pharmaceuticals

MDPI

Article

Structured Waters Mediate Small Molecule Binding to G-Quadruplex Nucleic Acids

Stephen Neidle

The School of Pharmacy, University College London, 29-39 Brunswick Square, London WC1N 1AX, UK; s.neidle@ucl.ac.uk

Abstract: The role of G-quadruplexes in human cancers is increasingly well-defined. Accordingly, G-quadruplexes can be suitable drug targets and many small molecules have been identified to date as G-quadruplex binders, some using computer-based design methods and co-crystal structures. The role of bound water molecules in the crystal structures of G-quadruplex-small molecule complexes has been analyzed in this study, focusing on the water arrangements in several G-quadruplex ligand complexes. One is the complex between the tetrasubstituted naphthalene diimide compound MM41 and a human intramolecular telomeric DNA G-quadruplex, and the others are in substituted acridine bimolecular G-quadruplex complexes. Bridging water molecules form most of the hydrogen-bond contacts between ligands and DNA in the parallel G-quadruplex structures examined here. Clusters of structured water molecules play essential roles in mediating between ligand side chain groups/chromophore core and G-quadruplex. These clusters tend to be conserved between complex and native G-quadruplex structures, suggesting that they more generally serve as platforms for ligand binding, and should be taken into account in docking and in silico studies.

Keywords: G-quadruplex; ligands; naphthalene diimides; acridines; crystal structures; water clusters; hydrogen bonding

Citation: Neidle, S. Structured Waters Mediate Small Molecule Binding to G-Quadruplex Nucleic Acids. *Pharmaceuticals* **2022**, *15*, 7. https://doi.org/10.3390/ph15010007

Academic Editor: Alfredo Berzal-Herranz

Received: 29 November 2021
Accepted: 20 December 2021
Published: 22 December 2021

Publisher's Note: MDPI stays neutral with regard to jurisdictional claims in published maps and institutional affiliations.

1. Introduction

G-quadruplexes (GQs) are higher-order four-stranded structures that can form in DNA and RNA sequences by the folding of repetitive short guanine (G)-tracts [1–4]. These are typically interspersed with short runs of general sequence. The G-tracts self-associate by Hoogsteen hydrogen bonding to form guanine (G)-quartets, several of which stack on one another to form a G-quadruplex (GQ). The quartets are held together by the phosphodiester backbones, and by the general sequences, which typically form extra-helical loops. GQ folding and hence the nature of the loop can occur in several ways, such that the four strands can be all-parallel, all-anti-parallel, or various combinations [2,5]. GQ prevalence in the human genome is non-random [6,7]. In eukaryotic cells, they occur in telomeres [8,9] and are over-represented in genomic promoter [10–14] and untranslated sequences [15–17], especially in genes and pathways involved in cancer initiation and progression [14,18–22]. GQ-forming sequences have been found in the genomes of many organisms, ranging from viruses [23–27] to bacteria [28–31] and malaria [32–34]. GQs have been visualized in fixed [35–38] and in live cells [39–42], where their existence may be more than transient, with several roles in gene function [3,4,12]; for example, GQ folding is associated with sites of active transcription and precedes transcription itself [13,22].

The presence of GQ sequences in the promoters of oncogenes, such as h*TERT*, *MYC*, *KRAS*, *BCL2*, and c-*KIT* as well as their potential for destabilizing telomere maintenance in cancer cells (and their involvement in replication and genomic instability [12,43–45]), has focused attention on GQs as promoter or telomeric therapeutic targets in viral and bacterial diseases as well as human cancers. The strategy of stabilizing them with appropriate small-molecule compounds has resulted in many chemically diverse chemotypes being

investigated, notably against oncogene GQ promoter targets [12,46,47]. Formation of high-affinity GQ-ligand complexes within promoters has been demonstrated to reduce or even abolish transcription of a target gene or genes, see for example refs. [48–58], which can result in cell growth inhibition and in vivo anti-cancer activity in animal models of human cancers [59].

Most of these GQ-binding compounds share common structural features of planar heteroaromatic groups and side chains carrying cationic groups, albeit in a wide variety of chemotypes [60–66]. Structure-activity studies have optimized activity for numerous series of these compounds, for for example several libraries of acridine [67–76] and naphthalene diimide [18,19,77–94] derivatives. The lead naphthalene diimide derivatives CM03 and SOP1812 (Figure 1) show evidence of target engagement and in vivo anti-tumor activity [18,19]. Crystallographic and NMR studies have provided detailed information on GQ-ligand interactions in these and other GQ-ligand complexes [50,51,71,74–76,80,81,95–100], some of which have been used in several hit-to-lead optimization projects [18,19,50,81,90,97,99]. Quantitative and semi-quantitative computer modeling methods including docking procedures [101–111] have also been extensively employed to screen virtual libraries, with the aims of aiding optimization and identifying plausible new chemotypes and lead compounds for future drug development and eventual clinical trial.

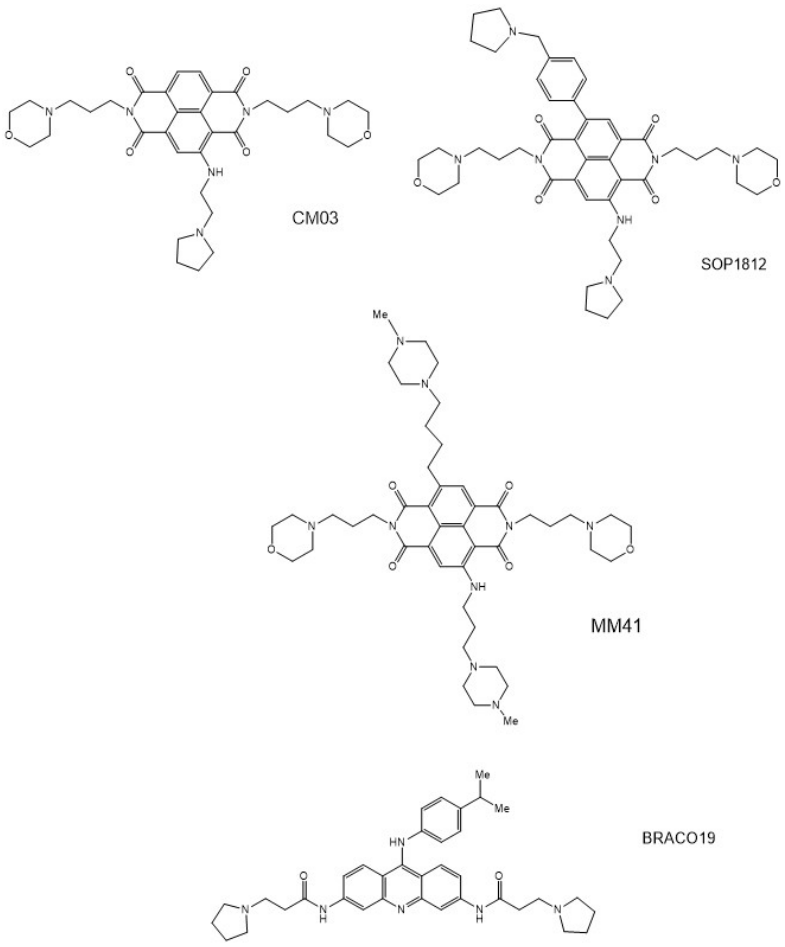


Figure 1. Structures of the tri-substituted naphthalene diimide derivatives (CM03), two tetra-substituted derivatives (MM41 and SOP1812), and the trisubstituted acridine compound BRACO19. CM03: 2,7-bis(3-morpholinopropyl)-4-((2-(pyrrolidin-1-yl)ethyl)amino) benzo[lmn][3,8] phenanthroline-1,3,6,8(2H,7H)-tetraone; SOP1812: 2,7-bis(3-morpholinopropyl)- 4-((2-(pyrrolidin-1-yl)ethyl)amino)-9-(4-(pyrrolidin-1-ylmethyl)phenyl) benzo[lmn][3,8] phenanthroline-1,3,6,8-(2H,7H)-tetraone); MM41: (4,9-bis((3-(4-methylpiperazin-1-yl)propyl)amino)-2,7-bis(3-morpholinopropyl) benzo[*lmn*][3,8] phenanthroline-1,3,6,8(2H,7H)-tetraone); BRACO19: 3,6-bis(3-pyrrolidin-1-ylpropionamido)-9-(4-dimethylaminophenylamino) acridine.

Several high-resolution crystallographic studies of native GQs [112–115] and GQ-ligand complexes [71,74,75,80,81,98,100,116,117] have reported that water molecules are intimately associated with GQ sites and with bound ligand. It is well-established for DNA-, RNA- [118], and protein-ligand complexes [119–123] that water molecules can play critical target-ligand mediation roles. The present work analyzes water arrangements in several GQ-ligand complexes (PDB id 3CE5, 3NZ7, and 3UYH) [74,81,100] that are of sufficient resolution to give confidence in the significance of the arrangements, to better understand the role and the consequences of discrete water molecules associated with ligand binding. The naphthalene diimide complex [81] was also chosen since it involves the tetra-substituted compound MM41 (Figure 1). This is the direct precursor of two other more recently designed naphthalene diimide compounds, CM03 and SOP1812 [18,19] (Figure 1), which are currently in pre-clinical development stages. Hydration features of this complex and other complexes examined here, it is suggested, have wider implications for GQ-based drug design and hit/lead selection.

