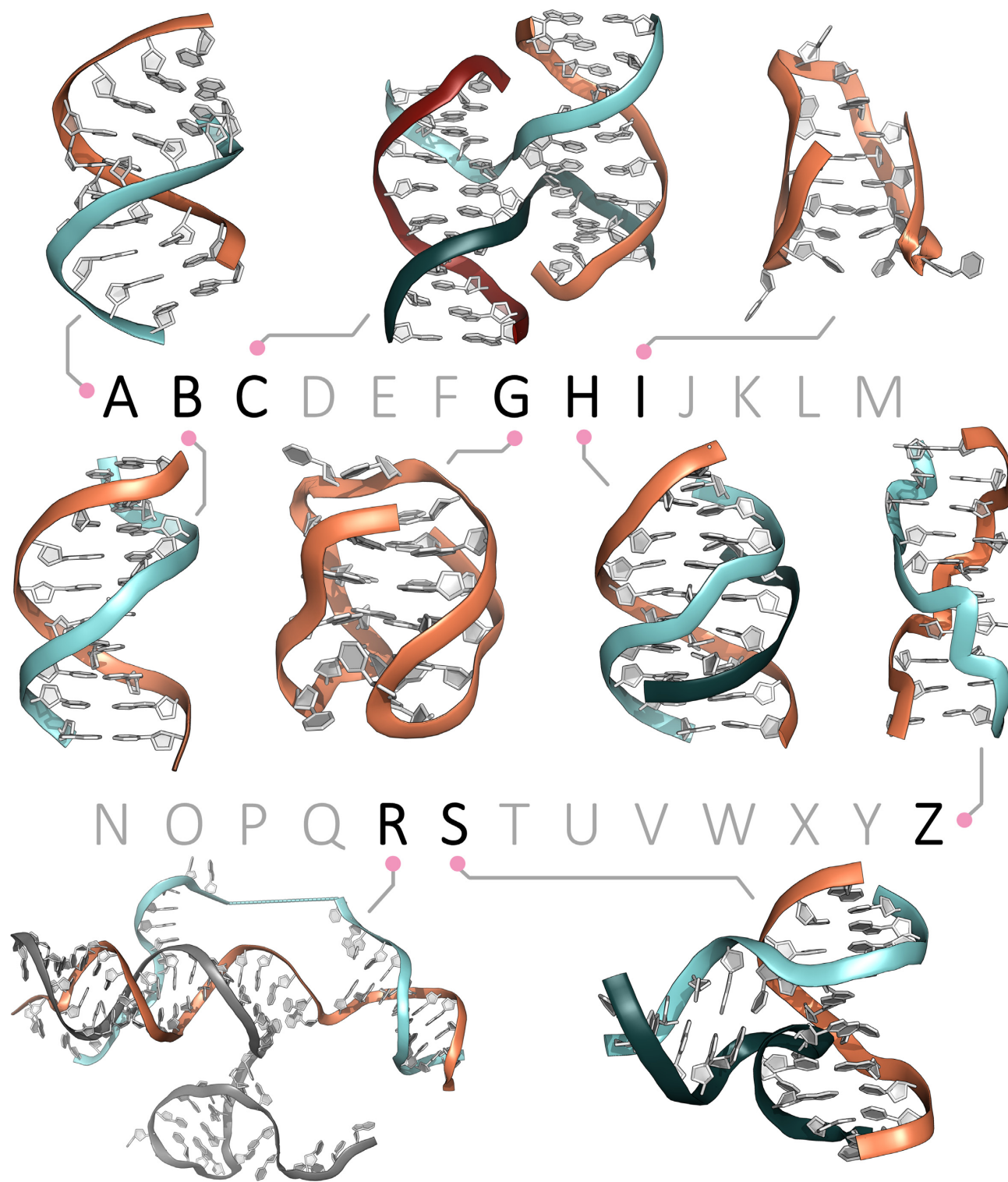
The DNA alphabet is naturally restricted to four letters, i.e. A for adenine, C for cytosine, G for guanine and T for thymine (1,2) (although some creative scientists have achieved the juggling act to expand it to 6 in living semisynthetic organisms (3,4), and then 8 (5), mimicking what was discovered in bacteriophages >4 decades ago) (6–9). In contrast, the DNA structure alphabet is far richer, with >20 letters used to date as descriptors of secondary structures

(10). The canonical, so called Watson–Crick, structure is referred to as B-DNA (as the X-ray crystallographic structure was obtained by Rosalind Franklin after hydration of a first sample ('A') of high-quality DNA provided by Rudolf Signer to Maurice Wilkins, consequently termed 'B') (11–14). Every other structure has consequently been categorized as a non-B-DNA structure, spanning from A-DNA (thus, dehydrated duplex) (11) to Z-DNA (duplex of inverted helicity, Z for zigzag) (10,15).

The central dogma of biology, heralded by Francis Crick in 1957 (16,17), has placed the B-DNA at the very centre of all molecular biology efforts invested, and discoveries made, after the elucidation of its double helix structure (2,13,18). This has kept the limelight away from reports published in succeeding years on the ability of DNA to fold into a variety of non-B-DNA structures (Figure 1), including: the triple helix (or triplex (19), termed H-DNA given its homopurine (hPu)/homopyrimidine (hPy) nature, also referring to hinged DNA) (20), first identified in RNA in 1957 (21) before being characterized in DNA in 1979; (22) the G-quartet in 1962 (23), the constitutive unit of the quadruple helix G-quadruplex (G4-DNA, or G-DNA) whose formation was demonstrated in 1988; (24) the tetra-stranded four-way DNA junction proposed as a model to explain gene conversion in 1964 by Robin Holliday (consequently called the Holliday junction) (25), predicted in 1966 (26) and demonstrated *in vitro* (and termed cruciform DNA, C-DNA) in the early 80s; (27,28) the Z-DNA first detected in 1967 (as a B-DNA of inverted, left-handed helicity) (29) before being firmly confirmed in 1979 (30), etc.

The topological diversity of DNA stems from supramolecular chemistry considerations: nucleobases (A, C, G, T) associate through the formation of hydrogen bonds (H-bonds), two in the A = C base pair, three in the G ≡ C base pair, allowing for a dynamic assembly/disassembly without substantial energy penalty. However, nucleobases are not exclusive in their H-bond-mediated association, and >20 different pairing modes are possible involving two of the four letters (31), offering a

\*To whom correspondence should be addressed. Tel: +33 380 399 043; Email: david.monchaud@u-bourgogne.fr  
Correspondence may also be addressed to Christine J. Cardin. Tel: +44 118 378 8215; Email: c.j.cardin@reading.ac.uk



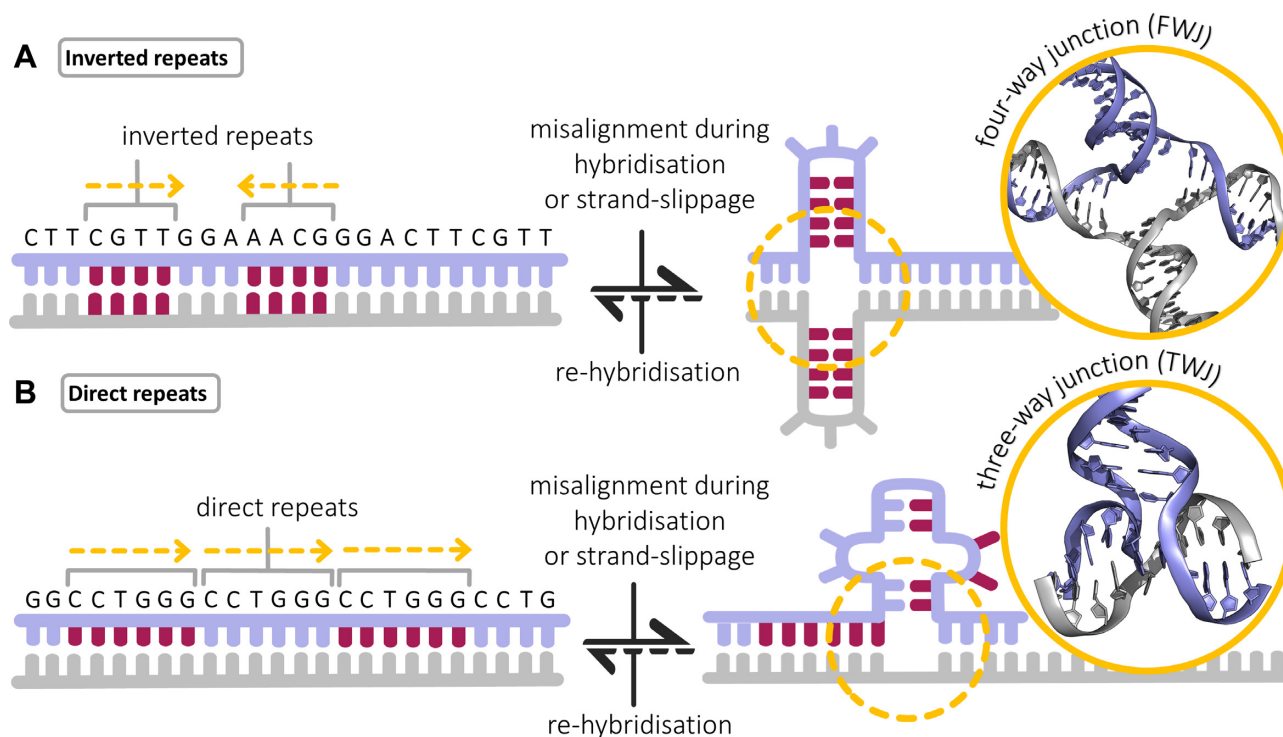
**Figure 1.** The DNA alphabet including the canonical B-DNA (or Watson–Crick duplex) along with a series of non-canonical structures including A-DNA (dehydrated duplex DNA), C-DNA (cruciform DNA, or four-way DNA junction), G-DNA (G-quadruplex DNA), H-DNA (for hinged DNA, or triplex DNA), I-DNA (i-motif), R-DNA (R-loop), S-DNA (for slipped DNA, or three-way DNA junction) and Z-DNA (for zigzag DNA).

first degree of topological diversity (base pairing). Also, nucleobases can donate/accept H-bonds through both their Watson–Crick (2) and Hoogsteen (32,33) faces, which are not mutually exclusive and thus permit the formation of base triplets (34) and quartets (23), offering a second degree of topological diversity (strandness). On the basis of the above, the structural pluralism of nucleic acids is anything but surprising.

The reason why these alternative structures have been overlooked for years is that their existence in a cellular context is challenged by the chromatin environment, in which the dense DNA packaging massively favours B-DNA. This has made the demonstration of their existence in *in vivo*-like conditions daunting, and the chemical biology tools required to do so have reached the necessary degree of maturity only recently (e.g. the visualisation of G4s in fixed cells by the antibody BG4 in 2013 (35), in live cells by the small molecule SiR-PyPDS in 2020) (36). Also, the rules that dictate their formation in the genome are poorly understood even if the sequences these structures fold from (i.e. repeated sequences) are known to be both abundant and widespread in the human genome (repetitive DNA elements cover roughly half of our genome) (37,38). Non-B-DNA-prone sequences can be inverted repeats (IR, involved in the formation of four-way DNA junction (FWJ), or cruciform DNA, Figures 1 and 2A), direct repeats (DR, involved in the formation of three-way DNA junction (TWJ), also termed slipped-strand DNA, slipped DNA or S-DNA (39), Figures 1 and 2B), mirror repeats (MR, involved in the formation of H-DNA) (38,40,41), or short tandem repeats (STR, involved in the formation of hairpin DNA, G4-DNA and i-motif (iM or I-DNA, Figure 1) depending of the repeated motif) (42). Thanks to recent computational analyses of an entirely sequenced human genome, the abundance of the non-B-DNA-prone sequences can be calculated: it is established that the least abundant motif is MR, with ca. 70 000 occurrences (ca. 2/100 kb), while the most abundant motif, IR, has >6 000 000 occurrences (ca. 206/100 kb) (41). However, despite the abundance of sequences that could give rise to non-B-DNA structures, the timing and kinetics of their folding are still poorly understood. Also, this folding requires the involved sequences to be freed from the duplex constraint (i.e. open chromatin, damaged DNA), their formation is thus transient only and subjected to a permanent and careful surveillance by *ad hoc* enzymes (e.g. helicases) (43), which make their formation even less likely. This thus explains why non-B-DNA structures have not been considered as reliable genetic elements (and targets) for years. Despite all these constraints, some of these structures (G4s, iMs) have been isolated and identified by sequencing methods. Efforts invested *in vitro* (that is, with purified DNA fragments) confirmed their widespread formation in the human genome (with >500 000 G4- (44) and iM-forming (45) sequences), but not only (46) (e.g. >25 000 G4- (47) and iM-forming (48) sequences in plants) (49). Investigations performed *in vivo* distinctly highlighted the suppressive role of chromatin for non-B-DNA structure formation as only ca. 1% of the G4s detected *in vitro* (G4-seq) (44) were detected *in vivo* (G4 ChIP-seq) (50). More globally, a combination of footprinting experiments (KMnO<sub>4</sub>/S1 nuclease) and genome-wide sequencing (ssDNA-seq) allowed to link

ssDNA regions in mammalian cells with non-B-DNA structures (C-, G-, H- and Z-DNA, ca. 20 000 motifs each) and then, these structures with regions involved in gene expression regulation (51). Further combining nuclease cleavage (S1 and P1) and sequencing (S1-END-seq and P1-END-seq, respectively) revealed that C- and H-DNA formation (at dinucleotide (TA)<sub>n</sub> repeats and hPu/hPy repeats (e.g. (GAA)<sub>n</sub>), respectively) is both widespread in human cells but also strongly dependent on the cell status (cancerous versus non-cancerous cells) (52,53). These results, beyond confirming the existence on non-B-DNA structures *in vivo*, provide also a strong correlation between structure-prone sequences and both mutability and genetic instability (52) (often referred to as to RIM, for repeat-induced mutagenesis) (54,55). Indeed, their widespread and nonuniform distribution across our genome is significantly enriched in regulatory regions, and the distribution quite often overlaps with that of reported hotspots, i.e. regions prone to undergo breakage, deletions and translocations (38,41,55,56). This explains why repeated sequences are carefully patrolled and tightly controlled by genome surveillance systems, notably the DNA damage response (DDR) machinery (57–59).

As indicated above, to adopt higher-order structures, these sequences must be relieved from their duplex constraint and reannealed with misalignment or slippage. Their folding is thus dependent on, and coupled with, DNA transactions (transcription, replication) and repair. This makes them interesting targets for therapeutic approaches aimed at inflicting damage to highly active, that is, rapidly dividing cells. Indeed, once folded, these thermodynamically stable structures might act as roadblocks to polymerases, pausing and/or stalling their processivity, which is recognized as a situation of crisis (DNA damage) (60–63). It is therefore unsurprising that chemical biologists soon envisioned an original strategy in which the transient stabilization of non-B-DNA structures by small molecules (so called ligands) could be exploited to foster this situation of crisis notably in rapidly dividing, that is, cancer cells (62,64). The relevance of this approach is further substantiated by the fact that cancer cells are generally DDR-impaired (57), which makes them more sensitive to DNA damage-inducing agents than healthy cells. This explains why many of the therapeutics currently used in the clinic damage DNA, although through different modalities (65). However, these therapeutics target B-DNA (or their associated proteins such as topoisomerases) but none of them target non-B-DNA structures. This is clearly the major caveat of this field, which deserves to be addressed soon in order to lend credence to its strategic relevance. The most advanced molecule, the G4-stabilizer CX-5461 is currently in phase I (clinicaltrials.gov NCT02719977) (66) against advanced solid tumours (its parent molecule CX-3543, also known as Quarfloxin, was stopped after phase I). Without a successful example of non-B-DNA targeting agent in the clinic, this field will keep on suffering from a lack of legitimacy. Yet, in contrary to B-DNA, which offers poorly defined binding sites only (i.e. minor and major grooves, intercalation in between 2 bp), non-B-DNA structures offer a broad variety of structurally well-defined ligand binding sites, which makes highly selective targeting with small molecules possible. Chief among them are the G4-DNA (67–69), which display two accessi-



**Figure 2.** Schematic representation of junction-forming sequences, i.e. inverted repeats that can give rise to a four-way DNA junction (A) or direct repeats that can fold into a three-way DNA junction (B).

ble G-quartets surrounded by flexible loops, and DNA junctions (70–72), in which the junction point between the duplex arms (3 duplex arms for TWJ; four arms for FWJ) creates a cavity prone to welcome small molecules (*vide infra*).

The recent developments described in this review will be focused on the targeting of DNA junctions, as that of G4-DNA is regularly covered by authoritative reviews (68,69,73–77) to which interested readers are invited to refer. These developments offer this field a shining message of hope, as these new ligands allow for a specific targeting of DNA junctions (78), yet in *in vitro* and cell-based assays only to date, which bears significant potential for delivering soon a new generation non-B-DNA targeting therapeutics for which the demonstration of clinical efficacy is greatly expected.