2. Results

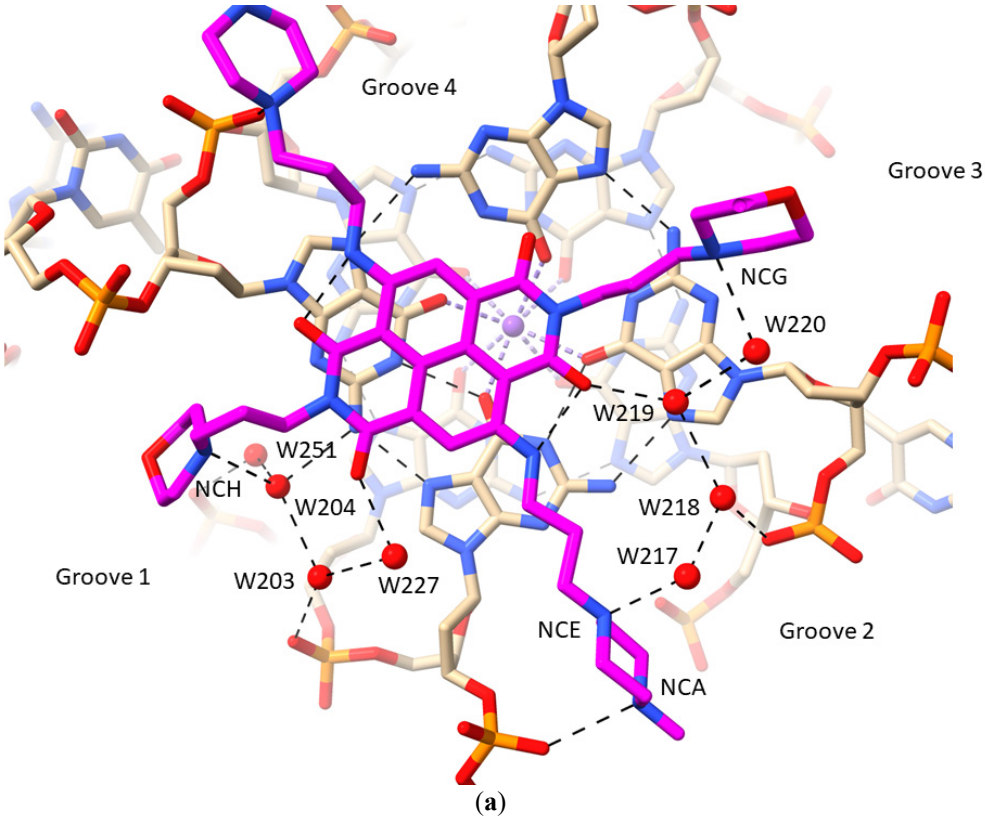
Only the crystal structures containing acridine, berberine, and naphthalene diimide (ND) derivatives fulfilled the acceptance criteria summarized in the preceding section. Two co-crystal structures are available (Table 1) for the tetrasubstituted naphthalene diimide compound MM41 (Figure 1) complexed to intramolecular human telomeric GQs, for which it has high binding affinity. Structure PDB id 3UYH is at the higher resolution of the two and consequently a greater number of ligand-associated water molecules were observed in electron density maps and included in the final refined crystal structure [81]. Hence it was chosen for further detailed analysis (Tables 1 and 2). Structure 3CDM [124] has 2 G-quadruplexes, 158 water molecules, and 4 substituted naphthalene diimides in the asymmetric unit, i.e., 79 waters per G-quadruplex. However, few water molecules in this structure are resolved in the vicinity of the two stacked naphthalene diimide ligands compared to structure 3UYH and so this structure was not chosen for detailed analysis. In addition, the nature of naphthalene diimide substituents in 3CDM is not directly relevant to MM41 and hence not to CM03 or SOP1812, so this structure was not considered any further in the present analysis.

2.1. MM41 Side Chain Contacts and Water Environment

MM41 has two side chains terminating in N-methyl-piperazine groups and two with terminal morpholino groups. Each of these groups can be assumed to be protonated at physiological pH, with N-methyl-piperazine having a pK of 8.5 compared to the slightly less basic morpholino group, with a pK of 9.2 [81]. Figure 2a shows a view of structure 3UYH projected onto the planes of the G-quartets and the naphthalene diimide core, highlighting the grooves of the GQ. Each MM41 side chain is positioned in or close to the mouth of a GQ groove, although only three of the four end groups are actually situated within a groove. The fourth, having a terminal morpholino ring, is oriented away from the quartet plane and a detailed examination of the crystal structure has indicated that rotation of the side chain to place the morpholino group into groove 4 is sterically hindered by the small surface area of the naphthalene diimide core compared to that of the quartet [81].

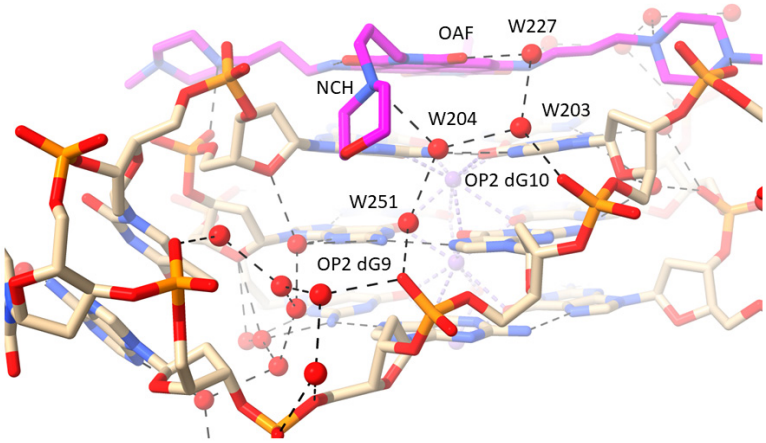
Table 1. G-quadruplex-small molecule crystal structures, taken from the Protein Data Bank, for which at the minimum a first shell of water molecules around the DNA and ligand have been reported. The number of water molecules associated with each complete G-quadruplex is quoted, as taken from the PDB entry. The structures discussed here (PDB ids 3CE5, 3NZ7 and 3UYH) are highlighted in bold.

PDB Id	G-Quadruplex Type	Compound	Resoln (Å)	No. of Waters/AU	Ref.
3CE5	**12-mer bimolecular human telomeric**	**3,6,9- trisubstituted acridine BRACO19**	**2.5**	**54**	**74**
3NZ7	**12-mer bimolecular *Oxytricha nova* telomeric**	**3,6- disubstituted acridine, F substituents**	**1.10**	**187**	**100**
3NYP	12-mer bimolecular *Oxytricha nova* telomeric	3,6- disubstituted acridine, F substituents	1.18	176	100
3EM2	12-mer bimolecular *Oxytricha nova* telomeric	3,6- disubstituted acridine	2.3	64	75
3EQW	12-mer bimolecular *Oxytricha nova* telomeric	3,6- disubstituted acridine	2.2	66	75
3EUI	12-mer bimolecular *Oxytricha nova* telomeric	3,6- disubstituted acridine	2.2	159	75
3ERU	12-mer bimolecular *Oxytricha nova* telomeric	3,6- disubstituted acridine	2.0	71	75
3ES0	12-mer bimolecular *Oxytricha nova* telomeric	3,6- disubstituted acridine	2.2	56	75
3ET8	12-mer bimolecular *Oxytricha nova* telomeric	3,6- disubstituted acridine	2.45	51	75
3EUM	12-mer bimolecular *Oxytricha nova* telomeric	3,6- disubstituted acridine	1.78	52	75
1L1H	12-mer bimolecular *Oxytricha nova* telomeric	3,6- disubstituted acridine	1.75	146	75
3UYH	**22-mer human telomeric**	**Tetrasubstituted naphthalene diimide MM41**	**1.95**	**51**	**81**
3T5E	22-mer human telomeric	Tetrasubstituted naphthalene diimide BMSG-SH-4	2.10	38	80
3CCO	11-mer biomolecular human telomeric	Tetrasubstituted naphthalene diimide	2.20	28	124
3CDM	22-mer human telomeric	Tetrasubstituted naphthalene diimide	2.10	158	124
4DA3	21-mer human telomeric	Tetrasubstituted naphthalene diimide MM41	2.40	25	81
6S15	12-mer bimolecular human telomeric	Pyridine derivative of berberine	1.70	23	98

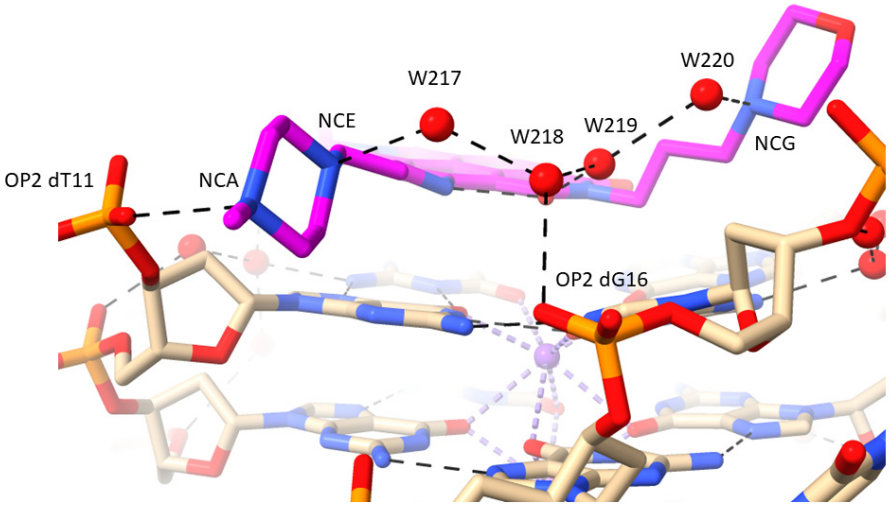


(a)

Figure 2. *Cont.*



(**b**)



(**c**)

Figure 2. Views of the crystal structure of MM41 (with its carbon atoms colored magenta) bound to a human intramolecular telomeric G-quadruplex, PDB id 3UYH. (**a**) The view is projected onto the G-quartet plane and shows the extent of overlap with the naphthalene diimide core chromophore. The water molecules that are in direct or indirect contact with the MM41 molecule, are shown as red spheres, with hydrogen bonds indicated by dashed lines. This and the subsequent figures were drawn using the ChimeraX package (https://www.cgl.ucsf.edu/chimerax/, last accessed on 16 December 2021) [125]. (**b**) A view of the 3UYH complex looking into groove 1. The four water molecules are shown that are hydrogen bonded to the morpholino group in this groove and the OAF carbonyl oxygen atom of the naphthalene diimide core. A nearby cluster of water molecules is also shown, embedded deep in the groove and adjacent to a TTA loop. (**c**) A view of the 3UYH complex midway between grooves 2 and 3, highlighting the group of four water molecules hydrogen bonding to an N-methyl-piperazine and a morpholino group in these grooves.

Table 2. Hydrogen bond interactions. (**a**) In structure 3UYH, involving the tetrasubstituted naphthalene diimide MM41, a human intramolecular telomeric G-quadruplex, and water molecules. Hydrogen-bond distances are shown ($d_{1\text{-}2}$ in Å), together with the reported crystallographic B factor values (in $Å^2$) for MM41-bound waters and associated MM41 and DNA atoms. MM41 atoms are highlighted in bold red type. Waters in direct contact with MM41 atoms are highlighted in bold blue type. Numbering is as in the PDB entry. (**b**) In structure PDB id 3CE5 [74], involving the trisubstituted compound BRACO19, a human intermolecular bimolecular telomeric G-quadruplex, and water molecules. Parameter definitions and color coding are as in Table 2a.

		(a)			
	$Atom_1$	**$Atom_2$**	**$d_{1\text{-}2}$**	**B Factor $Atom_1$**	**B Factor $Atom_2$**
Groove 1					
	NCH	**W204**	3.2	**64**	**32**
	W204	W203	3.0	**32**	32
	W203	OP2 dG10	2.9	32	32
	W203	**W227**	3.4	32	**45**
	W227	**OAF**	2.5	**45**	**29**
	W251	OP2 dG9	2.9	27	30
	W204	N2 dG4	2.9	**32**	20
	W204	W251	2.8	**32**	27
Groove 2					
	NCA	OP2 dT11	3.1	**46**	35
	NCE	**W217**	2.7	**40**	**46**
	W217	W218	2.6	**46**	48
	W218	**W219**	3.0	48	**43**
	W219	**ODX**	2.7	**43**	**31**
	W218	OP2 dG16	3.4	48	45
Groove 3					
	NCG	**W220**	2.9	**48**	**50**
	W219	**W220**	2.9	**43**	**50**
Groove 4					
	NCF	OP2 dG4	2.9	**58**	39
		(b)			
	$Atom_1$	**$Atom_2$**	**$d_{1\text{-}2}$**	**B Factor $Atom_1$**	**B Factor $Atom_2$**
Groove 1					
	N7	**W52**	2.7	**11**	**21**
	W52	N3 dT24	2.9	**21**	12
	W52	O6 dG5	3.0	**21**	18
	W52	**W53**	3.1	**21**	**33**
	W53	O2 dT24	3.3	**33**	18
	W53	**O52**	3.1	**33**	**14**
Groove 2					
	W56	**N39**	3.4	**23**	**27**
	N17	**W55**	3.0	**14**	**38**
	W55	O2 dT12	3.3	**38**	17
Groove 4					
	N21	O4 dT24	3.0	13	17
	N47	**W41**	3.0	**21**	**28**
	W41	W44	2.9	**28**	29
	W44	OP2 dG23	2.5	29	21
	W41	N2 dG17	2.9	**28**	18

All four basic end groups of MM41 have their protonated nitrogen atoms in hydrogen bond/electrostatic contact with atoms in the GQ grooves (Table 2). However, only two of

these contacts, each involving the terminal nitrogen atom of an N-methyl-piperazine group, has a direct nitrogen-phosphate group hydrogen bond interaction (N . . . OP distances of 2.9 and 3.1 Å). The three other end groups all have water contacts with ring nitrogen atoms (Figure 2b,c), which in the case of the morpholino groups, are presumed to be protonated. Groove 1 (Figure 2b) has the morpholino group positioned at the mouth of the groove. A small linear cluster of four water molecules extends from the morpholino basic nitrogen atom, with one water contacting two further waters, which contact with two neighboring phosphate oxygen atoms. One of these waters OP2 dG10 also contacts a fourth water molecule, which in turn contacts the adjacent oxygen substituent on the naphthalene diimide chromophore. The other water, contacting a phosphate oxygen atom (OP2 G9), also contacts and is thus the link to a water molecule in a second water network that fills the rest of this groove. The N-methyl-piperazine group in groove 2 (Figure 2c), which is situated at the mouth of the groove, has its terminal nitrogen atom (NCA) in close contact with a phosphate oxygen atom, OP2 dT11, suggesting that this nitrogen atom carries a proton. The inner piperazine ring nitrogen atom contacts another linear group of four water molecules which extends into groove 3 and terminates with a hydrogen bond contact with the second morpholino group. The second water in this array has a contact with a phosphate oxygen atom (OP2dG16), and the third is in hydrogen bond contact with carbonyl oxygen atom OAG of the ND core. The second N-methyl-piperazine ring, also situated at the mouth of groove 4, has a direct contact involving the outer ring nitrogen and a phosphate oxygen atom (Figure 2a), but does not have any associated water molecules.

The above section describes water molecules involved in ligand contacts; other waters fill out the remaining space in the grooves (the non-completeness in some grooves is most likely due to limitations of the crystal structure at 1.95 Å with only a fraction of the potential total number of water molecules located in electron density). Groove 2 is one of the more completely resolved grooves in terms of hydration (Figure 3), with waters embedded deep into the groove. This complex array of 12 water molecules, of which the majority are first-shell, hydrogen bond with phosphate oxygens, O4′ and O5′ atoms, and guanine base edges (which form the floor of the groove). The net effect is to maintain the relative positions of TTA loop and the groove.

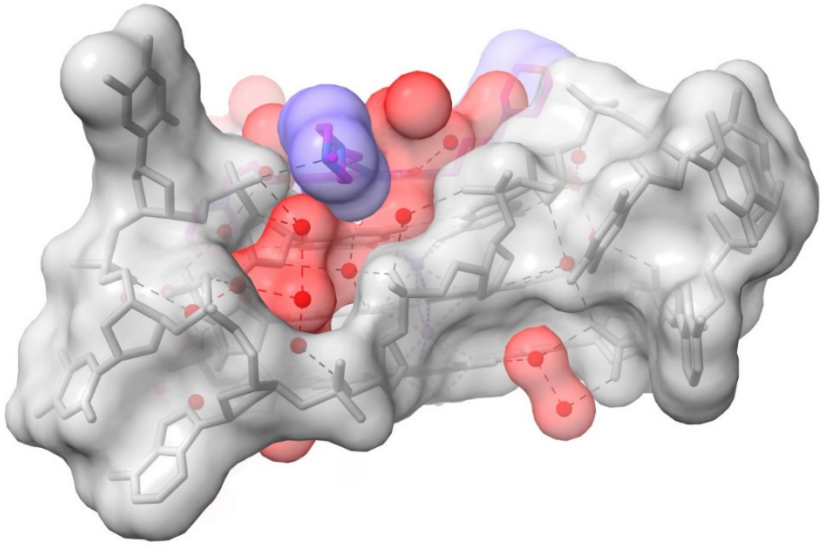


Figure 3. Surface representation of a view into groove 2 in the 3UYH complex. The N-methyl-piperazine substituent of MM41 is oriented end-on and is colored blue. Note that the groove space is filled out by water molecules, colored red. The semi-transparent surface of the G-quadruplex is colored grey.