## STRUCTURE, FUNCTIONS AND TARGETING OF DNA JUNCTIONS

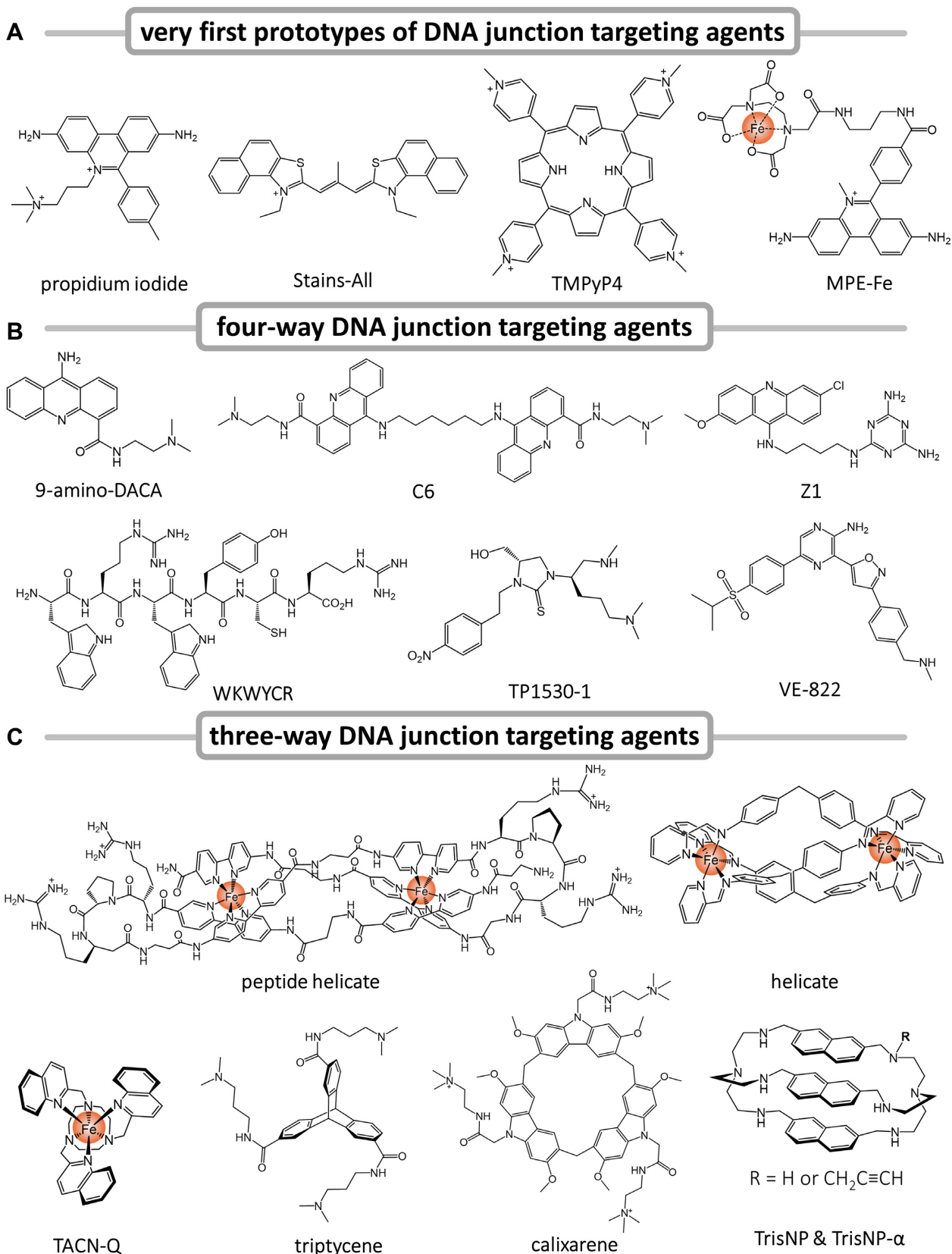
### The early days of DNA junction targeting

The field of non-B-DNA structures and related ligands was undoubtedly pioneered by Neville Kallenbach (72,79). In a series of articles published in the mid 80's-mid 90's, he provided both biophysical characterizations (e.g. by calorimetry (80), gel electrophoresis (81), chemical footprinting (82,83), etc.) of DNA junctions (but also of G4-DNA) (84,85) and the first insights into how small molecules (Figure 3A), mostly fluorescent dyes, i.e. the DNA labelling agent propidium iodide (86), the cyanine stains-all (87), the porphyrin TMPyP4 (88), but also cleaving agents such as methidium-propyl-EDTA-Fe(II) (MPE-Fe) (89), inter-

act with them (and with G4s) (90). Of course, these studies were limited to *in vitro* investigations and the selectivity for DNA junctions over B-DNA, of utmost importance for chemical biology and medicinal chemistry, was not investigated. Anyway, these studies were instrumental in that they spurred on research aiming at ultimately investigating the cellular effects on these new genetic targets and molecular tools (further discussed below).

### Targeting four-way DNA junctions

As indicated above, the existence of a four-stranded DNA structure was postulated by Robin Holliday in 1964 as an intermediate in homologous recombination (HR) (25,91). The possible formation of a central four-way junction intermediate (a 'chiasma') had already been discussed by Joseph G. Gall in 1954 (92), as a basis for the crossing-over mechanism, i.e. the reciprocal exchange of segments along pairs of homologous chromosomes. However, heroic efforts were needed to obtain a structural confirmation of the FWJ organization of this intermediate: various biochemical and physicochemical methods were implemented, including gel electrophoresis (93–95), fluorescence energy transfer (FET (96,97), then reported as FRET (98,99), for fluorescence resonance energy transfer), birefringence decay (100), and nuclear magnetic resonance (NMR) (101,102), along with some imaging techniques such as scanning electron microscopy (SEM) (103) and atomic force measurements (AFM) (104). Among these techniques, X-ray crystallography played a central role: (105) solid state analyses were indeed instrumental to solve the structure of FWJ, alone (106,107) or in interaction with enzymes in charge



**Figure 3.** (A) Chemical structures of the very first DNA junction targeting agents used by Kallenbach *et al.* (79), i.e. propidium iodide, Stains-all, TMPyP4 and MPE-Fe (counter-ions are removed for clarity). (B) FWJ ligands used for either *in vitro* studies, i.e. 9-amino-DACA (115), C6 (118) and Z1 (122), or for cell-based assays, i.e. WKWYCR (132), TP1530-1 (143) and VE-822 (146). (C) TWJ ligands including diiron helicate (171) and peptide helicate (177), the poly-aza-macrocycle TACN-Q (178) along with a triptycene (181), a calixarene (186), and two azacryptands TrisNP and TrisNP- $\alpha$  (156).

of the recombination (the resolvase RuvA (108,109) or the recombinase Cre (110), *vide infra*) (111) and then, small molecules (*vide infra*). All these techniques concurred in demonstrating the organisation of the FWJ, with its four arms that either extend from opposite sides of the central cavity, which is therefore wide open (the so-called open planar X-structure), or stack on each other two by two (the so-called stacked X-structure), which consequently closes the central cavity (Figure 4A) (70).

The initial junction structure was obtained while studying DNA mismatches, and a later study revealed that, in the crystal structure, junction formation is sequence dependent (112). These authors carried out a systematic study of the decamer sequences  $CCn_1n_2n_3N_1N_2N_3GG$ , where  $N$  can be any of the four common nucleotides, and  $n$  are specified accordingly to maintain the IR motif and thus self-complementarity of the sequences. From the 64 possible combinations, they concluded that the d(CCGGTACCGG) sequence gave the junction form under all conditions tested, whereas other sequences could be in equilibrium with A- or B-DNA forms. A way to control the conformation is the covalent binding of psoralen: (113) the closely related sequence d(CCGTAGCGG) was found to crystallise as a B-form duplex, but the binding of psoralen after laser irradiation at the TA/TA step led to the stabilisation of the junction (Figure 4B). More generally, the X-ray structures obtained in presence of enzymes and ligands showed that these effectors can modulate the shape of the FWJ in order to structurally optimise the ligand/DNA association (induced-fit).

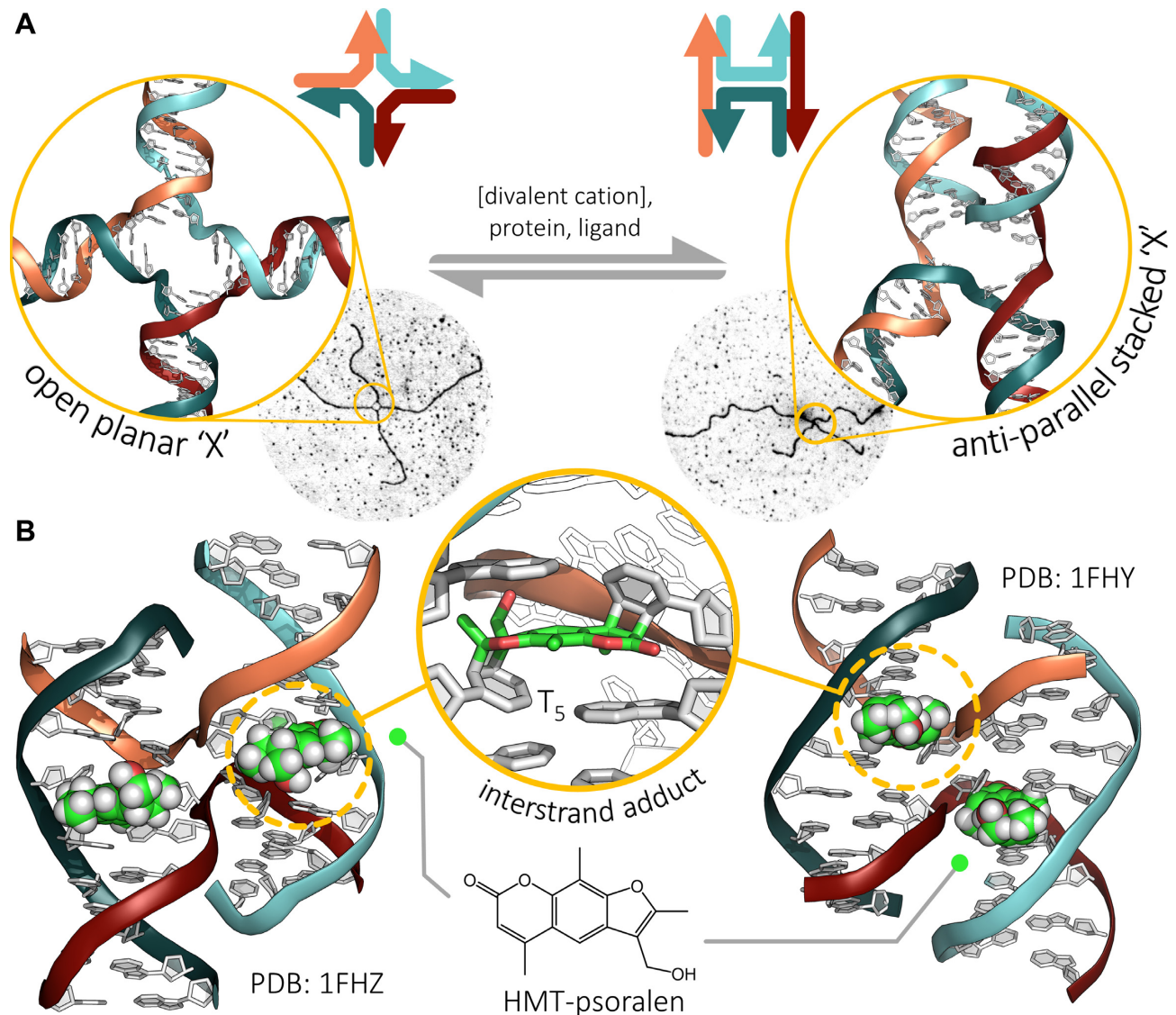
Regarding the way small molecules interact with FWJs (Figure 3B), the first insights were obtained through the ligand-induced dimerisation of a short, 6-nucleotide (6-nt) duplex. In this study (114), incubation of the self-complementary d(CGTACG) sequence, or the closely related d(CG[5-BrU]ACG), with acridines (e.g. DACA, 9-amino-DACA, Figures 3B and 5A) (114) and phenazine (115) derivatives provided an asymmetric unit in which the terminal C of one strand invades the second helix at the ligand binding site, therefore creating a cavity in which two acridines nestle in between four base pairs (Figure 5A). We thus designed dimeric derivatives of DACA (linked through 1,6-diaminohexane or 1,8-diaminooctane linkers, named C6 and C8, respectively, Figures 3B and 5B) as possible bis-intercalators and determined two types of X-ray structures with d(CGTACG): a structure in which, behaving similarly to its monomeric counterpart (see above), C8 induces a FWJ-like assembly on the basis of a terminal C exchange and creates a H-bond with the residue G<sub>6</sub> (Figure 5B); (116) and a structure in which C6 crosslinks two separate duplexes, each of its acridine units being intercalated in between two CG base pairs of a given duplex (Figure 5B) (117). Finally, the X-ray structure of C6 in interaction with a full FWJ resulting from the self-assembly of the d(CCGGTACCGG) sequence was successfully solved and refined (Figure 5B) (118). This clearly defined binding mode affirmed the suitability of the central cavity of FWJs to accommodate close contacts with dedicated ligands. In this example, the acridine units cause two A residues (A<sub>6</sub>) at the central TA step of the crossover strands to flip out to generate the ligand binding site in which both a long H-bond (be-

tween the DACA sidechain and the residue C<sub>7</sub>) and stacking interactions (in between a T and a C) with surrounding nucleobases are created. Later, it was demonstrated that DACA derivatives were also able to induce the formation of FWJs (119); however, these studies, though elegant and insightful, were limited to structural investigations only, owing to the known tendency of acridine derivatives to intercalate within duplexes (120). Their use in cell-based assays would have been troublesome because of their strong B-DNA interaction, which would have blurred the relevance of FWJs as therapeutic target: indeed, it would have been impossible to discriminate the origins of the cellular effects observed between those originating in B-DNA intercalation (the most likely event) and those originating in FWJ interaction (the less likely event). This is mostly due to the fact that these ligands do not target the unique structural component of FWJs (i.e. the central cavity), but instead bind to the duplex arms at the proximity of the cavity, which makes specific interaction unattainable. One possibility would have been to structurally fine-tune acridines to preclude intercalation in a manner reminiscent of what was done with G4 ligands, which led to the design of the 3,6,9-trisubstituted acridine BRACO-19 (121). Another possibility was to target the central cavity of FWJs; as discussed further below, this was done with a series of short peptides but without firm structural characterization of their FWJ binding mode.

As with the original Holliday junction structure, which was a serendipitous discovery while studying DNA mismatches, a recent mismatch study also uncovered junction formation (122). A potent DNA-binding compound triaminotriazine-acridine conjugate (Z1, Figure 3B) functions by targeting T:T mismatches in CTG trinucleotide repeats that are responsible for neurological diseases such as myotonic dystrophy type 1, but its binding mechanism remains unclear. The crystal structure of this ligand was solved in a complex with DNA containing three consecutive CTG repeats with three T:T mismatches. Surprisingly, direct intercalation of two Z1 molecules at both ends of the CTG repeats induced T base flipping (T<sub>4</sub>), a H-bond between Z1 and a T (T<sub>10</sub>) and DNA backbone deformation to form a four-way junction (Figure 5C). The core of the complex unexpectedly adopts a U-shaped head-to-head topology to form a crossover of each chain at the junction site. The crossover junction is held together by two stacked G:C pairs at the central core that rotate with respect to each other in an X-shape to form two nonplanar minor-groove-aligned G:C:G:C tetrads. Two stacked G:C pairs on both sides of the central core are involved in the formation of pseudo-continuous duplex DNA. However, and again, these studies were limited to structural investigations only owing to the lack of information concerning the specificity of Z1 for FWJs that, combined with the known health-threatening nature of triazines (notably widely used as herbicides) (123), is not the best guarantee of a bright future for this series of compounds.

### The relevance of FWJ in cells

Although the central position of HJs in HR has been firmly established for decades (91), it is only recently that they were visualized in cells (124). While HR is only one of two major



**Figure 4.** (A) The conformational plurality of a four-way DNA junction ('open' versus 'stacked' junction) that can be modulated by external mediators including cations (e.g.  $Mg^{2+}$ ), enzymes (e.g. RuvA or Cre) and ligands (e.g. the peptide WKHYNY). SEM images from (103). (B) X-ray crystal structures (PDB: 1FHZ and 1FHY) (113) showing that the photochemical reaction of psoralen with the d(CCGTAGCGG) sequence led to a stable four-way DNA junction (1FHY), whereas the native sequence crystallises as a B-DNA. The sequence d(CCGGTACCGG) gives a four-way junction both in native conditions and after reaction with the psoralen (1FHZ).