2.2. MM41 and Water Mobility

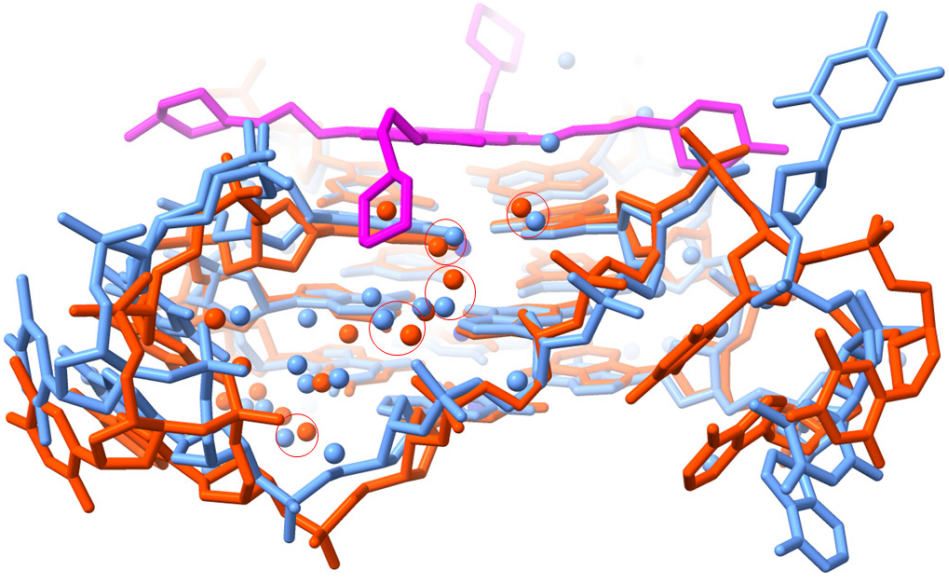
It is notable that there is remarkably little overlap between the ND four-ring core and the individual guanines in the top quartet, as seen in Figure 2a. The surface area of the ND core of MM41 is too small to allow simultaneous overlap with more than one guanine of the top G-quartet. Computational experiments were undertaken using two different energy minimization protocols (in the ARGUSLAB and AVOGADRO packages) to assess the effect of removing all the water molecules from the structure. Minimization in both cases resulted in movement of the ND core by ca 1.5 Å to produce improved overlap with a guanine base of the G-quartet. The cationic groups in the side chains tended to move closer to the phosphate oxygen atoms. Attempts to dock the MM41 molecule onto the water-free GQ using AutoDock Vina 1.1.2 as installed within the database G4LDB 2.2, resulted in a series of almost equi-energetic poses in which three out of four side chains were positioned away from the grooves. This was not pursued further.

The side chain heterocyclic groups in MM41 have greater mobility than the ND core, as revealed by their individual atomic temperature factors (see the PDB entry for 3UYH and Table 2). The five nitrogen atoms in these terminal rings, which are hydrogen-bonded to waters or phosphate groups, have a mean B value of 51 $\bar{\text{Å}}^2$, corresponding to a <U> of 0.8 $\bar{\text{Å}}$. The water molecules in the mini cluster around and contacting the morpholino ring in groove 1 have lower B values, with most in the range 27–32 $\bar{\text{Å}}^2$, corresponding to a <U> of 0.6 $\bar{\text{Å}}$. The cluster of waters in groove 2/3 have slightly greater mobility.

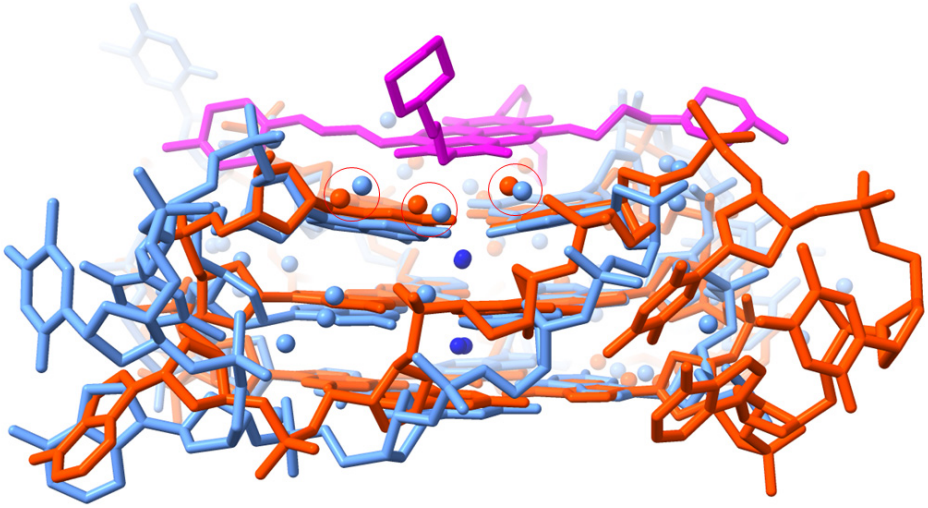
The extent to which water molecules located in the 3UYH crystal structure correspond to those found in the native structure was examined by superimposing on 3UYH the native crystal structure 1KF1 [126], overlaying the G-quartets (Figure 4a,b). Overlap of the quartets was good, as expected. However, the loops in the MM41-bound structure adopt distinct conformations compared to those in the native structure. Systematics of loop conformations in GQ crystal structures have been previously reported [95] and will not be further discussed here. Detailed comparison of water positions revealed that the cluster of four waters at the mouth of groove 1 that mediate between ono the morpholino end groups of MM41 and G-quadruplex is also present in the native structure (Figure 4a), with distances between each pair of waters (i.e., a 3UYH water . . . .1KF1 water) 0.6–1.0 $\bar{\text{Å}}$. Since these waters are mobile with a <U> of 0.6 $\bar{\text{Å}}$, they can be considered to likely occupy the same space. Three conserved water molecules are also present at the mouth of groove 3 (Figure 4b), close to the second morpholino end group.

2.3. Water Mediation in Acridine-G-Quadruplex Structures

The crystal structure (PDB id 3CE5) of the complex between the experimental drug BRACO19 and a bimolecular human GQ [74], shows that, in common with the MM41 complex, the GQ has adopted a parallel topology. In both instances, the ligand is stacked onto one end of the GQ, onto a terminal G-quartet. The BRACO19 molecule has three cationic charges at physiological pH, one in each side chain pyrrolidino ring and one on the central ring nitrogen atom in the acridine ring. None of these are directly hydrogen-bonded to anionic phosphate groups. Instead (Table 2 and Figure 5a), they are hydrogen bonded to water molecules. The waters hydrogen bonded to the pyrrolidino cationic nitrogen atoms do eventually link indirectly to phosphate groups, via further water molecules. Water molecule W52 hydrogen bonded to the acridine central ring nitrogen, appears to play a crucial role, hydrogen bonding both to an O6 of a guanine from the adjacent stacked G-quartet, and to N3 of a thymine in-plane with the acridine. W52 also hydrogen bonds to W53, which in turn hydrogen bonds to the carbonyl oxygen atom of one of the amide side-chains on BRACO19. The other amide group is trans to this and its amide nitrogen atom hydrogen bonds to O4 of this thymine—this is the sole direct drug—GQ hydrogen bond, with the remaining five being water-mediated. An overlay of the MM41 and BRACO19 complexes (Figure 5c) shows that several of the key water molecules are conserved between the two structures, with distances between pairs of waters < 1 Å, using the argument outlined in the previous section.



(**a**)



(**b**)

Figure 4. Superposition of the G-quartets of native and MM41-complexed G-quadruplex crystal structures, 1KF1 and 3UYH, respectively, viewed into groove 1. The native structure is colored light red, the ligand-bound is cyan, and the MM41 molecule is shown magenta. Only the water molecules in the groove are shown, colored as in their G-quadruplex structures. Those water molecules in the two structures that are <1.0 Å to each other are enclosed in red circles. (**a**) Viewed into groove 1. (**b**) Viewed into groove 3.