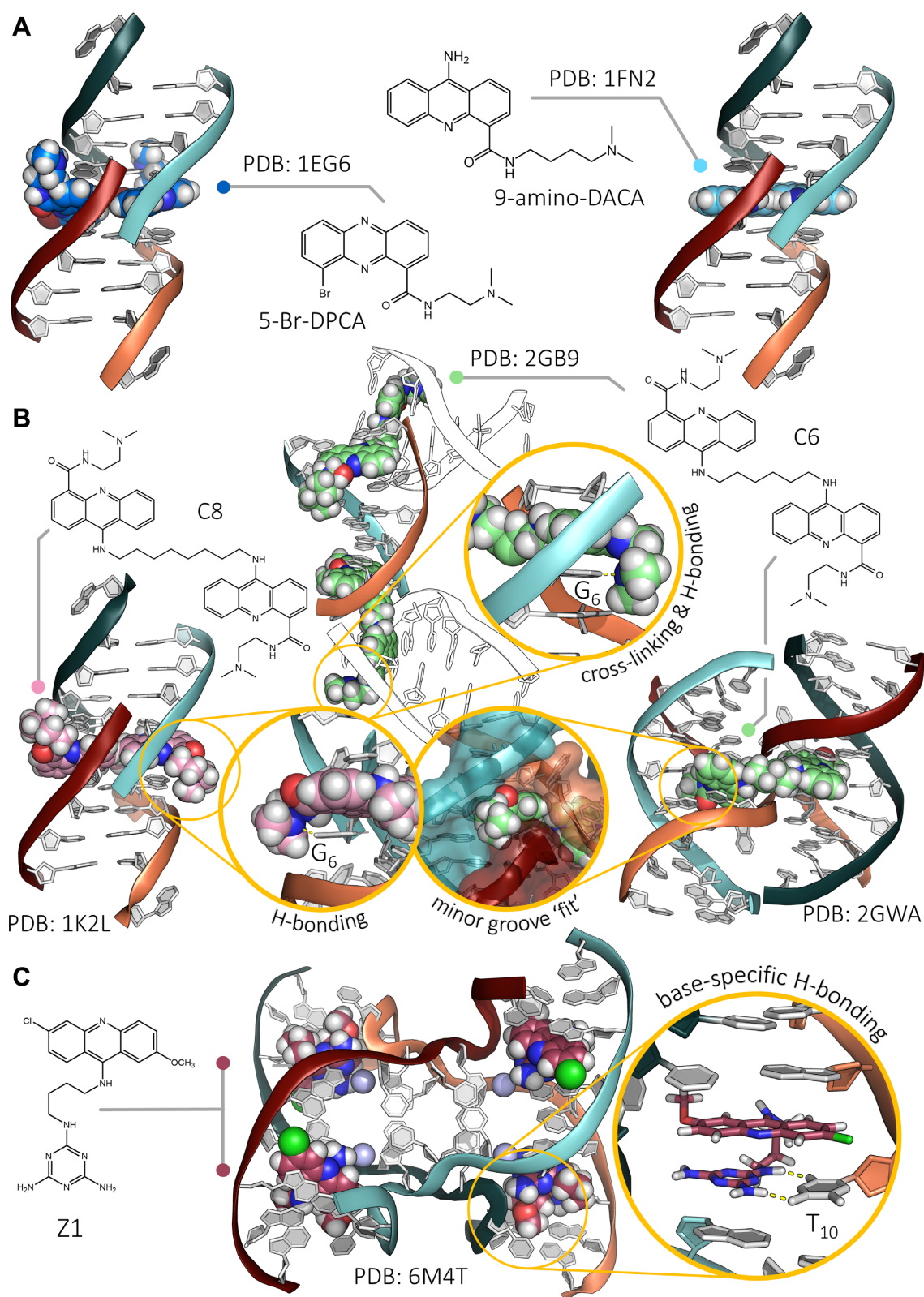
pathways leading to double strand DNA break (DSB) repair in mammalian cells, along with non-homologous end joining (NHEJ) (Figure 6) (125), it is indispensable in lower organisms (bacteria, yeasts) for ensuring DNA transfer and adapting to evolution (natural selection) (126). It is therefore unsurprising that a vast body of research has been dedicated to understanding the roles of HJ/FWJ in yeasts and bacteria.

These efforts have led to the discovery of the RuvABC complex that operates at HJs in prokaryotes (*Escherichia coli*) in a sequential manner: RuvA binds to HJs and targets RuvB to the junction, both RuvA and RuvB (RuvAB) promote branch migration and RuvC resolves the junction (127). An engineered synthetic protein RuvCDefGFP was recently used to investigate HJs in cells

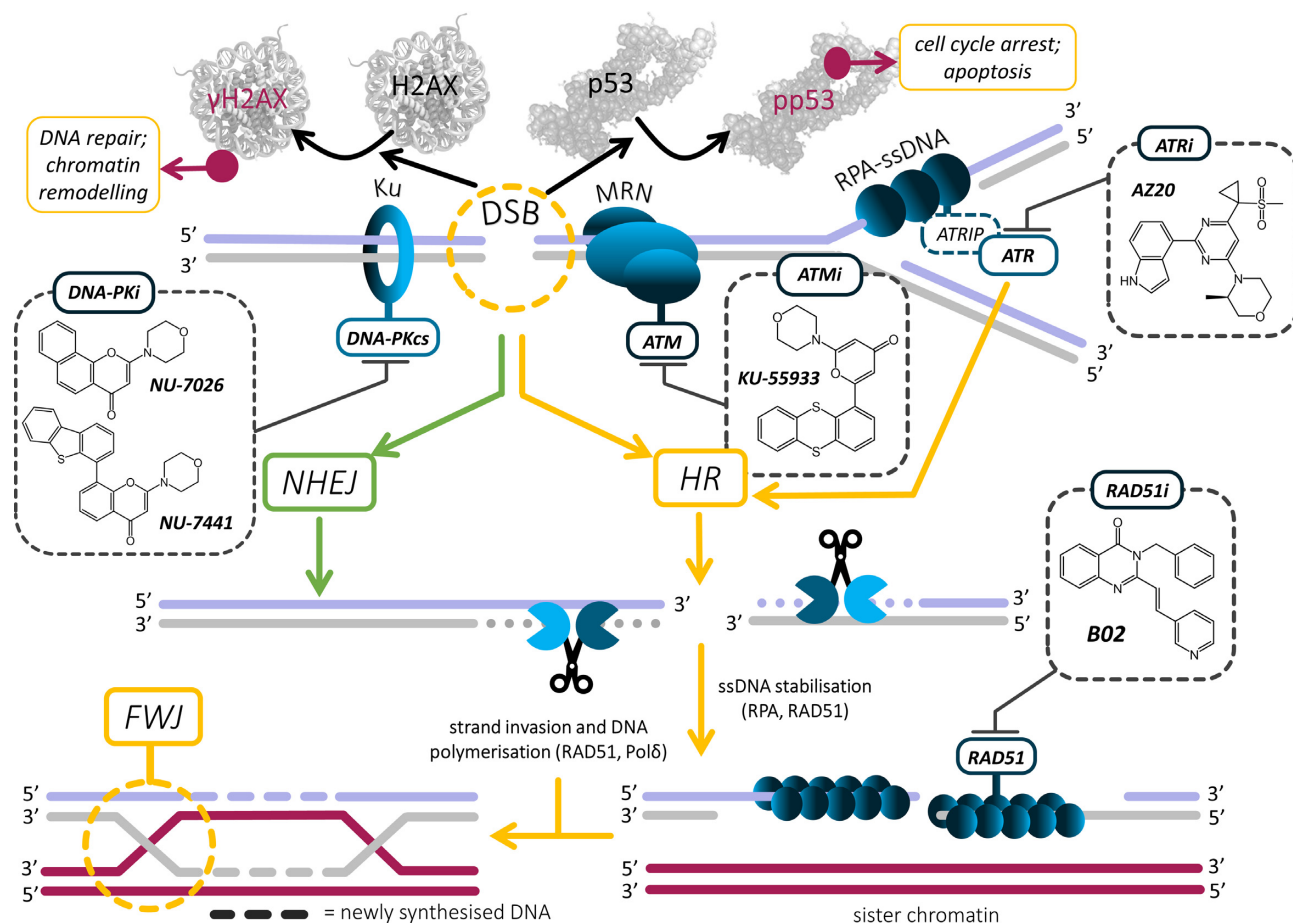
(128). RuvCDefGFP is a catalytically defective (Def) RuvC fused to green fluorescent protein (GFP); it was used to both map HJs in *E. coli* by HJ-ChIP-seq (using the anti-RuvC antibody), which demonstrated the accumulation of HJs near DSB sites, and visualize HJs *via* live-cell imaging, with a distribution correlated with homology-directed DSB repair, notably at single-stranded gaps.

These compelling results thus lend further credence to value of targeting HJs in bacteria by small molecules. The short hexapeptides developed by Anca M. Segall and coworkers were among the first reported HJ-ligands. It was first demonstrated that WKHYNY (Trp-Lys-His-Tyr-Asn-Tyr) target HJs *in vitro* and inhibits its resolution by the recombination proteins integration host factor (IHF, a heterodimeric protein that binds to and bends DNA) and the





**Figure 5.** (A) X-ray crystal structures (PDB: 1EG6 (115) and 1FN2 (114)) of quasi-junction cavities obtained with the d(GCTACG) sequence in interaction with phenazine or acridine derivatives. (B) X-ray crystal structures (PDB: 1K2L (116), 2GB9 (117) and 2GWA (118)) of DNA junction in interaction with acridine dimers. (C) X-ray crystal structure (PDB: 6M4T) of a DNA junction in interaction with the triazine Z1 (122). The nucleobases involved in ligand/FWJ interactions are identified in the zooms.



**Figure 6.** Double-strand breaks (DSBs) and stalled replication forks (RFs) are repaired by homologous recombination (HR) and non-homologous end joining (NHEJ) mechanisms, mediated by DNA-PK, ATM and ATR kinases, and RAD51. DSBs signalling involves the phosphorylation of the histone H2AX (Ser139) and the protein p53 (Ser15). The various inhibitors of these DNA repair pathways (NU-7026, NU-7441 for DNA-PK; KU-55933 for ATM; AZ20 for ATR; and B02 for RAD51) are highlighted, along with the FWJ intermediate of the HR. Figure adapted from (156).

excisionase (Xis) (129). Next, it was shown by X-ray crystallography that WKHYNY does indeed bind to the central cavity of HJs (loxP sequence) (130) when a recombinase Cre tetramer maintains the HJ in its open form (110). The exact position of the peptide was not accurately determined, but found to be distributed over two major positions located on both sides of the C2-symmetrical cavity. Interestingly, the presence of the peptide modifies the structural organization of the HJ, which makes it less prone to be enzymatically resolved, thus providing some interesting insights into the mechanism by which hexapeptides inhibit HJ resolution.

This blocking behaviour was further investigated with two other peptides, WRWYCR (Trp-Arg-Trp-Tyr-Cys-Arg, Figure 3B) and KWCCRW (Lys-Trp-Trp-Cys-Arg-Trp), whose ability to inhibit both RecG helicase (which binds to and unwinds forked DNA) and RuvABC resolvase (which displays DNA junction specific helicase-endonuclease activity) in *E. coli* (131) was scrutinised (132). Their efficiency was quite high (e.g. with  $IC_{50}$  values  $>100 \mu\text{M}$  for WKHYNY versus 5–20 nM for both WRWYCR and KWCCRW against RecG; 0.06  $\mu\text{M}$  for WRWYCR against RuvABC) but the exact mechanism remains unclear, as peptides do not prevent RuvA tetramer from binding to the

HJ, RuvB does not interact with the central cavity of HJ *per se* and the cleavage activity of RuvC alone is marginally affected by the presence of the ligand. However, the peptides inhibit the resolvase activity of RuvC when embedded in the RuvABC complex, probably since this complex opens wide the cavity and makes it accessible to the peptides, which then blocks its enzymatic resolution for steric reasons. The most active peptide, WRWYCR, was indeed found more- and very- active against Gram<sup>+</sup> bacteria (e.g. *S. aureus*, with minimal inhibitory concentration (MIC) between 4 and 32  $\mu\text{g/ml}$ ) than against Gram<sup>-</sup> bacteria (e.g. *E. coli*, MIC between 32 and 64  $\mu\text{g/ml}$ ) (133). Of note, the presence of a cysteine in the most potent peptides implies that their active form is the disulfide-linked dimer: this was first demonstrated by the loss of activity upon Ala scan or addition of reducing agents (dithiothreitol, or DTT) (134) and by the design and use of the disulfide-independent dodecamer WRWYRGGRYWRW, found to be quite active (135). A series of fluorescence investigations performed with 2-aminopurines (2-APs) located at different positions around the central cavity of the FWJ along with molecular modelling studies confirmed a binding mode based on two hexapeptides interacting on both sides of the FWJ cav-

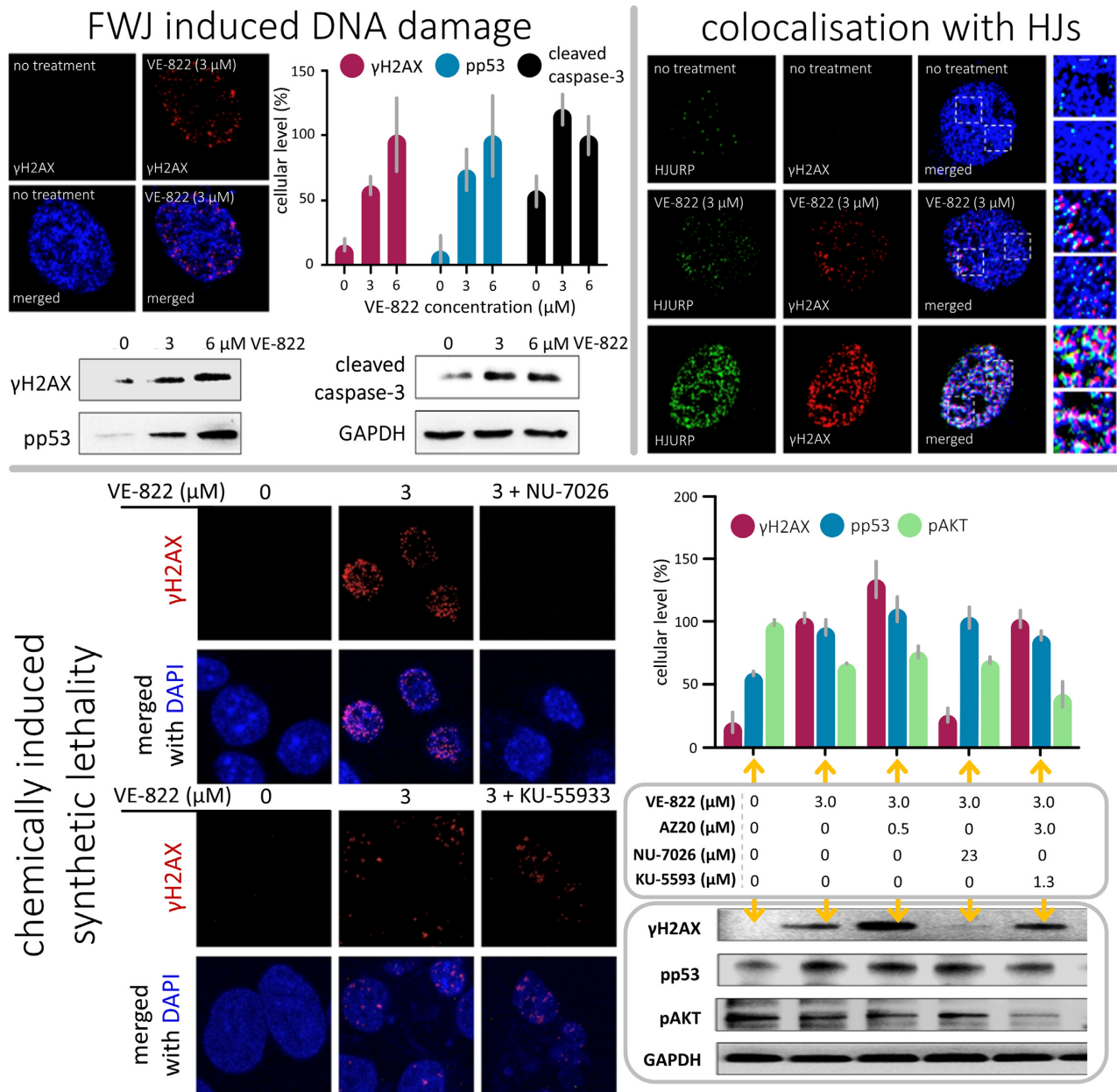
ity (136). Attempts to favour these interactions using cyclic hexa- and octa-peptides were made: while the synthetic challenge was addressed (137,138), the macrocyclopeptides were not found to be very biologically active.