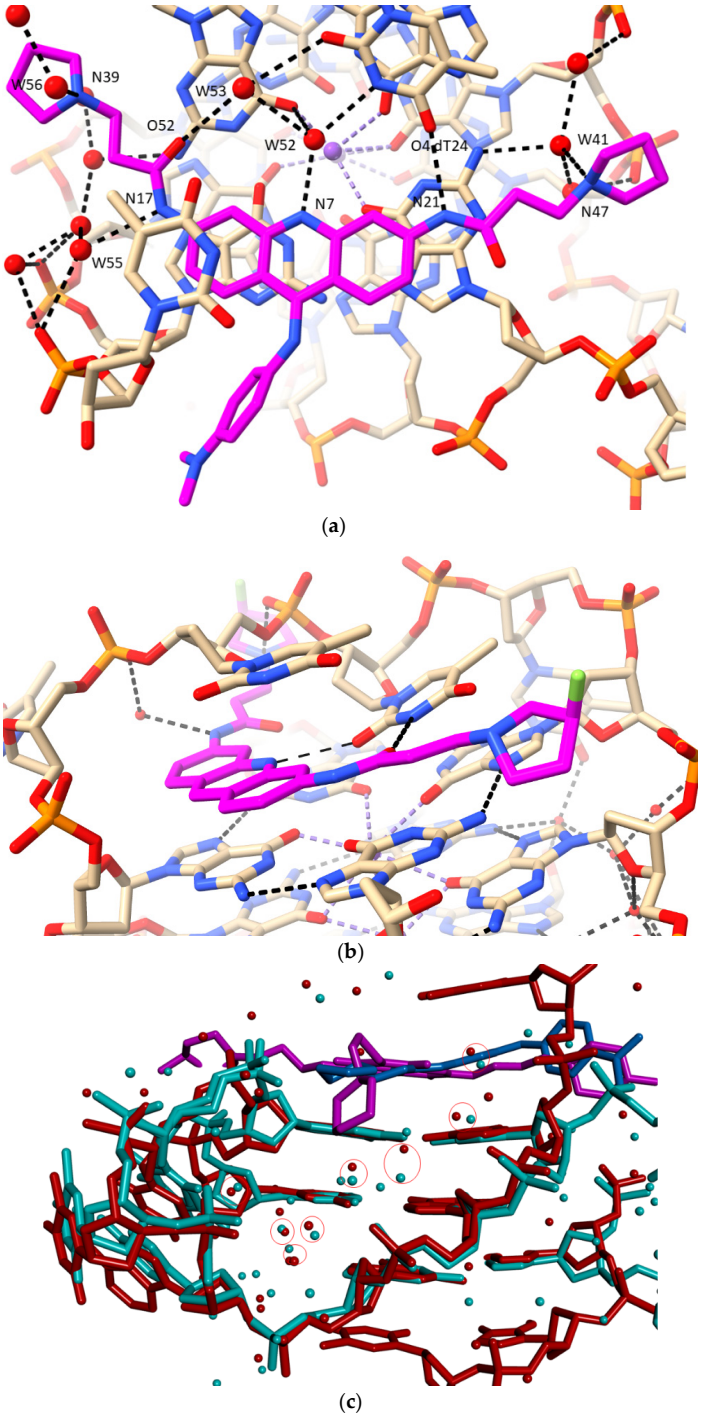


Figure 5. (**a**) View of the BRACO19 complex with a human telomeric bimolecular G-quadruplex, as observed in the crystal structure PDB id 3CE5 [74], projected onto the acridine plane. The carbon atoms of the ligand are colored magenta and water molecules are colored as red spheres. Hydrogen bonds are shown as dotted lines. Some water molecules not directly involved in ligand interactions have been omitted from this view to enhance clarity. (**b**) View of the complex involving a disubstituted acridine with a fluorine atom attached to each terminal side chain pyrrolidino ring, bound to an *Oxytricha nova* bimolecular G-quadruplex [100]. Color coding is as in other figures, with the fluorine atoms colored yellow. (**c**) Overlay of structures 3UYH (G-quadruplex in cyan and MM41 in magenta) and 3CE5 (G-quadruplex in red and BRACO19 in dark blue), superimposed on the G-quartets, viewed into groove 1 of the 3UYH structure. Water molecules in the groove are colored as in their G-quadruplex structures. Those water molecules in the two structures that are <1.0 Å distance to each other are enclosed in red circles.

Removal of the water molecules from this structure followed by energy minimization, resulted in movement of the acridine by ca 2.5 Å, enabling the acridine ring nitrogen atom to directly contact the thymine ring substituents. Such an arrangement has been observed in the series of co-crystal structures [75] involving disubstituted acridines with the bimolecular anti-parallel GQ from *Oxytricha nova* (Table 2). These structures, exemplified by the two high-resolution structures [100] with fluorine substituents in the pyrrolidino side chains, have direct O_2 thymine and N3 hydrogen-bond contacts with the acridine ring and an amide carbonyl oxygen atom (Figure 5b). A water molecule is involved in mediating between a side chain amide and a phosphate oxygen atom. However, the *Oxytricha nova* complexes all have a distinct anti-parallel GQ topology, with the acridine constrained within a tetranucleotide diagonal loop, with little room for any associated water molecules.

3. Discussion

The number of water molecules located in nucleic acid and protein crystal structures is invariably less than the total present in the crystal lattice. Thus, crystal structures 3UYH and 3CE5, whose reported water molecules are analyzed here, both have an estimated 56% solvent, which corresponds to >400 water molecules. A small fraction of this, 51 and 54 water molecules, respectively, were observed in electron density maps [74,81]. These are almost all first or second shell immobilized waters.

Three principal findings emerge from the current analysis:

1. The morpholino end groups of MM41, which are assumed to be basic in the buffering conditions of the crystallization experiment and in biological solution, do not directly contact the GQ. Hydrogen bonding/electrostatic interactions with negative backbone phosphate groups were anticipated but were not observed. Instead, the basic ring nitrogen in each morpholino group hydrogen bonds to one of a group of four water molecules positioned in the mouth of the relevant grooves (1 and 3). The waters are in hydrogen bond contact with backbone phosphates. Similarly, the basic pyrrolidino side chain terminal groups of BRACO19 do not directly contact phosphate groups in its GQ complex, with water mediation being observed in the crystal environment.
2. A nitrogen atom on both N-methyl-piperazine groups of MM41, by contrast, directly hydrogen bonds to a backbone phosphate oxygen atom, implying greater basicity than morpholino for this end group.
3. The water clusters associated with the two morpholino groups of MM41 are highly conserved between the native and the MM41-bound GQ structures. There is also conservation of a number of the ligand-associated waters between the MM41 and BRACO19 structures, and by implication, between the native and BRACO19 structures.

We suggest that the conserved water clusters have relevance to the observed structure-activity relationships for MM41 derivatives [81]. Thus, replacing the morpholino groups with isosteric groups such as hexose or ether groups, which lack the morpholino hydrogen bonding ability, results in an almost complete loss of GQ affinity and reduced biological activity compared to MM41. It is also notable that none of the four side chains are deeply embedded in the GQ grooves, and one might have therefore expected reduced GQ affinity compared to analogues with longer side chains. This is not the case since, as observed here, the short side chains are effectively captured by hydrogen bonded to the conserved water clusters, which would have to be displaced by longer side chains. This would be at a significant entropic cost. The strategy of replacing two of the four strongly basic end groups characteristic of earlier ND compounds (see, for example refs. [55,57,79]), by less strongly basic morpholino groups [18,19,81], has culminated in the design and evaluation of MM41, and subsequently compounds CM03 and SOP1812 (Figure 1), which are currently being assessed as pre-clinical candidates. The rationale for lowering the highly cationic nature of these ND compounds is that this could improve cellular uptake and tumor distribution while retaining GQ affinity. The present analysis has shown that this substitution has

preserved the water structure around the perimeter of the grooves in which the morpholino groups bind, with no diminution of GQ affinity [90].

The mediated waters in the BRACO19 structure also have relevance to the observed structure-activity relationships for BRACO-GQ interactions [73,74], and for BRACO19 biology. Although BRACO19 may not be further developed as an anti-cancer drug [127], its activity against a variety of anti-viral GQ targets [128–130] suggests the likelihood of further anti-viral analogue programs, for which the 3CE5 structure and its structured water features are of direct relevance.

The present analysis, although limited to two crystal structures, demonstrates that water molecules can play an active role in GQ-ligand recognition. This indicates that in silico and docking studies of ligand-GQ binding need to take account of reliably located explicit water molecules. It is concluded that their omission will lead to misleading conclusions on low-energy ligand binding states and interactions. Many such studies still tend to ignore the role of water and really require input from high-resolution crystal structures or reliable and well-validated water modelling/simulations. This has been recognized in several studies for example, refs. [118,131]. Prediction of water positions and mobilities in ligand complexes can be made using molecular dynamics [131], although this has only rarely been been used to date for GQ systems [132]. The prediction of water positions having low mobility in nucleic acids by use of a specially generated water force field together with statistical scoring has led to the development of an automated method, termed "SPLASH'EM" (Solvation Potential Laid around Statistical Hydration on Entire Macromolecules) [118], which has given results for duplex DNA and some RNA structures in good agreement with experiment. It will be interesting to see this type of approach used for those GQ ligand complexes for which there is high resolution structural data, as well as therapeutically important GQ drug targets such as that from the KRAS promoter sequence [114]. Conserved groove water molecules have been identified in the grooves of this crystal structure [114], as well as in other high-resolution GQ native structures [115]. By analogy with structures 3UYH and 3CE5, these conserved and structured waters should be retained in docking studies. The present analysis indicates that such water platforms for ligand binding can form an essential part of the total low-energy GQ interaction complex.

4. Materials and Methods

Crystal structures were downloaded from the Protein Data Bank and visualized by the ChimeraX (https://www.cgl.ucsf.edu/chimerax/, last accessed on 16 December 2021) [125] and BIOVIA Discovery Studio (https://www.3ds.com/products-services/biovia/, last accessed on 16 December 2021) programs.

Criteria for further consideration of individual structures were:

1. Resolution $\leq$ 2.5 Å,
2. Having at least one water molecule contacting a ligand,
3. Hydrogen bonds were accepted in a structure if

donor-acceptor distances $\leq$3.25 Å
donor-hydrogen . . . acceptor angles were $\leq$30° from ideality, and

4. Relevance to current drug discovery.

We excluded the daunomycin complexes with d(G_4) and d(TG_4T) (PDB ids 3TVB and 1O0K) from consideration, even though they are of high resolution and have large numbers of localized water molecules [116,117]. Their relationship to human GQ ligand complexes is unclear, because of their characteristics of multiple bound and stacked daunomycin molecules.

Root-mean-square displacement (RMSD) values, <U> (=$<U^2>)^{1/2}$, in Å, were calculated from the deposited experimental crystallographic isotropic temperature factors (B factors in $Å^2$), using the relationship

$$<U> = (B/8\pi^2)^{1/2}$$

B factors, obtained from the refinement of a crystal structure, indicate the relative isotropic thermal motions of individual atoms in a crystal structure.

Molecular mechanics and docking studies were performed on structure PDB id 3UYH using the ARGUSLAB (http://www.arguslab.com/arguslab.com/Publications.html. Last accessed on 23 August 2021) and AVOGADRO (https://avogadro.cc/, last accessed on 23 August 2021) packages. The UFF Universal Force Field [133] was used in these calculations. Docking studies were also undertaken within the G-quadruplex ligand database G4LDB 2.2 (https://www.g4ldb.com. Last accessed on 16 December 2021) [134], which utilizes the docking modules in AutoDock Vina 1.1.2 [135].

5. Conclusions

This study has analysed data from earlier crystal structure analyses and has shown that 1st and 2nd shell water molecules play an important role in the binding of two experimental small-molecule drugs, MM41 and BRACO19 to human telomeric G-quadruplexes. These waters mediate between cationic side-chain functional groups and phosphate backbones. They also directly bridge the chromophore core of the drugs and other G-quadruplex groups. Altogether, waters serve to maintain the drug molecules in their low-energy binding positions and their removal would result in incorrect drug positions. This has implications for drug design and virtual library screening and docking.

Funding: This research received no external funding.

Institutional Review Board Statement: Not applicable.

Informed Consent Statement: Not applicable.

Data Availability Statement: Data is contained within the article.

Acknowledgments: I am grateful to the former members of my group, Gavin Collie and Nancy Campbell, who determined and refined numerous G-quadruplex-small molecule crystal structures, resulting in high-quality and well-validated structures. Work in my laboratory supporting the development of ND ligands has been funded by grants from Cancer Research UK, the European Union, an MRC Confidence in Concept Award, and the UCL Technology Fund. I am grateful to all these funders for their generous support.

Conflicts of Interest: The author declares no conflict of interest.

References

1. Gellert, M.; Lipsett, M.N.; Davies, D.R. Helix formation by guanylic acid. *Proc. Natl. Acad. Sci. USA* **1962**, *48*, 2013–2018. [CrossRef]
2. Burge, S.; Parkinson, G.N.; Hazel, P.; Todd, A.K.; Neidle, S. Quadruplex DNA: Sequence, topology and structure. *Nucleic Acids Res.* **2006**, *34*, 5402–5415. [CrossRef]
3. Bochman, M.L.; Paeschke, K.; Zakian, V.A. DNA secondary structures: Stability and function of G-quadruplex structures. *Nature Rev. Genet.* **2012**, *13*, 770–780. [CrossRef] [PubMed]
4. Spiegel, J.; Adhikari, S.; Balasubramanian, S. The structure and function of DNA G-quadruplexes. *Trends Chem.* **2019**, *2*, 123–136. [CrossRef] [PubMed]
5. Winnerdy, F.R.; Phan, A.T. Quadruplex structure and diversity. *Ann. Rep. Med. Chem.* **2020**, *54*, 45–73.
6. Todd, A.K.; Johnston, M.; Neidle, S. Highly prevalent putative quadruplex sequence motifs in human DNA. *Nucleic Acids Res.* **2005**, *33*, 2901–2907. [CrossRef] [PubMed]
7. Huppert, J.L.; Balasubramanian, S. Prevalence of quadruplexes in the human genome. *Nucleic Acids Res.* **2005**, *33*, 2908–2916. [CrossRef] [PubMed]
8. Henderson, E.; Hardin, C.C.; Walk, S.K.; Tinoco, I., Jr.; Blackburn, E.H. Telomeric DNA oligonucleotides form novel intramolecular structures containing guanine-guanine base pairs. *Biochemistry* **1987**, *51*, 899–908.
9. Williamson, J.R. G-quartet structures in telomeric DNA. *Ann. Rev. Biophys.* **1994**, *23*, 703–730. [CrossRef] [PubMed]
10. Huppert, J.L.; Balasubramanian, S. G-quadruplexes in promoters throughout the human genome. *Nucleic Acids Res.* **2007**, *35*, 406–413. [CrossRef]
11. Siddiqui-Jain, A.; Grand, C.L.; Bearss, D.J.; Hurley, L.H. Direct evidence for a G-quadruplex in a promoter region and its targeting with a small molecule to repress c-*MYC* transcription. *Proc. Natl. Acad. Sci. USA* **2002**, *99*, 11593–11598. [CrossRef]
12. Varshney, D.; Spiegel, J.; Zyner, K.; Tannahill, D.; Balasubramanian, S. The regulation and functions of DNA and RNA G-quadruplexes. *Nature Rev. Mol. Cell Biol.* **2020**, *21*, 459–474. [CrossRef]

13. Spiegel, J.; Cuesta, S.M.; Adhikari, S.; Hänsel-Hertsch, R.; Tannahill, D.; Balasubramanian, S. G-quadruplexes are transcription factor binding hubs in human chromatin. *Genome Biol.* **2021**, *22*, 117. [CrossRef] [PubMed]
14. Hansel-Hertsch, R.; Beraldi, D.; Lensing, S.V.; Marsico, G.; Zyner, K.; Parry, A.; Di Antonio, M.; Pike, J.; Kimura, H.; Narita, M.; et al. G-quadruplex structures mark human regulatory chromatin. *Nature Genet.* **2016**, *48*, 1267–1272. [CrossRef]
15. Huppert, J.L.; Bugaut, A.; Kumari, S.; Balasubramanian, S. G-quadruplexes: The beginning and end of UTRs. *Nucleic Acids Res.* **2008**, *36*, 6260–6268. [CrossRef] [PubMed]
16. Bugaut, A.; Balasubramanian, S. 5′-UTR RNA G-quadruplexes: Translation regulation and targeting. *Nucleic Acids Res.* **2012**, *40*, 4727–4741. [CrossRef] [PubMed]
17. Lee, D.S.M.; Ghanem, L.R.; Barash, Y. Integrative analysis reveals RNA G-quadruplexes in UTRs are selectively constrained and enriched for functional associations. *Nat. Commun.* **2020**, *11*, 527. [CrossRef]
18. Marchetti, C.; Zyner, K.G.; Ohnmacht, S.A.; Robson, M.; Haider, S.M.; Morton, J.P.; Marsico, G.; Vo, T.; Laughlin-Toth, S.; Ahmed, A.A.; et al. Targeting multiple effector pathways in pancreatic ductal adenocarcinoma with a G-quadruplex-binding small molecule. *J. Med. Chem.* **2018**, *61*, 2500–2517. [CrossRef] [PubMed]
19. Ahmed, A.A.; Angell, R.; Oxenford, S.; Worthington, J.; Williams, N.; Barton, N.; Fowler, T.G.; O'Flynn, D.E.; Sunose, M.; McConville, M.; et al. Asymmetrically substituted quadruplex-binding naphthalene diimide showing potent activity in pancreatic cancer models. *ACS Med. Chem. Lett.* **2020**, *11*, 1634–1644. [CrossRef]
20. Zyner, K.G.; Mulhearn, D.S.; Adhikari, S.; Martínez Cuesta, S.; Di Antonio, M.; Erard, N.; Hannon, G.J.; Tannahill, D.; Balasubramanian, S. Genetic interactions of G-quadruplexes in humans. *eLife* **2019**, *8*, e46793. [CrossRef] [PubMed]
21. Hänsel-Hertsch, R.; Simeone, A.; Shea, A.; Hui, W.W.I.; Zyner, K.G.; Marsico, G.; Rueda, O.M.; Bruna, A.; Martin, A.; Zhang, X.; et al. Landscape of G-quadruplex DNA structural regions in breast cancer. *Nature Genet.* **2020**, *52*, 878–883. [CrossRef]
22. Shen, J.; Varshney, D.; Simeone, A.; Zhang, X.; Adhikari, S.; Tannahill, D.; Balasubramanian, S. Promoter G-quadruplex folding precedes transcription and is controlled by chromatin. *Genome Biol.* **2021**, *22*, 143. [CrossRef]
23. Perrone, R.; Nadai, M.; Poe, J.A.; Frasson, I.; Palumbo, M.; Palù, G.; Smithgall, T.E.; Richter, S.N. Formation of a unique cluster of G-quadruplex structures in the HIV-1 Nef coding region: Implications for antiviral activity. *PLoS ONE* **2013**, *8*, e73121.
24. Ruggiero, E.; Richter, S.N. Viral G-quadruplexes: New frontiers in virus pathogenesis and antiviral therapy. *Ann. Rep. Med. Chem.* **2020**, *54*, 101–131.
25. Abiri, A.; Lavigne, M.; Rezaei, M.; Nikzad, S.; Zare, P.; Mergny, J.-L.; Rahimi, H.R. Unlocking G-quadruplexes as antiviral targets. *Pharmacol. Rev.* **2021**, *73*, 897–923. [CrossRef] [PubMed]
26. Zhao, C.; Qin, G.; Niu, J.; Wang, Z.; Wang, C.; Ren, J.; Qu, X. Targeting RNA G-quadruplex in SARS-CoV-2: A promising therapeutic target for COVID-19? *Angew. Chem. Int. Ed. Engl.* **2021**, *60*, 432–438. [CrossRef] [PubMed]
27. Belmonte-Reche, E.; Serrano-Chacón, I.; Gonzalez, C.; Gallo, J.; Bañobre-López, M. Potential G-quadruplexes and i-Motifs in the SARS-CoV-2. *PLoS ONE* **2021**, *16*, e0250654. [CrossRef] [PubMed]
28. Rawal, P.; Kummarasetti, V.B.; Ravindran, J.; Kumar, N.; Halder, K.; Sharma, R.; Mukerji, M.; Das, S.K.; Chowdhury, S. Genome-wide prediction of G4 DNA as regulatory motifs: Role in *Escherichia coli* global regulation. *Genome Res.* **2006**, *16*, 644–655. [CrossRef]
29. Holder, I.T.; Hartig, J.S. A matter of location: Influence of G-quadruplexes on *Escherichia coli* gene expression. *Chem. Biol.* **2014**, *21*, 1511–1521. [CrossRef]
30. Du, X.; Wojtowicz, D.; Bowers, A.A.; Levens, D.; Benham, C.J.; Przytycka, T.M. The genome-wide distribution of non-B DNA motifs is shaped by operon structure and suggests the transcriptional importance of non-B DNA structures in *Escherichia coli*. *Nucleic Acids Res.* **2014**, *41*, 5965–5977. [CrossRef] [PubMed]
31. Yadav, P.; Kim, N.; Kumari, M.; Verma, S.; Sharma, T.K.; Yadav, V.; Kumar, A. G-quadruplex structures in bacteria—Biological relevance and potential as antimicrobial target. *J. Bacteriol.* **2021**, *203*, e0057720. [CrossRef] [PubMed]
32. Smargiasso, N.; Gabelica, V.; Damblon, C.; Rosu, F.; De Pauw, E.; Teulade-Fichou, M.-P.; Rowe, J.A.; Claessens, A. Putative DNA G-quadruplex formation within the promoters of *Plasmodium falciparum var* genes. *BMC Genom.* **2009**, *10*, 362. [CrossRef] [PubMed]
33. Harris, L.M.; Monsell, K.R.; Noulin, F.; Famodimu, M.T.; Smargiasso, N.; Damblon, C.; Horrocks, P.; Merrick, C.J. G-quadruplex DNA motifs in the malaria parasite plasmodium falciparum and their potential as novel antimalarial drug targets. *Antimicrob. Agents Chemother.* **2018**, *62*, e01828-17. [CrossRef] [PubMed]
34. Gazanion, E.; Lacroix, L.; Alberti, P.; Gurung, P.; Wein, S.; Cheng, M.; Mergny, J.-L.; Gomes, A.R.; Lopez-Rubio, J.J. Genome wide distribution of G-quadruplexes and their impact on gene expression in malaria parasites. *PLoS Genet.* **2020**, *16*, e1008917. [CrossRef] [PubMed]
35. Biffi, G.; Tannahill, D.; McCafferty, J.; Balasubramanian, S. Quantitative visualization of DNA G-quadruplex structures in human cells. *Nat. Chem.* **2013**, *5*, 182–186. [CrossRef]
36. Henderson, A.; Wu, Y.; Huang, Y.C.; Chavez, E.A.; Platt, J.; Johnson, F.B.; Brosh, R.M., Jr.; Sen, D.; Lansdorp, P.M. Detection of G-quadruplex DNA in mammalian cells. *Nucleic Acids Res.* **2013**, *42*, 860–869. [CrossRef]
37. Biffi, G.; Di Antonio, M.; Tannahill, D.; Balasubramanian, S. Visualization and selective chemical targeting of RNA G-quadruplex structures in the cytoplasm of human cells. *Nat. Chem.* **2014**, *6*, 75–80. [CrossRef]
38. Tseng, T.-Y.; Chien, C.-H.; Chu, J.-F.; Huang, W.-C.; Lin, M.-Y.; Chang, C.-C.; Chang, T.-C. Fluorescent probe for visualizing guanine-quadruplex DNA by fluorescence lifetime imaging microscopy. *J. Biomed. Opt.* **2013**, *18*, 101309. [CrossRef]