Given the central position of HJs in HR, it is unsurprising that HJ-trapping peptides have also been studied for their ability to impair recombination-based DNA repair, in both prokaryotes and eukaryotes. This was first scrutinized in bacteria, and it was found that the two most potent inhibitors (WRWYCR and KWWRW), which bind to free HJs, display broad-spectrum bactericidal activity, notably against Gram<sup>+</sup> bacteria (while the less potent WKHYY, which binds to protein-bound HJs only, is not active) (133). The treatment of bacteria with peptides triggered DNA filamentation and segregation abnormalities, along with DNA break accumulation, in line with what is observed in cells undergoing severe DNA damage. The synergistic effects observed between peptide treatments and DNA damaging events (incubation with the cross-linking agent mitomycin C, or upon UV irradiation) show that DNA damage create more targets to which peptides bind (DNA lesions and hence repair); this delineates a novel antibacterial strategy based on HJ-trapping-mediated inhibition of recombination-dependent DNA repair. These studies were then extended to eukaryotic cells, particularly to human tumour cells (e.g. cervical cancer HeLa cells, prostate cancer PC3 cells) (139). The ability of WRWYCR to trigger DNA damage was evidenced by the immunodetection of both the histone H2AX phosphorylated on its serine 139 ( $\gamma$ H2AX) (140) (Figure 6) and p53-binding protein 1 (53BP1), well-established markers of DSBs (141), along with the terminal dUTP transferase nick-end labelling (TUNEL) assay, likely due to the accumulation of unresolved DNA repair intermediates that leads to DNA breaks. Their toxicity was also potentiated by DNA damaging agents such as doxorubicin and etoposide, thus offering a new application to HJ-targeting agents as potential chemotherapeutics (either as standalone agents or in synthetic lethality cocktails). However, these studies were limited by the lack of information regarding the actual cellular specificity of these peptides for FWJs. Indeed, linear and cyclic short peptides are routine in medicinal chemistry (being used to fight against bacterial and viral infection, cancers and neuropathologies) (142), implying a broad variety of cellular targets. They also suffer from known limitations (conformational pluralism, limited bioavailability, short half-life *in vivo*, etc.) that, along with a low target selectivity, might limit and/or prevent their clinical use. None of these shortcomings were discussed here as these studies were intended to provide a proof-of-concept that targeting FWJs with small molecules can produce interesting cellular outcomes.

A way to tackle them is to use small molecules instead of peptides. To this end, Segall *et al.* also invested efforts to identify small molecules able to trap HJs and inhibit HJ resolution. Screening using the chemical libraries of the Torrey Pines Institute (TPI) led to the identification of a *N*-methyl aminocyclic thiourea referred to as TPI1530-1 (Figure 3B) as a potent inhibitor of RecG for instance ( $IC_{50} = 0.85$  versus  $0.12 \mu\text{g}/\text{mL}$  for WRWYCR). However, this compound was found inactive against bacterial growth, likely due to its poor bioavailability (143). A series of pyrrolidine bicyclic

guanidine derivatives (TPI1609-1, -3, -10 and -12) were also identified, still *via* HTS-screens (144). Among them, TPI1609-10 displayed a still lower affinity for HJ than the hexapeptide WRWYCR ( $K_d = 300$  versus  $12.5 \text{ nM}$ ; but far higher than TPI1530-1,  $K_d = 11.8 \mu\text{M}$ ), along with a lower RecG inhibitory activity ( $IC_{50} = 0.94 \mu\text{g}/\text{mL}$ ), but elicited comparable, sometimes better MICs against Gram<sup>+/−</sup> bacteria than WRWYCR (e.g. 16 versus 32–64  $\mu\text{g}/\text{mL}$  against *E. coli*, respectively). These results thus provide strong support for the hypothesis that low-molecular weight molecules can indeed target HJs, inhibit their resolution and potentially hamper bacterial growth. Unfortunately, no information is provided about the way these molecules interact with FWJs, and about their selectivity for FWJs versus other structures of DNA.

As indicated above, FWJ ligands have mostly been used to fight bacterial infections. But not only, with the DNA-damaging properties of WRWYCR also investigated in cancer PC3 and HeLa cells (139) and the antiproliferative activity of WRWYRGGRYWRW assessed in colorectal cancer HCT116 cells (135). To find a new FWJ-targeting scaffold, Searcey *et al.* developed an isothermal assay based on polyacrylamide gel electrophoresis (PAGE); this assay aimed at identifying small molecules able to trigger FWJ assembly from the four separated strands (119). It was initially implemented to screen a series of acridine-based ligands, whose chemical diversity was ensured by click chemistry (145); the resulting derivatives displayed only modest antiproliferative properties against human leukaemia HL60 cells, probably due to a poor cellular uptake, and no mechanistic rationale was provided about the origin of this cellular activity. Again, the demonstration of the specificity for FWJs of these molecules was lacking and their interaction (intercalation) with B-DNA not investigated; the authors also indicated a possible interaction with G4, thus complicating the interpretation of cell-based results. However, the goal of this study was clearly the development of a convenient *in vitro* assay to screen possible FWJ ligand candidates. It was incidentally implemented in an independent manner by Qikun Yin, Hongbo Wang and co-workers, which led to the discovery of a new FWJ-binding agent, VE-822 (Figure 3B) (146). This small molecule efficiently promotes FWJ assembly assessed by PAGE ( $EC_{50} = 7.6 \mu\text{M}$ ) and confirmed by isothermal FRET assay ( $EC_{50} = 5.4 \mu\text{M}$ ). It binds to FWJs with a high affinity ( $K_d = 8.6 \mu\text{M}$ ) and a good selectivity over both dsDNA and ssDNA ( $K_d > 50 \mu\text{M}$ ). In osteosarcoma (U2OS) cells, VE-822 was found to efficiently trigger DNA damage ( $\gamma$ H2AX induction (140) along with phosphorylation of P53 on its serine 15 (pp53) (147,148), and of AKT on its serine 473 (pAKT) (149), Figures 6 and 7), which leads to cellular apoptosis (see the apoptosis marker cleaved caspase-3). DNA damage is mitigated by both the overexpression of the junction resolving enzyme BLM and cell pre-treatment with the DNA-PK inhibitor NU-7026 (150) (more modestly with ATR inhibitor AZ20 and ATM inhibitor KU-55933, Figures 6 and 7). These results highlight the role of DNA-PKcs (crucial for DSB signalling) (59) in VE-822-induced DNA damage sensing. To provide an unambiguous demonstration that DNA damage occurs at FWJ sites, co-immunodetection experiments were performed to colocalize  $\gamma$ H2AX sites with the Holli-



**Figure 7.** Cellular activity of the FWJ-ligand VE-822 that triggers extensive DNA damage ( $\gamma$ H2AX and pp53 markers) eventually leading to apoptosis (cleaved caspase-3 marker), which colocalizes with Holliday junctions in U2OS cells (labelled using Holliday junction recognition protein, HJURP) and can be modulated by inhibitors of DNA repair (including AZ20, NU-7026 and KU-55933, which inhibit ATR, DNA-PKcs and ATM, respectively). Figure adapted from (146).

day Junction Recognition Protein (HJURP, Figure 7) (151). Not only were common *foci* detected, but also their abundance was found to be modulated by either VE-822 (increase) or BLM overexpression (decrease), which thus establishes a direct link between ligand-stabilized FWJs and DNA damage. VE-822 is rather active against osteosarcoma (U2OS) and glioblastoma (U251) cells ( $IC_{50} \sim 6 \mu M$ ), more active than in healthy cells (e.g. lung fibroblast (WI38) and hepatocyte (HL7702) cells, with  $IC_{50} > 20 \mu M$ ), which thus opens an interesting therapeutic window. Also, it synergistically interacts with the topoisomerase 2 (Top2) poison doxorubicin, an activity that can be modulated by either

pre-incubation with NU-7026 or BLM over-expression (as above). Though comprehensive from a molecular and cellular biology point of view, this study suffers from the fact that VE-822 was initially developed as an ATR inhibitor (152). ATR is known, among other things, to prevent replication fork collapse and DSBs formation, and orchestrate DSB repair by the homologous recombination repair (HRR) pathway (153,154). Therefore, the inhibition of ATR by itself is expected to lead also to an accumulation of DNA damage markers (including  $\gamma$ H2AX) and inhibit cancer cell growth, thereby casting doubt as to the actual origin(s) of the cellular effects observed here. However, this study is interesting

in that it provides insights into FWJ biology beyond their role in DNA repair: the abnormal stabilisation of FWJs can indeed trigger DNA damage, likely as a result of the ability of FWJ/ligand complexes to hamper proper motion of polymerases along the genomic duplex. This role of road-blocking has been thoroughly studied in regards to TWJs (and will be thoroughly discussed below).

### Targeting three-way DNA junctions

In the examples described above, the interaction of small molecules with FWJs (or HJs) was intended to either impair DNA repair (their cellular activity being potentiated by DNA damaging agents such as doxorubicin) or trigger DNA damage (their cellular activity being modulated by DDR inhibitors). The situation is simpler with another class of branched DNA, the TWJs, as their stabilization is intended to trigger genetic instability by stabilizing physical obstacles that impair proper DNA-related enzymes processivity (such as polymerases and topoisomerases, involved in replication, repair and transcription). Indeed, junctions might arise ahead of the enzymatic complexes in charge of DNA transactions as a result of torsional stress originating in the motion of enzymes along the duplex. Stabilised TWJs might thus act as physical roadblocks to polymerases, which is recognized and coped with as a form of DNA damage by the cellular surveillance machinery (60–62). Therefore, the cellular activity of TWJ-targeting agents could be potentiated by DNA repair inhibitors (*vide infra*) (155,156).

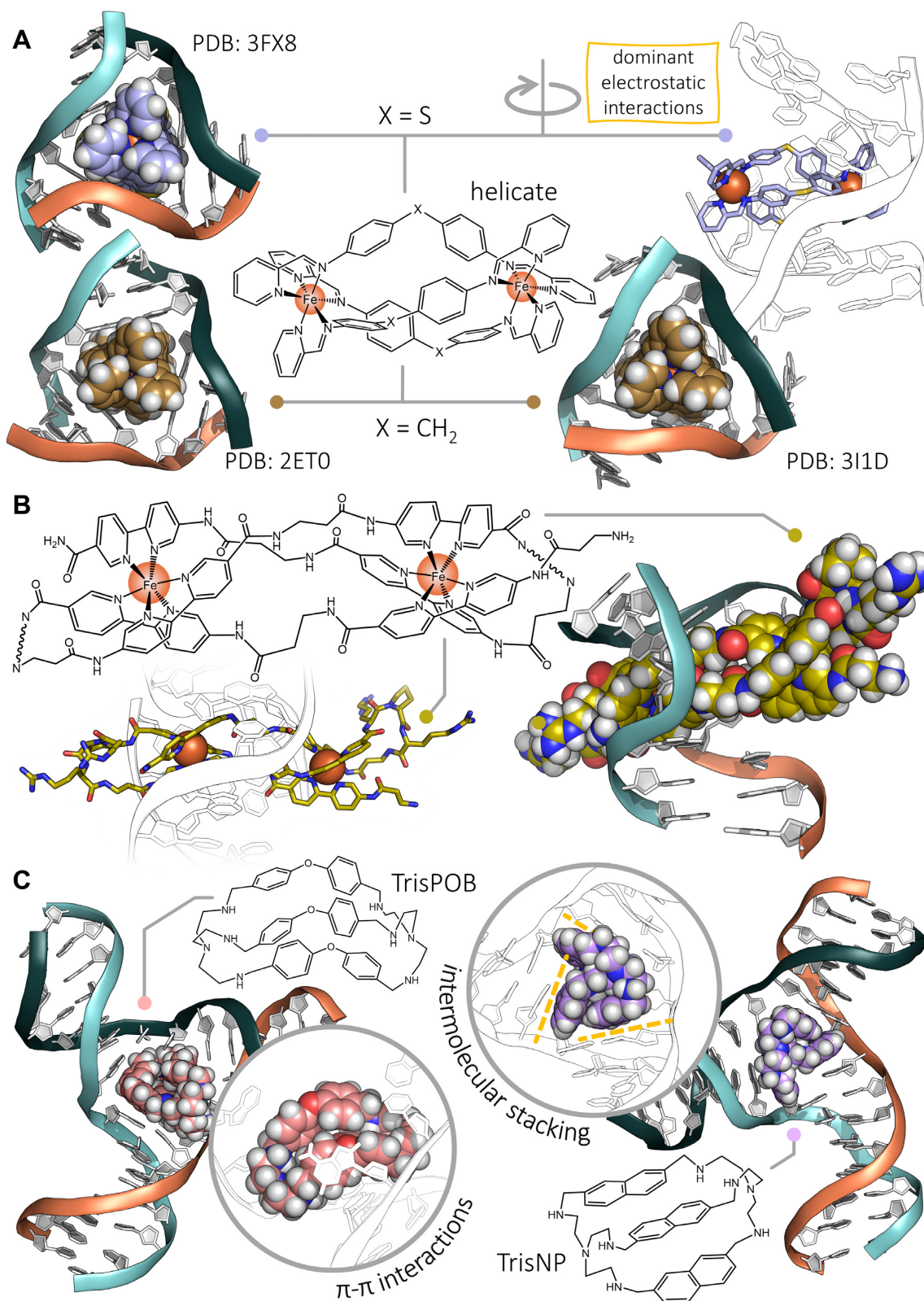
The discovery of DNA junctions in general, and of TWJs in particular, was concomitant with the advances in electron microscopy in the early 1960s, which allowed for highlighting a series of branched DNA, although in most cases the strandness was not determined accurately. For instance, branched DNAs were visualized in the DNA of *Pneumococcus* (1961) (157), of both T4 (1962) (158) and T7 bacteriophages (1964) (159) and of *Xenopus laevis* oocytes (1968) (160). A hypothetical model for the formation of TWJ was postulated in 1964 by Arthur Kornberg (161), as part of his longstanding study on the biological synthesis of DNA. TWJs started to be studied in more detail when designed and assembled *in vitro* (162,163), to assess the specificity of the phage T4 endonuclease VII (162), or investigate the effect of unpaired nucleotides at the branch point on the overall stability of the junction (163). Naturally occurring TWJs have also been characterized in adeno-associated viruses (AAVs) (164). However, contrarily to FWJs, no central, constitutive roles were described for TWJs in eukaryotic cells. TWJs are thus pletonemic structures that can form transiently ahead of replication forks (RFs, *vide infra*), which can form structural obstacles that threaten genetic stability, in line with their genotoxic roles in repeat expansion diseases (165).

With this in mind, it is thus unsurprising that TWJs were targeted to create DNA damage according to a novel strategy, i.e. the indirect inhibition of polymerases *via* the stabilisation of physical obstacles they cannot cope with efficiently. Beyond the pioneering work of Kallenbach (with iron and copper complexes, e.g. MPE-Fe, Figure 3A) (89), the proof of concept that TWJs could be targeted by small molecules was incidentally provided by the design and use

of aptamers as sensors for drugs (cocaine) (166–168) and steroids (e.g. cholic acid (169) and derivatives) (170). In these setups, TWJ-based aptamers were used firstly as turn-off (a dimeric TWJ is assembled upon small molecule binding), which brings the fluorophore (fluorescein) and the quencher (dabcyl) close to each other; (166,167) an externally added cyanine dye that is displaced from the cavity by the guest molecule) (168). Secondly, there are turn-on systems (a covalently linked fluorescein (quenched) that nestles within the TWJ cavity and is displaced by the guest molecule) (170). These results thus provided a basis for the design of hydrophobic TWJ ligands. The relevance of this approach was spectacularly demonstrated by Michael J. Hannon, Miquel Coll and coworkers with the elucidation of the X-ray structure of a TWJ/helicate complex (Figures 3C and 8A), in which the helicate (or triple-stranded dimetallic supramolecular cylinder)  $[\text{Fe}_2\text{L}_3]^{4+}$  triggers the assembly of a TWJ from a palindromic, 6-mer DNA sequence (171). This structure confirmed that the dimension of the cavity (ca. 1.1 nm diameter) makes it perfectly suited to accommodate rather large molecules (the size of the cylinder being ca. 2 nm length  $\times$  1 nm diameter). The exquisite recognition of TWJs by helicates was further demonstrated by another series of X-ray structure analyses (172), and confirmed by results collected *via* alternative techniques, including PAGE (173) and NMR (174). These results thus gave strong impetus to the use of sterically demanding small molecules, displaying shapes and volumes suited to ‘fill’ the TWJ cavity, including helical metalloproteins (studied by circular dichroism (CD), NMR and PAGE) (175–177), polyazamacrocycles such as TACN-Q (Figure 3C, by CD, PAGE, UV- and fluorescence resonance energy transfer (FRET)-melting assays) (178,179), porphyrins (by UV- and FRET-melting assays) (180), triptycenes (Figure 3C, by UV-melting and fluorescence quenching assays) (181,182), azacryptands such as TrisNP (Figure 3C) (155,156,179) and azacyclophanes (183,184) (by PAGE, FRET-melting and fluorescence quenching assays, and electrospray ionization mass spectrometry (ESI-MS) investigations), metallocages (by fluorescence titrations, fluorescence quenching assays and PAGE) (183–185), and calix[3]carbazoles (Figure 3C, by fluorescence titrations, CD, UV-melting assay and PAGE) (186).