39. Laguerre, A.; Hukezalie, K.; Winckler, P.; Katranji, F.; Chanteloup, G.; Pirrotta, M.; Perrier-Cornet, J.-M.; Wong, J.M.Y.; Monchaud, D. Visualization of RNA-quadruplexes in live cells. *J. Am. Chem. Soc.* **2015**, *137*, 8521–8525. [CrossRef]
40. Summers, P.A.; Lewis, B.W.; Gonzalez-Garcia, J.; Porreca, R.M.; Lim, A.H.M.; Cadinu, P.; Martin-Pintado, N.; Mann, D.J.; Edel, J.B.; Vannier, J.B.; et al. Visualising G-quadruplex DNA dynamics in live cells by fluorescence lifetime imaging microscopy. *Nat. Commun.* **2021**, *12*, 162. [CrossRef]
41. 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. Single-molecule visualization of DNA G-quadruplex formation in live cells. *Nat. Chem.* **2020**, *12*, 832–837. [CrossRef]
42. Long, W.; Zheng, B.-X.; Huang, X.-H.; She, M.-T.; Liu, A.-L.; Zhang, K.; Wong, W.-L.; Lu, Y.-J. Molecular recognition and imaging of human telomeric G-quadruplex DNA in live cells: A systematic advancement of thiazole orange scaffold to enhance binding specificity and inhibition of gene expression. *J. Med. Chem.* **2021**, *64*, 2125–2138. [CrossRef] [PubMed]
43. Wang, Y.; Yang, J.; Wild, A.T.; Wu, W.H.; Shah, R.; Danussi, C.; Riggins, G.J.; Kannan, K.; Sulman, E.P.; Chan, T.A.; et al. G-quadruplex DNA drives genomic instability and represents a targetable molecular abnormality in ATRX-deficient malignant glioma. *Nat. Commun.* **2019**, *10*, 943. [CrossRef]
44. Lerner, L.K.; Sale, J.E. Replication of G quadruplex DNA. *Genes* **2019**, *10*, 95. [CrossRef] [PubMed]
45. Amato, J.; Miglietta, G.; Morigi, R.; Iaccarino, N.; Locatelli, A.; Leoni, A.; Novellino, E.; Pagano, B.; Capranico, G.; Randazzo, A. Monohydrazone based G-quadruplex selective ligands induce DNA damage and genome instability in human cancer cells. *J. Med. Chem.* **2020**, *63*, 3090–3103. [CrossRef] [PubMed]
46. Balasubramanian, S.; Hurley, L.H.; Neidle, S. Targeting G-quadruplexes in gene promoters: A novel anticancer strategy? *Nat. Rev. Drug. Discov.* **2011**, *10*, 261–275. [CrossRef] [PubMed]
47. Rigo, R.; Palumbo, M.; Sissi, C. G-quadruplexes in human promoters: A challenge for therapeutic applications. *Biochim. Biophys. Acta* **2017**, *1861*, 1399–1413. [CrossRef]
48. Song, J.H.; Kang, H.-J.; Luevano, L.A.; Gokhale, V.; Wu, K.; Pandey, R.; Sherry Chow, H.-H.; Hurley, L.H.; Kraft, A.S. Small-molecule-targeting hairpin loop of hTERT promoter G-quadruplex induces cancer cell death. *Cell Chem. Biol.* **2019**, *26*, 1110–1121. [CrossRef]
49. Boddupally, P.V.; Hahn, S.; Beman, C.; De, B.; Brooks, T.A.; Gokhale, V.; Hurley, L.H. Anticancer activity and cellular repression of c-MYC by the G-quadruplex-stabilizing 11-piperazinylquindoline is not dependent on direct targeting of the G-quadruplex in the c-MYC promoter. *J. Med. Chem.* **2012**, *55*, 6076–6086. [CrossRef] [PubMed]
50. Calabrese, D.R.; Chen, X.; Leon, E.C.; Gaikwad, S.M.; Phyo, Z.; Hewitt, W.M.; Alden, S.; Hilimire, T.A.; He, F.; Michalowski, A.M.; et al. Chemical and structural studies provide a mechanistic basis for recognition of the MYC G-quadruplex. *Nat. Commun.* **2018**, *9*, 4229. [CrossRef] [PubMed]
51. Wang, K.B.; Elsayed, M.S.A.; Wu, G.; Deng, N.; Cushman, M.; Yang, D. Indenoisoquinoline topoisomerase inhibitors strongly bind and stabilize the MYC promoter G-quadruplex and downregulate MYC. *J. Am. Chem. Soc.* **2019**, *141*, 11059–11070. [CrossRef] [PubMed]
52. Lavrado, J.; Brito, H.; Borralho, P.M.; Ohnmacht, S.A.; Kim, N.S.; Leitão, C.; Pisco, S.; Gunaratnam, M.; Rodrigues, C.M.; Moreira, R.; et al. KRAS oncogene repression in colon cancer cell lines by G-quadruplex binding indolo[3,2-c]quinolines. *Sci. Rep.* **2015**, *5*, 9696. [CrossRef]
53. Brito, H.; Martins, A.C.; Lavrado, J.; Mendes, E.; Francisco, A.P.; Santos, S.A.; Ohnmacht, S.A.; Kim, N.S.; Rodrigues, C.M.; Moreira, R.; et al. KRAS oncogene in colon cancer cells with 7-carboxylate indolo[3,2-b]quinoline tri-alkylamine derivatives. *PLoS ONE* **2015**, *10*, e0126891. [CrossRef]
54. Wang, X.-D.; Ou, T.-M.; Lu, Y.-J.; Li, Z.; Xu, Z.; Xi, C.; Tan, J.-H.; Huang, S.-L.; An, L.-K.; Li, D.; et al. Turning off transcription of the bcl-2 gene by stabilizing the bcl-2 promoter quadruplex with quindoline derivatives. *J. Med. Chem.* **2010**, *53*, 4390–4398. [CrossRef] [PubMed]
55. Gunaratnam, M.; Collie, G.W.; Reszka, A.P.; Todd, A.K.; Parkinson, G.N.; Neidle, S. A naphthalene diimide G-quadruplex ligand inhibits cell growth and down-regulates BCL-2 expression in an imatinib-resistant gastrointestinal cancer cell line. *Bioorg. Med. Chem.* **2018**, *26*, 2958–2964. [CrossRef] [PubMed]
56. Pelliccia, S.; Amato, J.; Capasso, D.; Di Gaetano, S.; Massarotti, A.; Piccolo, M.; Irace, C.; Tron, G.C.; Pagano, B.; Randazzo, A.; et al. Bio-inspired dual-selective BCL-2/c-MYC G-quadruplex binders: Design, synthesis, and anticancer activity of drug-like imidazo[2,1-i]purine derivatives. *J. Med. Chem.* **2020**, *63*, 2035–2050. [CrossRef] [PubMed]
57. Gunaratnam, M.; Swank, S.; Haider, S.M.; Galesa, K.; Reszka, A.P.; Beltran, M.; Cuenca, F.; Fletcher, J.A.; Neidle, S. Targeting human gastrointestinal stromal tumor cells with a quadruplex-binding small molecule. *J. Med. Chem.* **2009**, *52*, 3774–3783. [CrossRef] [PubMed]
58. McLuckie, K.I.; Waller, Z.A.; Sanders, D.A.; Alves, D.; Rodriguez, R.; Dash, J.; McKenzie, G.J.; Venkitaraman, A.R.; Balasubramanian, S. G-quadruplex-binding benzo[a]phenoxazines down-regulate c-KIT expression in human gastric carcinoma cells. *J. Am. Chem. Soc.* **2011**, *133*, 2658–2663. [CrossRef] [PubMed]
59. Kosiol, N.; Juranek, S.; Brossart, P.; Heine, A.; Paeschke, K. G-quadruplexes: A promising target for cancer therapy. *Mol. Cancer* **2021**, *20*, 40. [CrossRef] [PubMed]
60. Monchaud, D.; Teulade-Fichou, M.-P. A hitchhiker's guide to G-quadruplex ligands. *Org. Biomol. Chem.* **2008**, *6*, 627–636. [CrossRef]