### The relevance of TWJ in cells

With this series of fully characterised molecular tools in hand, cellular studies and chemical biology investigations were made possible, notably to assess the prevalence of TWJs in the human genome and the strategic relevance of targeting TWJs for creating DNA damage. The antiproliferative activity against human cancer cells were profiled for most of the aforementioned ligands (e.g. ovarian cancer SKOV3 cells and leukaemia HL-60 cells (187), melanoma B16 cells (179), breast cancer MCF-7 and MDA-MB-231 cells (184,188), colon carcinoma HCT-116 cells (189), etc.), but the mechanistic basis for this activity was sparsely established. These investigations were in fact slow to begin as a result of a combination of several factors: first, the biological relevance of TWJs in the studies performed with the first validated TWJ ligand (the dimetallic supramolecular cylin-



**Figure 8.** (A) X-ray crystal structures (PDB: 2ET0 (171), 3FX8 and 311D (172)) of a three-way DNA junction in interaction with dimetallic supramolecular cylinders (helicates). (B) Molecular dynamics (MD) simulation of a peptide helicate (Fe(II).LLD) in interaction with a TWJ (177). (C) Representative conformations of two azacryptands, TrisNP and TrisPOB, in interaction with a TWJ, obtained by MD simulations (156).

der  $[\text{Fe}_2\text{L}_3]^{4+}$ ) was misleadingly ascribed to RFs (171,187), while the fine structural details of RFs indicate that it is indeed a Y-shaped structure but made of unpaired sequences at the junction, unlike TWJs (190). Second, these first cellular investigations failed in providing convincing results, and the ability of  $[\text{Fe}_2\text{L}_3]^{4+}$  to interact with DNA in cells was even questioned (189). Third, the very nature of  $[\text{Fe}_2\text{L}_3]^{4+}$  made these investigations irrelevant owing to the ability of this supramolecular complex to interact with other DNA structures including the major grooves of B-DNA (191) and G4s (192,193). Therefore, this prototype was invaluable to uncover and accurately characterize a completely new DNA binding mode (with a molecule nestling within the central cavity of a TWJ) but was not ideally suited for chemical biology investigations. A very interesting observation, however, was that the closely related cylinder  $[\text{Ru}_2\text{L}_3]^{4+}$  was able to inhibit DNA transactions by preventing polymerase processivity; (194) this laid the basis for investigating whether TWJs might fold during DNA transactions in cells, more particularly at replication sites in the nucleus. To this end, co-localization studies were performed with a fluorescently labelled Fe(II) helicate (TAMRA-LLD, Figure 8B) and PCNA (proliferating cell nuclear antigen, a component of the replication machinery) fused to the green fluorescent protein (GFP) in HeLa cells (177). The collected images showed well-defined common *foci*, indicating that TWJs might fold transiently during DNA replication, in the vicinity of the replisome, and be trapped by *ad hoc* ligands. Similar co-localization studies were recently performed with a TWJ probe comprising a central, C3-symmetrical fluorogenic template (1,3,5-tristyrylbenzene) surrounded by three peptidic arms designed to interact with the three duplex arms of the TWJs (AT-hooks), thus found to preferentially accumulate in permeabilized cells where PCNA-GFP accumulates (195). However, the relevance of these results needs to be substantiated by a clear demonstration of the selectivity of used ligands for TWJs, as off-target effects cannot be ruled out on the basis of the currently available *in vitro* data.

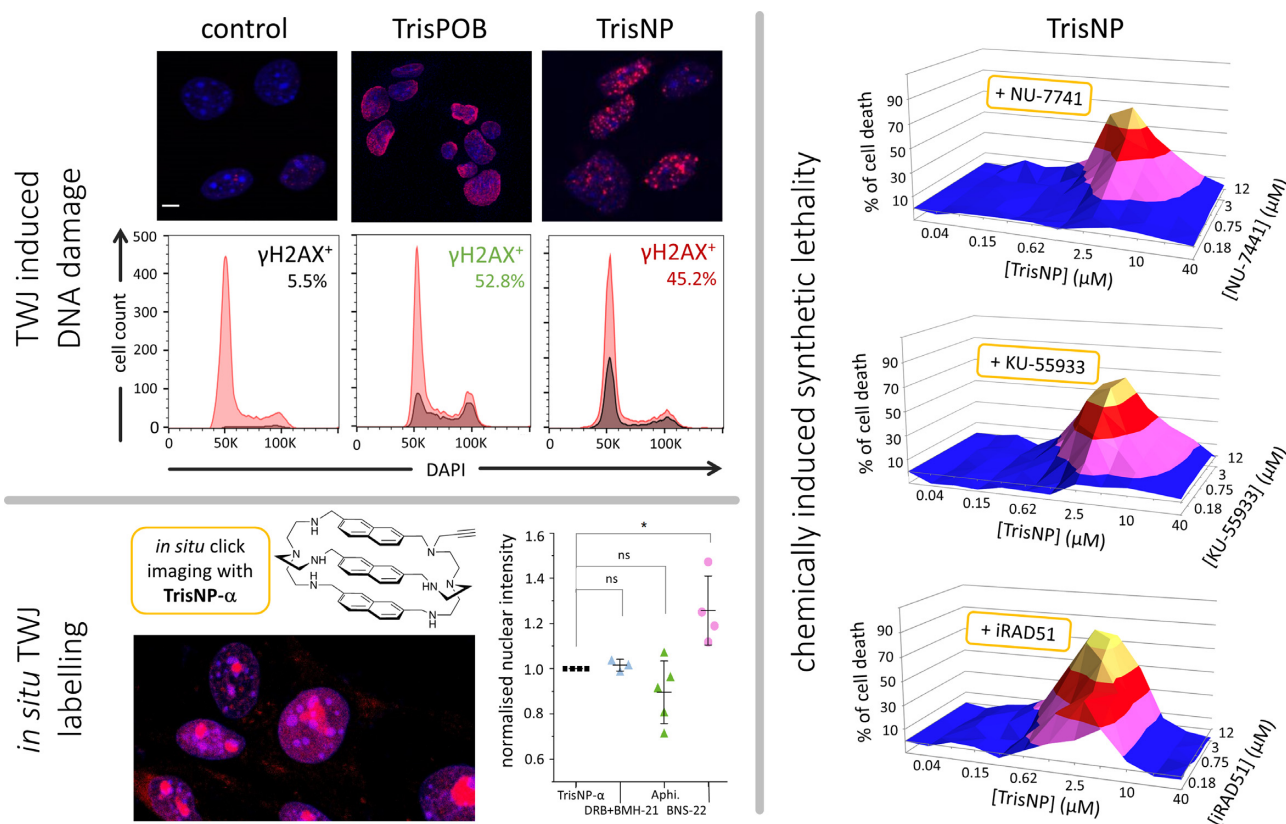
Studies were also performed to investigate whether TWJ ligands trigger DNA damage in dividing cells. This was firmly established *via* the immunodetection of  $\gamma\text{H2AX}$  by both optical imaging and flow cytometry; after a thorough characterization of their TWJ-interacting properties *in vitro* (affinity and specificity *via* a panel of different techniques), two azacryptands (TrisPOB, TrisNP) were studied in cells and shown to trigger an accumulation of DNA strand breaks (and particularly DSBs) (Figures 3C, 8C and 9) (155,156). When combined with the established ability of TWJ ligands to prevent polymerase activity *in vitro* (194), these results thus provided a strong rationale for a mechanistic model in which ligand-stabilized TWJs act as roadblocks that hamper proper processivity of DNA-related enzymes, slowing or even stopping their motion along the genomic duplex, which is recognized as a DNA damage and trigger the DDR and repair machineries. We further exploited this model by using drug combinations in which TWJ ligands were used to induce DNA damage and DNA repair inhibitors to act synergistically with them, in an approach referred to as chemically induced synthetic lethality (196). In eukaryotic cells, DSBs can be repaired by either HR (*vide supra*) or non-homologous end joining (NHEJ)

(Figure 6). The former relies on a broad array of protein effectors including RAD51 (which binds to single-stranded DNA, searches for sequence homology and favours strand exchange), and the kinases ATM and ATR (which bind to DSBs and stalled RFs, respectively); the latter also involves multiple processing enzymes, including kinases and ligases (e.g. DNA-PK and Lig4, respectively) (58,59,125). The antiproliferative activity of the TWJ ligands TrisPOB and TrisNP was found to be quite efficiently potentiated when combined with inhibitors of DNA-PK (NU-7441), ATM (KU-55933) and RAD51 (B02) (155,156), confirming that the cellular activity of TWJ ligands relies on the induction of strongly genotoxic DSBs (Figure 9). Interestingly, a synergy was also obtained with the Top2 inhibitor BNS-22: Top2 proteins resolve DNA topological stress (197) but are also involved in the recognition of alternative DNA structures and the formation of DSBs at these sites (198). Top2 also participates in the cleavage of hairpin structures formed from  $\alpha$ -satellite sequences *in vitro* (199), which are centromeric regions known to be highly repetitive (200) and to fold into secondary structures (201). This synergy, confirmed by cytotoxicity assessments and  $\gamma\text{H2AX}$  labelling, thus implies that Top2 inhibition favours TWJ formation, which was further demonstrated by bioorthogonal chemistry using the *in situ* clickable TWJ ligand TrisNP- $\alpha$ , Figure 3C) (156), which highlight the increase of the TWJ landscape (TrisNP- $\alpha$  labelling) upon Top2 inhibition (Figure 9). This opens brand new therapeutic opportunities, notably to treat cancers resistant to Top2 poisons.

As above, though comprehensive from a molecular and cellular biology point of view, these studies suffer from the fact that used TWJ ligands display a non-neglectable affinity for G4s. It was demonstrated *in vitro* that they display a preferential affinity for TWJ (i.e. in competitive setups where both TWJ and G4 are mixed, the ligands interact solely with TWJ) but this selectivity was not satisfying enough for an unambiguous interpretation of cellular outcomes. To confirm this interpretation, we showed by optical imaging that there is no overlaps between TWJ sites labelled with TrisNP- $\alpha$  (*in situ* click chemistry (202) with AF594-azide) and G4 sites labelled with the G4-specific antibody BG4 (35). These results are highly convincing but they are still not strong enough to unequivocally dispel doubts about the actual origins of the cellular effects monitored upon cell incubation with *bona fide* TWJ ligands. What the field now needs is a truly specific TWJ ligand; in light of the wealth of promising data described above, we can wager that a substantial research effort is currently being invested to identify the impatiently awaited game-changing TWJ ligand.

## CONCLUSION AND FUTURE DIRECTIONS

Targeting higher-order DNA structures with *ad hoc* ligands for chemical biology and/or medicinal chemistry purposes is now a commonly accepted strategy, undoubtedly as a consequence of the wealth of data accumulated about G4-DNA (68,69), the first-in-class example of a biologically relevant alternative DNA structure. While DNA junctions were discovered concomitantly with G4s (as indicated above, their basic structural unit was solved in 1962 (23) but their biological relevance was only discussed in the late



**Figure 9.** Cellular activity of the TWJ-ligands TrisNP and TrisPOB, which triggers extensive DNA damage in MCF7 cells (immunodetection of  $\gamma$ H2AX and flow cytometry) that can be potentiated by co-incubation with inhibitors of DNA repair (including NU-7441, KU-55933 and B02, which inhibit DNA-PK, ATM and RAD51, respectively) (155). The number of TWJ foci labelled by *in situ* click chemistry using TrisNP- $\alpha$  as TWJ-ligand is not modulated by transcription inhibitors DRB and BMH-21, weakly modulated (decrease) by the replication inhibitor aphidicolin and strongly modulated (increase) by the Top2 inhibitor BNS-22 (156).

1980s) (24,203,204), the investigations aiming at confirming the existence of DNA junctions in human cells, scrutinising their functional relevance and establishing the reliability of the therapeutic strategies based on their targeting with small molecules still lag way behind that of G4s. This might be attributed to several factors: first and foremost, right after the pioneering report by Stephen Neidle, Lawrence H. Hurley and co-workers on a small molecule able to interact with a G4 (in the aim of indirectly inhibiting telomerase *via* the sequestration of its telomeric substrate under a form that is not recognized by the enzyme) (205), hundreds of G4 ligands have been synthesized and studied (75,206). This incredible enthusiasm has provided a strong impetus for the discovery of truly specific ligands (e.g. PhenDC3 (207), PDS) (208) which have soon become invaluable molecular tools to decipher G4 biology in a highly accurate manner (209). The DNA junction field has not experienced such a keen interest, presumably because the biological relevance of DNA junctions is still poorly understood. Also, the search for genomic G4 sites *via in silico* techniques (e.g. QuadParser (210), G4hunter (211), etc.) was straightforward thanks to the very nature of G4-forming sequences (example of used algorithm:  $d(G_3+N_{1-7}G_3+N_{1-7}G_3+N_{1-7}G_3+)$ ) (212). They have provided strong arguments about the prevalence of G4s in the human genome, which was soon connected to a

widespread functional relevance. The nature of the DNA junction-forming sequences (direct or inverted sequence repeats, and not easily identifiable motif repeats such as GGG triplets) make them more difficult to be reliably predicted at a genome-wide scale (although some laudable attempts have been made, such as IRfinder (213), palindrome analyzer (214), etc.). Again, this has contributed to somewhat dampen enthusiasm for the search of genomic DNA junction-forming sites. Finally, the challenge of identifying genomic G4 sites by ChIP-seq-like techniques (such as G4 ChIP-seq) (50) was met, thanks to the development of the G4-specific antibody BG4 (35), and no similar antibody exists in the DNA junction field, which makes both their identification (ChIP) and visualisation (immunofluorescence) challenging and regularly questioned.

Despite these difficulties, the aforementioned efforts have started to pay off, but the prevalence of putative DNA junction-forming sequences described above (37,41) combined with the lack of fine details regarding DNA junction biology (which originates in the lack of reliable molecular tools) explain why this chemical biology quest is still in its infancy. Massive efforts must now be invested to keep on developing tools and technologies to interrogate and manipulate DNA junctions in a functional cellular context. The next steps will thus be to (i) identify truly specific ligands for DNA junctions, in order to establish a reliable link between



junction interaction in cells and the cellular outcomes monitored; (ii) create and/or optimize *in vitro* screening assays to blind screen commercially available chemolibraries in order to expand the portfolio of DNA junction ligand candidates far beyond rational design; (iii) develop on the basis of identified promising scaffolds multivalent molecular tools (bearing for instance an alkyne appendage that could be manipulated in a bioorthogonal manner in cells) to identify the sequences they interact with (direct repeats? inverted repeats? other?) and the proteins associated combining both sequencing (bioinformatic tools will have to be optimized as well) and proteomics. Of note, the identification of DNA junction-binding proteins will be a very important step as they are the cellular effectors by which the response to junction stabilization by *ad hoc* ligands is mediated; while hundreds of G4-binding proteins are now known (215), only a handful of junction-binding proteins have been reported to date (e.g. RuvABC and RecG, *vide supra*); (iv) produce DNA junction-specific antibodies for both immunoprecipitation and immunodetection purposes, and (v) develop cellular and small animal models to assess the properties of identified candidates in a standardized manner.

The ultimate goal of these investigations will be to rapidly provide a proof-of-concept that DNA junction-targeting molecules can be considered as key players in the field of therapeutic agents, to validate all the necessary pre-clinical milestones in a reliable manner in the aim of reaching the clinical stage rapidly and confidently. Without a doubt, these chemical biology investigations will lead to major advances in research on this new class of therapeutic targets, a momentum that will contribute to both better understand the biology of DNA junctions (which will find applications in diverse therapeutic areas such as cancers (62), neuropathologies (165) and infectious diseases) and unravel the fascinating structural and functional diversity of DNA.

## DATA AVAILABILITY

No new data were generated or analysed in support of this research.

## ACKNOWLEDGEMENTS

We are indebted to the different funding agencies that supported our work on DNA junctions over the past years, including the CNRS (D.M.), European Commission (H2020-MSCA-IF-2016–750638, D.M.), Plan Cancer INSERM/Avesian (19CP117-00, D.M.), iSITE BFC (COMUE UBFC, PIA2, UB21018.MUB.IS, D.M./A.P.), Agence Nationale de la Recherche (ANR-22 InJUNCTION, D.M./A.P.) and BBSRC (BB/T008342/1, K.M./C.J.C.). We also acknowledge our colleagues worldwide to make this field a lively and thrilling research area, especially Drs Anton Granzhan (Institut Curie, Orsay, FR), Sébastien Britton (IPBS Toulouse, FR) and Nicolas Chéron (ENS Paris, FR) for our close, stimulating and fruitful collaboration.

## FUNDING

Agence Nationale de la Recherche [ANR-22-CE44]; European Regional Development Fund [BG0021532];

ISITE BFC [UB21018.MUB.IS]. Funding for open access charge: ISTE BFC.

*Conflict of interest statement.* None declared.

## REFERENCES

- Chargaff, E., Zamenhof, S. and Green, C. (1950) Human deoxyribose nucleic acid: composition of human deoxyribose nucleic acid. *Nature*, **165**, 756–757.
- Watson, J.D. and Crick, F.H. (1953) Molecular structure of nucleic acids. *Nature*, **171**, 737–738.
- Malyshev, D.A., Dhama, K., Lavergne, T., Chen, T., Dai, N., Foster, J.M., Corrêa, I.R. and Romesberg, F.E. (2014) A semi-synthetic organism with an expanded genetic alphabet. *Nature*, **509**, 385–388.
- Zhang, Y., Lamb, B.M., Feldman, A.W., Zhou, A.X., Lavergne, T., Li, L. and Romesberg, F.E. (2017) A semisynthetic organism engineered for the stable expansion of the genetic alphabet. *Proc. Natl. Acad. Sci. U.S.A.*, **114**, 1317–1322.
- Hoshika, S., Leal, N.A., Kim, M.-J., Kim, M.-S., Karalkar, N.B., Kim, H.-J., Bates, A.M., Watkins, N.E., SantaLucia, H.A., Meyer, A.J. *et al.* (2019) Hachimoji DNA and RNA: a genetic system with eight building blocks. *Science*, **363**, 884–887.
- Kirnos, M.D., Khudyakov, I.Y., Alexandrushkina, N.I. and Vanyushin, B.F. (1977) 2-Amino adenine is an adenine substituting for a base in S-2L cyanophage DNA. *Nature*, **270**, 369–370.
- Pezo, V., Jaziri, F., Bourguignon, P.-Y., Louis, D., Jacobs-Sera, D., Rozenski, J., Pochet, S., Herdewijn, P., Hatfull, G.F., Kaminski, P.-A. *et al.* (2021) Noncanonical DNA polymerization by amino adenine-based siphoviruses. *Science*, **372**, 520–524.
- Sleiman, D., Garcia, P.S., Lagune, M., Loc'h, J., Haouz, A., Taib, N., Röthlisberger, P., Gribaldo, S., Marière, P. and Kaminski, P.A. (2021) A third purine biosynthetic pathway encoded by amino adenine-based viral DNA genomes. *Science*, **372**, 516–520.
- Zhou, Y., Xu, X., Wei, Y., Cheng, Y., Guo, Y., Khudyakov, I., Liu, F., He, P., Song, Z., Li, Z. *et al.* (2021) A widespread pathway for substitution of adenine by diaminopurine in phage genomes. *Science*, **372**, 512–516.
- Ghosh, A. and Bansal, M. (2003) A glossary of DNA structures from a to Z. *Acta Crystallogr. D: Biol. Crystallogr.*, **59**, 620–626.
- Franklin, R.E. and Gosling, R.G. (1953) The structure of sodium thymonucleate fibres. I. The influence of water content. *Acta Crystallogr.*, **6**, 673–677.
- Franklin, R.E. and Gosling, R.G. (1953) Evidence for 2-chain helix in crystalline structure of sodium deoxyribonucleate. *Nature*, **172**, 156–157.
- Franklin, R.E. and Gosling, R.G. (1953) Molecular configuration in sodium thymonucleate. *Nature*, **171**, 740–741.
- Elkin, L.O. (2003) Rosalind Franklin and the double helix. *Physics Today*, **56**, 42–48.
- Ravichandran, S., Subramani, V.K. and Kim, K.K. (2019) Z-DNA in the genome: from structure to disease. *Biophys. Rev.*, **11**, 383–387.
- Crick, F.H. (1958) On protein synthesis. *Symp. Soc. Exp. Biol.*, **12**, 138–163.
- Crick, F. (1970) Central dogma of molecular biology. *Nature*, **227**, 561.
- Wilkins, M.H.F., Stokes, A.R. and Wilson, H.R. (1953) Molecular structure of nucleic acids: molecular structure of deoxyribose nucleic acids. *Nature*, **171**, 738–740.
- Dalla Pozza, M., Abdullrahman, A., Cardin, C.J., Gasser, G. and Hall, J.P. (2022) Three's a crowd – stabilisation, structure, and applications of DNA triplexes. *Chem. Sci.*, **13**, 10193–10215.
- Htun, H. and Dahlberg, J.E. (1989) Topology and formation of triple-stranded H-DNA. *Science*, **243**, 1571–1576.
- Felsenfeld, G., Davies, D.R. and Rich, A. (1957) Formation of a three-stranded polynucleotide molecule. *J. Am. Chem. Soc.*, **79**, 2023–2024.
- Lee, J.S., Johnson, D.A. and Morgan, A.R. (1979) Complexes formed by (pyrimidine)*n*-(purine)*n* DNAs on lowering the pH are three-stranded. *Nucleic Acids Res.*, **6**, 3073–3091.
- Gellert, M., Lipsett, M.N. and Davies, R.D. (1962) Helix formation by guanylic acid. *Proc. Natl. Acad. Sci. U.S.A.*, **48**, 2013–2018.

24. Sen, D. and Gilbert, W. (1988) Formation of parallel four-stranded complexes by guanine-rich motifs in DNA and its implications for meiosis. *Nature*, **334**, 364–366.
25. Holliday, R. (1964) A mechanism for gene conversion in fungi. *Genet. Res.*, **5**, 282–304.
26. Gierer, A. (1966) Model for DNA and protein interactions and the function of the operator. *Nature*, **212**, 1480–1481.
27. Lilley, D.M. (1980) The inverted repeat as a recognizable structural feature in supercoiled DNA molecules. *Proc. Natl. Acad. Sci. U.S.A.*, **77**, 6468–6472.
28. Panayotatos, N. and Wells, R.D. (1981) Cruciform structures in supercoiled DNA. *Nature*, **289**, 466–470.
29. Pohl, F.M. (1967) Ein Modell der DNS-Struktur. *Naturwissenschaften*, **54**, 616–616.
30. Wang, A.H.J., Quigley, G.J., Kolpak, F.J., Crawford, J.L., van Boom, J.H., van der Marel, G. and Rich, A. (1979) Molecular structure of a left-handed double helical DNA fragment at atomic resolution. *Nature*, **282**, 680–686.
31. Sivakova, S. and Rowan, S.J. (2005) Nucleobases as supramolecular motifs. *Chem. Soc. Rev.*, **34**, 9–21.
32. Hoogsteen, K. (1959) The structure of crystals containing a hydrogen-bonded complex of 1-methylthymine and 9-methyladenine. *Acta Crystallogr.*, **12**, 822–823.
33. Hoogsteen, K.R. (1963) The crystal and molecular structure of a hydrogen-bonded complex between 1-methylthymine and 9-methyladenine. *Acta Crystallogr.*, **16**, 907–916.
34. Arnott, S. and Bond, P.J. (1973) Structures for Poly(U), poly(A), poly(U) triple stranded polynucleotides. *Nat. New Biol.*, **244**, 99–101.
35. Biffi, G., Tannahill, D., McCafferty, J. and Balasubramanian, S. (2013) Quantitative visualization of DNA G-quadruplex structures in human cells. *Nat. Chem.*, **5**, 182–186.
36. Di Antonio, M., Ponjavic, A., Radzevičius, A., Ranasinghe, R.T., Catalano, M., Zhang, X., Shen, J., Needham, L.-M., Lee, S.F., Klenerman, D. *et al.* (2020) Single-molecule visualization of DNA G-quadruplex formation in live cells. *Nat. Chem.*, **12**, 832–837.
37. Treangen, T.J. and Salzberg, S.L. (2012) Repetitive DNA and next-generation sequencing: computational challenges and solutions. *Nat. Rev. Genet.*, **13**, 36–46.
38. Wang, G. and Vasquez, K.M. (2006) Non-B DNA structure-induced genetic instability. *Mutat. Res.*, **598**, 103–119.
39. Pearson, C.E. and Sinden, R.R. (1996) Alternative structures in duplex DNA formed within the trinucleotide repeats of the myotonic dystrophy and fragile x loci. *Biochem.*, **35**, 5041–5053.
40. Wells, R.D. (2007) Non-B DNA conformations, mutagenesis and disease. *Trends Biochem. Sci.*, **32**, 271–278.
41. Georgakopoulos-Soares, I., Morganello, S., Jain, N., Hemberg, M. and Nik-Zainal, S. (2018) Noncanonical secondary structures arising from non-B DNA motifs are determinants of mutagenesis. *Genome Res.*, **28**, 1264–1271.
42. Murat, P., Guilbaud, G. and Sale, J.E. (2020) DNA polymerase stalling at structured DNA constrains the expansion of short tandem repeats. *Genome Biol.*, **21**, 209.
43. Singleton, M.R., Dillingham, M.S. and Wigley, D.B. (2007) Structure and mechanism of helicases and nucleic acid translocases. *Annu. Rev. Biochem.*, **76**, 23–50.
44. Chambers, V.S., Marsico, G., Boutell, J.M., Di Antonio, M., Smith, G.P. and Balasubramanian, S. (2015) High-throughput sequencing of DNA G-quadruplex structures in the human genome. *Nat. Biotechnol.*, **33**, 877–881.
45. Peña Martínez, C.D., Zeraati, M., Rouet, R., Mazigi, O., Gloss, B., Chan, C.-L., Bryan, T.M., Smith, N.M., Dinger, M.E., Kummerfeld, S. *et al.* (2022) Human genomic DNA is widely interspersed with i-motif structures. bioRxiv doi: <https://doi.org/10.1101/2022.04.14.488274>, 14 April 2022, preprint: not peer reviewed.
46. Marsico, G., Chambers, V.S., Sahakyan, A.B., McCauley, P., Boutell, J.M., Antonio, M.D. and Balasubramanian, S. (2019) Whole genome experimental maps of DNA G-quadruplexes in multiple species. *Nucleic Acids Res.*, **47**, 3862–3874.
47. Feng, Y., Tao, S., Zhang, P., Rota Sperti, F., Liu, G., Cheng, X., Zhang, T., Yu, H., Wang, X., Chen, C. *et al.* (2022) Epigenomic features of DNA G-quadruplexes and their roles in regulating rice gene transcription. *Plant Physiol.*, **188**, 1632–1648.
48. Ma, X., Feng, Y., Yang, Y., Li, X., Shi, Y., Tao, S., Cheng, X., Huang, J., Wang, X., Chen, C. *et al.* (2022) Genome-wide characterization of i-motifs and their potential roles in the stability and evolution of transposable elements in rice. *Nucleic Acids Res.*, **50**, 3226–3238.
49. Li, M., Tian, R., Monchaud, D. and Zhang, W. (2022) Omics studies of DNA G-/C-quadruplexes in plants. *Trends Genet.*, **38**, 999–1002.
50. Hänsel-Hertsch, R., Beraldi, D., Lensing, S.V., Marsico, G., Zyner, K., Parry, A., Di Antonio, M., Pike, J., Kimura, H. and Narita, M. (2016) G-quadruplex structures mark human regulatory chromatin. *Nat. Genet.*, **48**, 1267–1272.
51. Kouzine, F., Wojtowicz, D., Baranello, L., Yamane, A., Nelson, S., Resch, W., Kieffer-Kwon, K.-R., Benham, C.J., Casellas, R., Przytycka, T.M. *et al.* (2017) Permanganate/S1 nuclease footprinting reveals Non-B DNA structures with regulatory potential across a mammalian genome. *Cell Syst.*, **4**, 344–356.
52. van Wietmarschen, N., Sridharan, S., Nathan, W.J., Tubbs, A., Chan, E.M., Callen, E., Wu, W., Belinky, F., Tripathi, V., Wong, N. *et al.* (2020) Repeat expansions confer WRN dependence in microsatellite-unstable cancers. *Nature*, **586**, 292–298.
53. Matos-Rodrigues, G., van Wietmarschen, N., Wu, W., Tripathi, V., Koussa, N.C., Pavani, R., Nathan, W.J., Callen, E., Belinky, F., Mohammed, A. *et al.* (2022) S1-END-seq reveals DNA secondary structures in human cells. *Mol. Cell*, **82**, 3538–3552.
54. Shah, K.A., Shishkin, A.A., Voineagu, I., Pavlov, Y.I., Shcherbakova, P.V. and Mirkin, S.M. (2012) Role of DNA polymerases in repeat-mediated genome instability. *Cell Rep.*, **2**, 1088–1095.
55. Shah, K.A. and Mirkin, S.M. (2015) The hidden side of unstable DNA repeats: mutagenesis at a distance. *DNA Repair*, **32**, 106–112.
56. Wang, G. and Vasquez, K.M. (2014) Impact of alternative DNA structures on DNA damage, DNA repair, and genetic instability. *DNA Repair*, **19**, 143–151.
57. Jackson, S.P. and Bartek, J. (2009) The DNA-damage response in human biology and disease. *Nature*, **461**, 1071–1078.
58. Ciccio, A. and Elledge, S.J. (2010) The DNA damage response: making it safe to play with knives. *Mol. Cell*, **40**, 179–204.
59. Blackford, A.N. and Jackson, S.P. (2017) ATM, ATR, and DNA-PK: the trinity at the heart of the DNA damage response. *Mol. Cell*, **66**, 801–817.
60. Mirkin, E.V. and Mirkin, S.M. (2007) Replication fork stalling at natural impediments. *Microbiol. Mol. Biol. Rev.*, **71**, 13–35.
61. Belotserkovskii, B.P., Mirkin, S.M. and Hanawalt, P.C. (2013) DNA sequences that interfere with transcription: implications for genome function and stability. *Chem. Rev.*, **113**, 8620–8637.
62. Zell, J., Rota Sperti, F., Britton, S. and Monchaud, D. (2021) DNA folds threaten genetic stability and can be leveraged for chemotherapy. *RSC Chem. Biol.*, **2**, 47–76.
63. Kaushal, S. and Freudenreich, C.H. (2019) The role of fork stalling and DNA structures in causing chromosome fragility. *Genes Chromosomes Cancer*, **58**, 270–283.
64. del Mundo, I.M., Vasquez, K.M. and Wang, G. (2019) Modulation of DNA structure formation using small molecules. *Biochim. Biophys. Acta - Mol. Cell Res.*, **1866**, 118539.
65. Roos, W.P., Thomas, A.D. and Kaina, B. (2016) DNA damage and the balance between survival and death in cancer biology. *Nat. Rev. Cancer*, **16**, 20–33.
66. Hilton, J., Gelmon, K., Bedard, P.L., Tu, D., Xu, H., Tinker, A.V., Goodwin, R., Laurie, S.A., Jonker, D., Hansen, A.R. *et al.* (2022) Results of the phase I CCTG IND.231 trial of CX-5461 in patients with advanced solid tumors enriched for DNA-repair deficiencies. *Nat. Commun.*, **13**, 3607.
67. Burge, S., Parkinson, G.N., Hazel, P., Todd, A.K. and Neidle, S. (2006) Quadruplex DNA: sequence, topology and structure. *Nucleic Acids Res.*, **34**, 5402–5415.
68. Rhodes, D. and Lipps, H.J. (2015) G-quadruplexes and their regulatory roles in biology. *Nucleic Acids Res.*, **43**, 8627–8637.
69. Varshney, D., Spiegel, J., Zyner, K., Tannahill, D. and Balasubramanian, S. (2020) The regulation and functions of DNA and RNA G-quadruplexes. *Nat. Rev. Mol. Cell Biol.*, **21**, 459–474.
70. Lilley, D.M. (2000) Structures of helical junctions in nucleic acids. *Q. Rev. Biophys.*, **33**, 109–159.
71. Altona, C. (1996) Classification of nucleic acid junctions. *J. Mol. Biol.*, **263**, 568–581.

72. Seeman, N.C. and Kallenbach, N.R. (1994) DNA branched junctions. *Annu. Rev. Biophys. Biomol. Struct.*, **23**, 53–86.
73. Monchaud, D. and Teulade-Fichou, M.-P. (2008) A hitchhiker's guide to G-quadruplex ligands. *Org. Biomol. Chem.*, **6**, 627–636.
74. Le, T.V.T., Han, S., Chae, J. and Park, H.-J. (2012) G-Quadruplex binding ligands: from naturally occurring to rationally designed molecules. *Curr. Pharm. Des.*, **18**, 1948–1972.
75. Neidle, S. (2016) Quadruplex nucleic acids as novel therapeutic targets. *J. Med. Chem.*, **59**, 5987–6011.
76. Spiegel, J., Adhikari, S. and Balasubramanian, S. (2020) The structure and function of DNA G-quadruplexes. *Trends Chem.*, **2**, 123–136.
77. Kosiol, N., Juranek, S., Brossart, P., Heine, A. and Paeschke, K. (2021) G-quadruplexes: a promising target for cancer therapy. *Mol. Cancer*, **20**, 40.
78. Ivens, E., Cominetti, M.M.D. and Searcey, M. (2022) Junctions in DNA: underexplored targets for therapeutic intervention. *Bioorg. Med. Chem.*, **69**, 116897.
79. Lu, M., Guo, Q. and Kallenbach, N.R. (1992) Interaction of drugs with branched DNA structures. *Crit. Rev. Biochem. Mol. Biol.*, **27**, 157–190.
80. Marky, L.A., Kallenbach, N.R., McDonough, K.A., Seeman, N.C. and Breslauer, K.J. (1987) The melting behavior of a DNA junction structure: a calorimetric and spectroscopic study. *Biopolymers*, **26**, 1621–1634.
81. Seeman, N.C., Chen, J.H. and Kallenbach, N.R. (1989) Gel electrophoretic analysis of DNA branched junctions. *Electrophoresis*, **10**, 345–354.
82. Guo, Q., Seeman, N.C. and Kallenbach, N.R. (1989) Site-specific interaction of intercalating drugs with a branched DNA molecule. *Biochem.*, **28**, 2355–2359.
83. Churchill, M., Tullius, T.D., Kallenbach, N.R. and Seeman, N.C. (1988) A holliday recombination intermediate is twofold symmetric. *Proc. Natl. Acad. Sci. U.S.A.*, **85**, 4653–4656.
84. Lu, M., Guo, Q. and Kallenbach, N.R. (1993) Thermodynamics of G-tetraplex formation by telomeric DNAs. *Biochem.*, **32**, 598–601.
85. Lu, M., Guo, Q. and Kallenbach, N.R. (1992) Structure and stability of sodium and potassium complexes of dT4G4 and dT4G4T. *Biochem.*, **31**, 2455–2459.
86. Hernández, L.I., Zhong, M., Courtney, S.H., Marky, L.A. and Kallenbach, N.R. (1994) Equilibrium analysis of ethidium binding to DNA containing base mismatches and branches. *Biochem.*, **33**, 13140–13146.
87. Lu, M., Guo, Q., Seeman, N.C. and Kallenbach, N.R. (1990) Drug binding by branched DNA: selective interaction of the dye stains-all with an immobile junction. *Biochem.*, **29**, 3407–3412.
88. Lu, M., Guo, Q., Pasternack, R.F., Wink, D.J., Seeman, N.C. and Kallenbach, N.R. (1990) Drug binding by branched DNA: selective interaction of tetrapyrrolyl porphyrins with an immobile junction. *Biochem.*, **29**, 1614–1624.
89. Guo, Q., Lu, M., Churchill, M., Tullius, T. and Kallenbach, N.R. (1990) Asymmetric structure of a three-arm DNA junction. *Biochem.*, **29**, 10927–10934.
90. Guo, Q., Lu, M., Marky, L.A. and Kallenbach, N.R. (1992) Interaction of the dye ethidium bromide with DNA containing guanine repeats. *Biochem.*, **31**, 2451–2455.
91. Liu, Y. and West, S.C. (2004) Happy hollidays: 40th anniversary of the holliday junction. *Nat. Rev. Mol. Cell Biol.*, **5**, 937–944.
92. Gall, J.G. (1954) Lampbrush chromosomes from oocyte nuclei of the newt. *J. Morphol.*, **94**, 283–351.
93. Cooper, J.P. and Hagerman, P.J. (1987) Gel electrophoretic analysis of the geometry of a DNA four-way junction. *J. Mol. Biol.*, **198**, 711–719.
94. Duckett, D.R., Murchie, A.I., Diekmann, S., von Kitzing, E., Kemper, B. and Lilley, D.M. (1988) The structure of the holliday junction, and its resolution. *Cell*, **55**, 79–89.
95. Gough, G.W. and Lilley, D.M. (1985) DNA bending induced by cruciform formation. *Nature*, **313**, 154–156.
96. Cooper, J.P. and Hagerman, P.J. (1990) Analysis of fluorescence energy transfer in duplex and branched DNA molecules. *Biochem.*, **29**, 9261–9268.
97. Murchie, A.I., Clegg, R.M., von Krtzing, E., Duckett, D.R., Diekmann, S. and Lilley, D.M. (1989) Fluorescence energy transfer shows that the four-way DNA junction is a right-handed cross of antiparallel molecules. *Nature*, **341**, 763–766.
98. Clegg, R.M., Murchie, A.I., Zechel, A., Carlberg, C., Diekmann, S. and Lilley, D.M. (1992) Fluorescence resonance energy transfer analysis of the structure of the four-way DNA junction. *Biochem.*, **31**, 4846–4856.
99. Lilley, D.M. and Wilson, T.J. (2000) Fluorescence resonance energy transfer as a structural tool for nucleic acids. *Curr. Opin. Chem. Biol.*, **4**, 507–517.
100. Cooper, J.P. and Hagerman, P.J. (1989) Geometry of a branched DNA structure in solution. *Proc. Natl. Acad. Sci. U.S.A.*, **86**, 7336–7340.
101. Chen, S.M. and Chazin, W.J. (1994) Two-dimensional <sup>1</sup>H NMR studies of immobile holliday junctions: nonlabile proton assignments and identification of crossover isomers. *Biochem.*, **33**, 11453–11459.
102. Pikkemaat, J.A., van den Elst, H., van Boom, J.H. and Altona, C. (1994) NMR studies and conformational analysis of a DNA four-way junction formed in a linear synthetic oligonucleotide. *Biochem.*, **33**, 14896–14907.
103. Potter, H. and Dressler, D. (1977) On the mechanism of genetic recombination: the maturation of recombination intermediates. *Proc. Natl. Acad. Sci. U.S.A.*, **74**, 4168–4172.
104. Mao, C., Sun, W. and Seeman, N.C. (1999) Designed two-dimensional DNA holliday junction arrays visualized by atomic force microscopy. *J. Am. Chem. Soc.*, **121**, 5437–5443.
105. Lilley, D.M. and Norman, D.G. (1999) The holliday junction is finally seen with crystal clarity. *Nat. Struct. Biol.*, **6**, 897–899.
106. Ortiz-Lombardía, M., González, A., Erić, R., Aymami, J., Azorín, F. and Coll, M. (1999) Crystal structure of a DNA holliday junction. *Nat. Struct. Biol.*, **6**, 913–917.
107. Eichman, B.F., Vargason, J.M., Mooers, B.H. and Ho, P.S. (2000) The holliday junction in an inverted repeat DNA sequence: sequence effects on the structure of four-way junctions. *Proc. Natl. Acad. Sci. U.S.A.*, **97**, 3971–3976.
108. Roe, S.M., Barlow, T., Brown, T., Oram, M., Keeley, A., Tsaneva, I.R. and Pearl, L.H. (1998) Crystal structure of an octameric ruva–Holliday junction complex. *Mol. Cell*, **2**, 361–372.
109. Hargreaves, D., Rice, D.W., Sedelnikova, S.E., Artymiuk, P.J., Lloyd, R.G. and Rafferty, J.B. (1998) Crystal structure of *e. coli* RuvA with bound DNA holliday junction at 6 Å resolution. *Nat. Struct. Biol.*, **5**, 441–446.
110. Gopaul, D.N., Guo, F. and Van Duyne, G.D. (1998) Structure of the holliday junction intermediate in Cre–loxP site-specific recombination. *EMBO J.*, **17**, 4175–4187.
111. Lilley, D.M. (2017) Holliday junction-resolving enzymes—structures and mechanisms. *FEBS Lett.*, **591**, 1073–1082.
112. Hays, F.A., Teegarden, A., Jones, Z.J.R., Harms, M., Raup, D., Watson, J., Cavaliere, E. and Ho, P.S. (2005) How sequence defines structure: a crystallographic map of DNA structure and conformation. *Proc. Natl. Acad. Sci. U.S.A.*, **102**, 7157–7162.
113. Eichman, B.F., Mooers, B.H.M., Alberti, M., Hearst, J.E. and Ho, P.S. (2001) The crystal structures of psoralen cross-linked DNAs: drug-dependent formation of holliday junctions. *J. Mol. Biol.*, **308**, 15–26.
114. Adams, A., Guss, J.M., Collyer, C.A., Denny, W.A. and Wakelin, L.P.G. (2000) A novel form of intercalation involving four DNA duplexes in an acridine-4-carboxamide complex of d(CGTACG)<sub>2</sub>. *Nucleic Acids Res.*, **28**, 4244–4253.
115. Thorpe, J.H., Hobbs, J.R., Todd, A.K., Denny, W.A., Charlton, P. and Cardin, C.J. (2000) Guanine specific binding at a DNA junction formed by d CG(5-BrU)ACG 2 with a topoisomerase poison in the presence of Co<sup>2+</sup> ions. *Biochem.*, **39**, 15055–15061.
116. Teixeira, S.C.M., Thorpe, J.H., Todd, A.K., Powell, H.R., Adams, A., Wakelin, L.P.G., Denny, W.A. and Cardin, C.J. (2002) Structural characterisation of bisintercalation in higher-order DNA at a junction-like quadruplex. *J. Mol. Biol.*, **323**, 167–171.
117. Hopcroft, N.H., Brogden, A.L., Searcey, M. and Cardin, C.J. (2006) X-ray crystallographic study of DNA duplex cross-linking: simultaneous binding to two d(CGTACG)<sub>2</sub> molecules by a bis(9-aminoacridine-4-carboxamide) derivative. *Nucleic Acids Res.*, **34**, 6663–6672.
118. Brogden, A.L., Hopcroft, N.H., Searcey, M. and Cardin, C.J. (2007) Ligand bridging of the DNA holliday junction: molecular recognition of a stacked-X four-way junction by a small molecule. *Angew. Chem. Int. Ed.*, **46**, 3850–3854.

119. Howell, L.A., Waller, Z.A.E., Bowater, R., O'Connell, M. and Searcey, M. (2011) A small molecule that induces assembly of a four way DNA junction at low temperature. *Chem. Commun.*, **47**, 8262–8264.
120. Singh, H., Singh, H., Sharma, S., Mohinder, P. and Bedi, S. (2015) Chemotherapeutic potential of acridine analogs: an ample review. *Heterocycles*, **91**, 2043–2085.
121. Read, M., Harrison, R.J., Romagnoli, B., Tanious, F.A., Gowan, S.H., Reszka, A.P., Wilson, W.D., Kelland, L.R. and Neidle, S. (2001) Structure-based design of selective and potent  $\gamma$  quadruplex-mediated telomerase inhibitors. *Proc. Natl. Acad. Sci. U.S.A.*, **98**, 4844–4849.
122. Chien, C.-M., Wu, P.-C., Satange, R., Chang, C.-C., Lai, Z.-L., Hagler, L.D., Zimmerman, S.C. and Hou, M.-H. (2020) Structural basis for targeting T:T mismatch with triaminotriazine-acridine conjugate induces a U-Shaped Head-to-Head four-way junction in CTG repeat DNA. *J. Am. Chem. Soc.*, **142**, 11165–11172.
123. Hernández, A.F., Parrón, T., Tsatsakis, A.M., Requena, M., Alarcón, R. and López-Guarnido, O. (2013) Toxic effects of pesticide mixtures at a molecular level: their relevance to human health. *Toxicology*, **307**, 136–145.
124. Xia, J., Mei, Q. and Rosenberg, S.M. (2019) Tools to live by: bacterial DNA structures illuminate cancer. *Trends Genet.*, **35**, 383–395.
125. Scully, R., Panday, A., Elango, R. and Willis, N.A. (2019) DNA double-strand break repair-pathway choice in somatic mammalian cells. *Nat. Rev. Mol. Cell Biol.*, **20**, 698–714.
126. Vos, M. (2009) Why do bacteria engage in homologous recombination? *Trends Microbiol.*, **17**, 226–232.
127. West, S.C. (2003) Molecular views of recombination proteins and their control. *Nat. Rev. Mol. Cell Biol.*, **4**, 435–445.
128. Xia, J., Chen, L.-T., Mei, Q., Ma, C.-H., Halliday, J.A., Lin, H.-Y., Magnan, D., Pribis, J.P., Fitzgerald, D.M., Hamilton, H.M. *et al.* (2016) Holliday junction trap shows how cells use recombination and a junction-guardian role of RecQ helicase. *Sci. Adv.*, **2**, e1601605.
129. Cassell, G.D. and Segall, A.M. (2003) Mechanism of inhibition of site-specific recombination by the holliday junction-trapping peptide WKHYNV: insights into phage  $\lambda$  Integrase-mediated strand exchange. *J. Mol. Biol.*, **327**, 413–429.
130. Ghosh, K., Lau, C.K., Guo, F., Segall, A.M. and Van Duyne, G.D. (2005) Peptide trapping of the holliday junction intermediate in Cre-loxP Site-specific recombination\*. *J. Biol. Chem.*, **280**, 8290–8299.
131. McGlynn, P. and Lloyd, R.G. (2001) Rescue of stalled replication forks by recg: simultaneous translocation on the leading and lagging strand templates supports an active DNA unwinding model of fork reversal and holliday junction formation. *Proc. Natl. Acad. Sci. U.S.A.*, **98**, 8227–8234.
132. Kepple, K.V., Boldt, J.L. and Segall, A.M. (2005) Holliday junction-binding peptides inhibit distinct junction-processing enzymes. *Proc. Natl. Acad. Sci. U.S.A.*, **102**, 6867–6872.
133. Gunderson, C.W. and Segall, A.M. (2006) DNA repair, a novel antibacterial target: holliday junction-trapping peptides induce DNA damage and chromosome segregation defects. *Mol. Microbiol.*, **59**, 1129–1148.
134. Boldt, J.L., Pinilla, C. and Segall, A.M. (2004) Reversible inhibitors of  $\lambda$  Integrase-mediated recombination efficiently trap holliday junction intermediates and form the basis of a novel assay for junction resolution\*. *J. Biol. Chem.*, **279**, 3472–3483.
135. Rideout, M.C., Naili, I., Boldt, J.L., Flores-Fujimoto, A., Patra, S., Rostron, J.E. and Segall, A.M. (2013) wrwrggrywrw is a single-chain functional analog of the holliday junction-binding homodimer, (wrwycr) 2. *Peptides*, **40**, 112–122.
136. Kepple, K.V., Patel, N., Salamon, P. and Segall, A.M. (2008) Interactions between branched DNAs and peptide inhibitors of DNA repair. *Nucleic Acids Res.*, **36**, 5319.
137. Bolla, M.L., Azevedo, E.V., Smith, J.M., Taylor, R.E., Ranjit, D.K., Segall, A.M. and McAlpine, S.R. (2003) Novel antibiotics: macrocyclic peptides designed to trap holliday junctions. *Org. Lett.*, **5**, 109–112.
138. Pan, P.-S., Curtis, F.A., Carroll, C.L., Medina, I., Liotta, L.A., Sharples, G.J. and McAlpine, S.R. (2006) Novel antibiotics: C-2 symmetrical macrocycles inhibiting holliday junction DNA binding by *E. coli* RuvC. *Bioorg. Med. Chem.*, **14**, 4731–4739.
139. Dey, M., Patra, S., Su, L.Y. and Segall, A.M. (2013) Tumor cell death mediated by peptides that recognize branched intermediates of DNA replication and repair. *PLoS One*, **8**, e78751.
140. Bonner, W.M., Redon, C.E., Dickey, J.S., Nakamura, A.J., Sedelnikova, O.A., Solier, S. and Pommier, Y. (2008)  $\gamma$ H2AX and cancer. *Nat. Rev. Cancer*, **8**, 957.
141. Panier, S. and Boulton, S.J. (2014) Double-strand break repair: 53BP1 comes into focus. *Nat. Rev. Mol. Cell Biol.*, **15**, 7–18.
142. Apostolopoulos, V., Bojarska, J., Chai, T.-T., Elnagdy, S., Kaczmarek, K., Matsoukas, J., New, R., Parang, K., Lopez, O.P., Parhiz, H. *et al.* (2021) A global review on short peptides: frontiers and perspectives. *Molecules*, **26**, 430.
143. Ranjit, D.K., Rideout, M.C., Nefzi, A., Ostresh, J.M., Pinilla, C. and Segall, A.M. (2010) Small molecule functional analogs of peptides that inhibit  $\lambda$  site-specific recombination and bind holliday junctions. *Bioorg. Med. Chem. Lett.*, **20**, 4531–4534.
144. Rideout, M.C., Boldt, J.L., Vahi-Ferguson, G., Salamon, P., Nefzi, A., Ostresh, J.M., Giulianotti, M., Pinilla, C. and Segall, A.M. (2011) Potent antimicrobial small molecules screened as inhibitors of tyrosine recombinases and holliday junction-resolving enzymes. *Mol. Diversity*, **15**, 989–1005.
145. Howell, L.A., Bowater, R.A., O'Connell, M.A., Reszka, A.P., Neidle, S. and Searcey, M. (2012) Synthesis of small molecules targeting multiple DNA structures using click chemistry. *ChemMedChem*, **7**, 792–804.
146. Yin, Q., Liu, X., Hu, L., Song, Q., Liu, S., Huang, Q., Geng, Z., Zhu, Y., Li, X., Fu, F. *et al.* (2021) VE-822, a novel DNA holliday junction stabilizer, inhibits homologous recombination repair and triggers DNA damage response in osteogenic sarcomas. *Biochem. Pharmacol.*, **193**, 114767.
147. Kasthuber, E.R. and Lowe, S.W. (2017) Putting p53 in context. *Cell*, **170**, 1062–1078.
148. Maclaine, N.J. and Hupp, T.R. (2011) How phosphorylation controls p53. *Cell Cycle*, **10**, 916–921.
149. Liu, Q., Turner, K.M., Yung, Alfred, Chen, W.K. and Zhang, W. (2014) Role of AKT signaling in DNA repair and clinical response to cancer therapy. *Neuro-Oncology*, **16**, 1313–1323.
150. Willmore, E., de Caux, S., Sunter, N.J., Tilby, M.J., Jackson, G.H., Austin, C.A. and Durkacz, B.W. (2004) A novel DNA-dependent protein kinase inhibitor, NU7026, potentiates the cytotoxicity of topoisomerase II poisons used in the treatment of leukemia. *Blood*, **103**, 4659–4665.
151. Kato, T., Sato, N., Hayama, S., Yamabuki, T., Ito, T., Miyamoto, M., Kondo, S., Nakamura, Y. and Daigo, Y. (2007) Activation of holliday junction-recognizing protein involved in the chromosomal stability and immortality of cancer cells. *Cancer Res.*, **67**, 8544–8553.
152. Fokas, E., Prevo, R., Pollard, J.R., Reaper, P.M., Charlton, P.A., Cornelissen, B., Vallis, K.A., Hammond, E.M., Olcina, M.M., Gillies McKenna, W. *et al.* (2012) Targeting ATR in vivo using the novel inhibitor VE-822 results in selective sensitization of pancreatic tumors to radiation. *Cell Death. Dis.*, **3**, e441.
153. O'Connor, M.J. (2015) Targeting the DNA damage response in cancer. *Mol. Cell*, **60**, 547–560.
154. Pilié, P.G., Tang, C., Mills, G.B. and Yap, T.A. (2019) State-of-the-art strategies for targeting the DNA damage response in cancer. *Nat. Rev. Clin. Oncol.*, **16**, 81–104.
155. Duskova, K., Lejault, P., Benchimol, É., Guillot, R., Britton, S., Granzhan, A. and Monchaud, D. (2020) DNA junction ligands trigger DNA damage and are synthetic lethal with DNA repair inhibitors in cancer cells. *J. Am. Chem. Soc.*, **142**, 424–435.
156. Zell, J., Duskova, K., Chouh, L., Bossaert, M., Chéron, N., Granzhan, A., Britton, S. and Monchaud, D. (2021) Dual targeting of higher-order DNA structures by azacryptands induces DNA junction-mediated DNA damage in cancer cells. *Nucleic Acids Res.*, **49**, 10275–10288.
157. Hall, C.E. and Cavalieri, L.F. (1961) Four-stranded DNA as determined by electron microscopy. *J. Cell Biol.*, **10**, 347–351.
158. Kozinski, A.W. and Beer, M. (1962) Effect of concentration on the formation of molecular hybrids from T4 DNA. *Biophys. J.*, **2**, 129–141.
159. Richardson, C.C., Inman, R. and Kornberg, A. (1964) Enzymic synthesis of deoxyribonucleic acid: XVIII. The repair of partially single-stranded DNA templates by DNA polymerase. *J. Mol. Biol.*, **9**, 46–69.

160. Dawid, I.B. and Wolstenholme, D.R. (1968) Renaturation and hybridization studies of mitochondrial DNA. *Biophys. J.*, **8**, 65–81.
161. Schildkraut, C.L., Richardson, C.C. and Kornberg, A. (1964) Enzymic synthesis of deoxyribonucleic acid: XVII. Some unusual physical properties of the product primed by native DNA templates. *J. Mol. Biol.*, **9**, 24–45.
162. Jensch, F. and Kemper, B. (1986) Endonuclease VII resolves Y-junctions in branched DNA in vitro. *EMBO J.*, **5**, 181–189.
163. Kadrmas, J.L., Ravin, A.J. and Leontis, N.B. (1995) Relative stabilities of DNA three-way, four-way and five-way junctions (multi-helix junction loops): unpaired nucleotides can be stabilizing or destabilizing. *Nucleic Acids Res.*, **23**, 2212–2222.
164. Leonard, C.J. and Berns, K.I. (1994) In: Cohn, W.E. and Moldave, K. (eds). *Prog. Nucleic Acid Res. Mol. Biol.* Academic Press, Vol. **48**, pp. 29–52.
165. Khristich, A.N. and Mirkin, S.M. (2020) On the wrong DNA track: molecular mechanisms of repeat-mediated genome instability. *J. Biol. Chem.*, **295**, 4134–4170.
166. Stojanovic, M.N., de Prada, P. and Landry, D.W. (2000) Fluorescent sensors based on aptamer self-assembly. *J. Am. Chem. Soc.*, **122**, 11547–11548.
167. Stojanovic, M.N., De Prada, P. and Landry, D.W. (2001) Aptamer-based folding fluorescent sensor for cocaine. *J. Am. Chem. Soc.*, **123**, 4928–4931.
168. Stojanovic, M.N. and Landry, D.W. (2002) Aptamer-based colorimetric probe for cocaine. *J. Am. Chem. Soc.*, **124**, 9678–9679.
169. Kato, T., Yano, K., Ikebukuro, K. and Karube, I. (2000) Interaction of three-way DNA junctions with steroids. *Nucleic Acids Res.*, **28**, 1963–1968.
170. Stojanović, M.N., Green, E.G., Semova, S., Nikić, D.B. and Landry, D.W. (2003) Cross-reactive arrays based on three-way junctions. *J. Am. Chem. Soc.*, **125**, 6085–6089.
171. Oleksi, A., Blanco, A.G., Boer, R., Usón, I., Aymami, J., Rodger, A., Hannon, M.J. and Coll, M. (2006) Molecular recognition of a three-way DNA junction by a metallosupramolecular helicate. *Angew. Chem. Int. Ed.*, **45**, 1227–1231.
172. Boer, D.R., Kerckhoffs, J.M., Parajo, Y., Pascu, M., Usón, I., Lincoln, P., Hannon, M.J. and Coll, M. (2010) Self-Assembly of functionalizable two-component 3D DNA arrays through the induced formation of DNA three-way-junction branch points by supramolecular cylinders. *Angew. Chem. Int. Ed.*, **49**, 2336–2339.
173. Malina, J., Hannon, M.J. and Brabec, V. (2007) Recognition of DNA three-way junctions by metallosupramolecular cylinders: gel electrophoresis studies. *Chem. Eur. J.*, **13**, 3871–3877.
174. Cerasino, L., Hannon, M.J. and Sletten, E. (2007) DNA three-way junction with a dinuclear iron (II) supramolecular helicate at the center: a NMR structural study. *Inorg. Chem.*, **46**, 6245–6251.
175. Cardo, L., Sadovnikova, V., Phongtongpasuk, S., Hodges, N.J. and Hannon, M.J. (2011) Arginine conjugates of metallo-supramolecular cylinders prescribe helicity and enhance DNA junction binding and cellular activity. *Chem. Commun.*, **47**, 6575–6577.
176. Gamba, I., Rama, G., Ortega-Carrasco, E., Maréchal, J.-D., Martínez-Costas, J., Vázquez, M.E. and López, M.V. (2014) Programmed stereoselective assembly of DNA-binding helical metallopeptides. *Chem. Commun.*, **50**, 11097–11100.
177. Gómez-González, J., Pérez, Y., Sciortino, G., Roldan-Martín, L., Martínez-Costas, J., Maréchal, J.-D., Alfonso, I., Vázquez López, M. and Vázquez, M.E. (2021) Dynamic stereoselection of peptide helicates and their selective labeling of DNA replication foci in cells. *Angew. Chem. Int. Ed.*, **60**, 8859–8866.
178. Vuong, S., Stefan, L., Lejault, P., Rousselin, Y., Denat, F. and Monchaud, D. (2012) Identifying three-way DNA junction-specific small-molecules. *Biochimie*, **94**, 442–450.
179. Novotna, J., Laguerre, A., Granzhan, A., Pirrotta, M., Teulade-Fichou, M.-P. and Monchaud, D. (2015) Cationic azacryptands as selective three-way DNA junction binding agents. *Org. Biomol. Chem.*, **13**, 215–222.
180. Stefan, L., Bertrand, B., Richard, P., Le Gendre, P., Denat, F., Picquet, M. and Monchaud, D. (2012) Assessing the differential affinity of small molecules for noncanonical DNA structures. *ChemBioChem*, **13**, 1905–1912.
181. Barros, S.A. and Chenoweth, D.M. (2014) Recognition of nucleic acid junctions using triptycene-based molecules. *Angew. Chem. Int. Ed.*, **53**, 13746–13750.
182. Barros, S.A. and Chenoweth, D.M. (2015) Triptycene-based small molecules modulate (CAG)-(CTG) repeat junctions. *Chem. Sci.*, **6**, 4752–4755.
183. Guyon, L., Pirrotta, M., Duskova, K., Granzhan, A., Teulade-Fichou, M.-P. and Monchaud, D. (2018) TWJ-Screen: an isothermal screening assay to assess ligand/DNA junction interactions in vitro. *Nucleic Acids Res.*, **46**, e16.
184. Duskova, K., Lamarche, J., Amor, S., Caron, C., Queyriaux, N., Gaschard, M., Penouilh, M.-J., de Robillard, G., Delmas, D., Devillers, C.H. *et al.* (2019) Identification of three-way DNA junction ligands through screening of chemical libraries and validation by complementary in vitro assays. *J. Med. Chem.*, **62**, 4456–4466.
185. Zhu, J., Haynes, C.J.E., Kieffer, M., Greenfield, J.L., Greenhalgh, R.D., Nitschke, J.R. and Keyser, U.F. (2019) FeII4L4 tetrahedron binds to nonpaired DNA bases. *J. Am. Chem. Soc.*, **141**, 11358–11362.
186. Yang, Z., Chen, Y., Li, G., Tian, Z., Zhao, L., Wu, X., Ma, Q., Liu, M. and Yang, P. (2018) Supramolecular recognition of three way junctions DNA by a cationic calix [3]carbazole. *Chem. Eur. J.*, **24**, 6087–6093.
187. Hotze, A.C., Hodges, N.J., Hayden, R.E., Sanchez-Cano, C., Paines, C., Male, N., Tse, M.-K., Bunce, C.M., Chipman, J.K. and Hannon, M.J. (2008) Supramolecular iron cylinder with unprecedented DNA binding is a potent cytostatic and apoptotic agent without exhibiting genotoxicity. *Chem. Biol.*, **15**, 1258–1267.
188. Pope, A.J., Bruce, C., Kysela, B. and Hannon, M.J. (2010) Issues surrounding standard cytotoxicity testing for assessing activity of non-covalent DNA-binding metallo-drugs. *Dalton Trans.*, **39**, 2772–2774.
189. Brabec, V., Howson, S.E., Kaner, R.A., Lord, R.M., Malina, J., Phillips, R.M., Abdallah, Q.M.A., McGowan, P.C., Rodger, A. and Scott, P. (2013) Metallohelices with activity against cisplatin-resistant cancer cells: does the mechanism involve DNA binding? *Chem. Sci.*, **4**, 4407–4416.
190. Liu, B., Lin, J. and Steitz, T.A. (2013) Structure of the polii $\alpha$ - $\tau$  c-DNA complex suggests an atomic model of the replisome. *Structure*, **21**, 658–664.
191. Parajo, Y., Malina, J., Meistermann, I., Clarkson, G.J., Pascu, M., Rodger, A., Hannon, M.J. and Lincoln, P. (2009) Effect of bridging ligand structure on the thermal stability and DNA binding properties of iron (II) triple helicates. *Dalton Trans.*, 4868–4874.
192. Yu, H., Wang, X., Fu, M., Ren, J. and Qu, X. (2008) Chiral metallo-supramolecular complexes selectively recognize human telomeric G-quadruplex DNA. *Nucleic Acids Res.*, **36**, 5695–5703.
193. Zhao, C., Geng, J., Feng, L., Ren, J. and Qu, X. (2011) Chiral metallo-supramolecular complexes selectively induce human telomeric G-Quadruplex formation under salt-deficient conditions. *Chem. Eur. J.*, **17**, 8209–8215.
194. Ducani, C., Leczkowska, A., Hodges, N.J. and Hannon, M.J. (2010) Noncovalent DNA-Binding metallo-supramolecular cylinders prevent DNA transactions in vitro. *Angew. Chem. Int. Ed.*, **49**, 8942–8945.
195. Gómez-González, J., Martínez-Castro, L., Tolosa-Barrilero, J., Alcalde-Ordóñez, A., Learte-Aymami, S., Mascareñas, J.L., García-Martínez, J.C., Martínez-Costas, J., Maréchal, J.-D. and López, M.V. (2022) Selective recognition of A/T-rich DNA 3-way junctions with a three-fold symmetric tripeptide. *Chem. Commun.*, **58**, 7769–7772.
196. McLuckie, K.I.E., Di Antonio, M., Zecchini, H., Xian, J., Caldas, C., Krippendorff, B.F., Tannahill, D., Lowe, C. and Balasubramanian, S. (2013) G-Quadruplex DNA as a molecular target for induced synthetic lethality in cancer cells. *J. Am. Chem. Soc.*, **135**, 9640–9643.
197. Wang, J.C. (2002) Cellular roles of DNA topoisomerases: a molecular perspective. *Nat. Rev. Mol. Cell Biol.*, **3**, 430–440.
198. Szlachta, K., Manukyan, A., Raimer, H.M., Singh, S., Salamon, A., Guo, W., Lobachev, K.S. and Wang, Y.-H. (2020) Topoisomerase II contributes to DNA secondary structure-mediated double-stranded breaks. *Nucleic Acids Res.*, **48**, 6654–6671.
199. Jonstrup, A.T., Thomsen, T., Wang, Y., Knudsen, B.R., Koch, J. and Andersen, A.H. (2008) Hairpin structures formed by alpha satellite DNA of human centromeres are cleaved by human topoisomerase II $\alpha$ . *Nucleic Acids Res.*, **36**, 6165–6174.

200. Thakur, J., Packiaraj, J. and Henikoff, S. (2021) Sequence, chromatin and evolution of satellite DNA. *Int. J. Mol. Sci.*, **22**, 4309.
201. Chardon, F., Japaridze, A., Witt, H., Velikovskiy, L., Chakraborty, C., Wilhelm, T., Dumont, M., Yang, W., Kikuti, C. and Gangnard, S. (2022) CENP-B-mediated DNA loops regulate activity and stability of human centromeres. *Mol. Cell*, **82**, 1751–1767.
202. Cañeque, T., Müller, S. and Rodriguez, R. (2018) Visualizing biologically active small molecules in cells using click chemistry. *Nat. Rev. Chem.*, **2**, 202–215.
203. Williamson, J.R., Raghuraman, M. and Cech, T.R. (1989) Monovalent cation-induced structure of telomeric DNA: the G-quartet model. *Cell*, **59**, 871–880.
204. Sundquist, W.I. and Klug, A. (1989) Telomeric DNA dimerizes by formation of guanine tetrads between hairpin loops. *Nature*, **342**, 825.
205. Sun, D.Y., Thompson, B., Cathers, B.E., Salazar, M., Kerwin, S.M., Trent, J.O., Jenkins, T.C., Neidle, S. and Hurley, L.H. (1997) Inhibition of human telomerase by a G-quadruplex-interactive compound. *J. Med. Chem.*, **40**, 2113–2116.
206. Wang, Y.-H., Yang, Q.-F., Lin, X., Chen, D., Wang, Z.-Y., Chen, B., Han, H.-Y., Chen, H.-D., Cai, K.-C., Li, Q. *et al.* (2021) G4LDB 2.2: a database for discovering and studying G-quadruplex and i-Motif ligands. *Nucleic Acids Res.*, **50**, D150–D160.
207. De Cian, A., DeLemos, E., Mergny, J.-L., Teulade-Fichou, M.-P. and Monchaud, D. (2007) Highly efficient G-quadruplex recognition by bisquinolinium compounds. *J. Am. Chem. Soc.*, **129**, 1856–1857.
208. Rodriguez, R., Mueller, S., Yeoman, J.A., Trentesaux, C., Riou, J.-F. and Balasubramanian, S. (2008) A novel small molecule that alters shelterin integrity and triggers a DNA-Damage response at telomeres. *J. Am. Chem. Soc.*, **130**, 15758–15758.
209. Raguseo, F., Chowdhury, S., Minard, A. and Di Antonio, M. (2020) Chemical-biology approaches to probe DNA and RNA G-quadruplex structures in the genome. *Chem. Commun.*, **56**, 1317–1324.
210. Huppert, J.L. and Balasubramanian, S. (2005) Prevalence of quadruplexes in the human genome. *Nucleic Acids Res.*, **33**, 2908–2916.
211. Bedrat, A., Lacroix, L. and Mergny, J.-L. (2016) Re-evaluation of G-quadruplex propensity with G4Hunter. *Nucleic Acids Res.*, **44**, 1746–1759.
212. Puig Lombardi, E. and Londoño-Vallejo, A. (2019) A guide to computational methods for G-quadruplex prediction. *Nucleic Acids Res.*, **48**, 1–15.
213. Strawbridge, E.M., Benson, G., Gelfand, Y. and Benham, C.J. (2010) The distribution of inverted repeat sequences in the *Saccharomyces cerevisiae* genome. *Curr. Genet.*, **56**, 321–340.
214. Čutová, M., Manta, J., Porubiaková, O., Kaura, P., Štastný, J., Jagelská, E.B., Goswami, P., Bartas, M. and Brázda, V. (2020) Divergent distributions of inverted repeats and G-quadruplex forming sequences in *Saccharomyces cerevisiae*. *Genomics*, **112**, 1897–1901.
215. Meier-Stephenson, V. (2022) G4-quadruplex-binding proteins: review and insights into selectivity. *Biophys. Rev.*, **14**, 635–654.