61. Müller, S.; Rodriguez, R. G-quadruplex interacting small molecules and drugs: From bench towards bedside. *Expert. Rev. Clin. Pharmacol.* **2014**, *7*, 663–679. [CrossRef]
62. Islam, M.K.; Jackson, P.J.; Rahman, K.M.; Thurston, D.E. Recent advances in targeting the telomeric G-quadruplex DNA sequence with small molecules as a strategy for anticancer therapies. *Future Med. Chem.* **2016**, *8*, 1259–1290. [CrossRef]
63. Neidle, S. Quadruplex nucleic acids as novel therapeutic targets. *J. Med. Chem.* **2016**, *59*, 5987–6011. [CrossRef]
64. Duarte, A.R.; Cadoni, E.; Ressurreição, A.S.; Moreira, R.; Paulo, A. Design of modular G-quadruplex ligands. *ChemMedChem* **2018**, *13*, 869–893. [CrossRef]
65. Asamitsu, S.; Bando, T.; Sugiyama, H. Ligand design to acquire specificity to intended G-quadruplex structures. *Chemistry* **2019**, *25*, 417–430. [CrossRef] [PubMed]
66. Savva, L.; Georgiades, S.N. Recent developments in small-molecule ligands of medicinal relevance for harnessing the anticancer potential of G-quadruplexes. *Molecules* **2021**, *26*, 841. [CrossRef]
67. Harrison, R.J.; Gowan, S.M.; Kelland, L.R.; Neidle, S. Human telomerase inhibition by substituted acridine derivatives. *Bioorg. Med. Chem. Lett.* **1999**, *9*, 2463–2468. [CrossRef]
68. Read, M.A.; Wood, A.A.; Harrison, J.R.; Gowan, S.M.; Kelland, L.R.; Dosanjh, H.S.; Neidle, S. Molecular modelling studies on G-quadruplex complexes of telomerase inhibitors: Structure-activity relationships. *J. Med. Chem.* **1999**, *42*, 4538–4546. [CrossRef]
69. Read, M.A.; Harrison, R.J.; Romagnoli, B.; Tanious, F.A.; Gowan, S.H.; Reszka, A.P.; Wilson, W.D.; Kelland, L.R.; Neidle, S. Structure-based design of selective and potent G quadruplex-mediated telomerase inhibitors. *Proc. Natl. Acad. Sci. USA* **2001**, *98*, 4844–4849. [CrossRef] [PubMed]
70. Heald, R.A.; Modi, C.; Cookson, J.C.; Hutchinson, I.; Laughton, C.A.; Gowan, S.M.; Kelland, L.R.; Stevens, M.F.G. Antitumor polycyclic acridines. 8. Synthesis and telomerase-inhibitory activity of methylated pentacyclic acridinium salts. *J. Med. Chem.* **2002**, *45*, 590–597. [CrossRef] [PubMed]
71. Haider, S.M.; Parkinson, G.N.; Neidle, S. Structure of a G-quadruplex–ligand complex. *J. Mol. Biol.* **2003**, *326*, 117–125. [CrossRef]
72. Burger, A.M.; Dai, F.; Schultes, C.M.; Reszka, A.P.; Moore, M.J.; Double, J.A.; Neidle, S. The G-quadruplex-interactive molecule BRACO-19 inhibits tumor growth, consistent with telomere targeting and interference with telomerase function. *Cancer Res.* **2005**, *65*, 1489–1496. [CrossRef] [PubMed]
73. Moore, M.J.; Schultes, C.M.; Cuesta, J.; Cuenca, F.; Gunaratnam, M.; Tanious, F.A.; Wilson, W.D.; Neidle, S. Trisubstituted acridines as G-quadruplex telomere targeting agents. Effects of extensions of the 3,6- and 9-side chains on quadruplex binding, telomerase activity, and cell proliferation. *J. Med. Chem.* **2006**, *49*, 582–599. [CrossRef] [PubMed]
74. Campbell, N.H.; Parkinson, G.N.; Reszka, A.P.; Neidle, S. Structural basis of DNA quadruplex recognition by an acridine drug. *J. Am. Chem. Soc.* **2008**, *130*, 6722–6724. [CrossRef] [PubMed]
75. Campbell, N.H.; Patel, M.; Tofa, A.B.; Ghosh, R.; Parkinson, G.N.; Neidle, S. Selectivity in ligand recognition of G-quadruplex loops. *Biochemistry* **2009**, *48*, 1675–1680. [CrossRef]
76. Collie, G.W.; Sparapani, S.; Parkinson, G.N.; Neidle, S. Structural basis of telomeric RNA quadruplex-acridine ligand recognition. *J. Am. Chem. Soc.* **2011**, *133*, 2721–2728. [CrossRef]
77. Cuenca, F.; Greciano, O.; Gunaratnam, M.; Haider, S.; Munnur, D.; Nanjunda, R.; Wilson, W.D.; Neidle, S. Tri- and tetra-substituted naphthalene diimides as potent G-quadruplex ligands. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 1668–1673. [CrossRef]
78. Hampel, S.M.; Sidibe, A.; Gunaratnam, M.; Riou, J.-F.; Neidle, S. Tetrasubstituted naphthalene diimide ligands with selectivity for telomeric G-quadruplexes and cancer cells. *Bioorg. Med. Chem. Lett.* **2010**, *20*, 6459–6463. [CrossRef]
79. Gunaratnam, M.; de la Fuente, M.; Hampel, S.M.; Todd, A.K.; Reszka, A.P.; Schatzlein, A.; Neidle, S. Targeting pancreatic cancer with a G-quadruplex ligand. *Bioorg. Med. Chem.* **2011**, *19*, 7151–7157. [CrossRef] [PubMed]
80. Collie, G.W.; Promontorio, R.; Hampel, S.M.; Micco, M.; Neidle, S.; Parkinson, G.N. Structural basis for telomeric G-quadruplex naphthalene diimide ligand targeting. *J. Am. Chem. Soc.* **2012**, *134*, 2723–2731. [CrossRef]
81. Micco, M.; Collie, G.W.; Dale, A.G.; Ohnmacht, S.A.; Pazitna, I.; Gunaratnam, M.; Reszka, A.P.; Neidle, S. Structure-based design and evaluation of naphthalene diimide G-quadruplex ligands as telomere targeting agents in pancreatic cancer cells. *J. Med. Chem.* **2013**, *56*, 2959–2974. [CrossRef]
82. Nadai, M.; Cimino-Reale, G.; Sattin, G.; Doria, F.; Butovskaya, E.; Zaffaroni, N.; Freccero, M.; Palumbo, M.; Richter, S.N.; Folini, M. Assessment of gene promoter Gquadruplex binding and modulation by a naphthalene diimide derivative in tumor cells. *Int. J. Oncol.* **2015**, *46*, 369–380. [CrossRef]
83. Ohnmacht, S.A.; Marchetti, C.; Gunaratnam, M.; Besser, R.J.; Haider, S.M.; Di Vita, G.; Lowe, H.L.; Mellinas-Gomez, M.; Diocou, S.; Robson, M.; et al. A G-quadruplex-binding compound showing anti-tumour activity in an *in vivo* model for pancreatic cancer. *Sci. Rep.* **2015**, *5*, 11385. [CrossRef]
84. Spinello, A.; Barone, G.; Grunenberg, J. Molecular recognition of naphthalene diimide ligands by telomeric quadruplex-DNA: The importance of the protonation state and mediated hydrogen bonds. *Phys. Chem. Chem. Phys.* **2016**, *18*, 2871–2877. [CrossRef]
85. Lopergolo, A.; Perrone, R.; Tortoreto, M.; Doria, F.; Beretta, G.L.; Zuco, V.; Freccero, M.; Borrello, M.G.; Lanzi, C.; Richter, S.N.; et al. Targeting of RET oncogene by naphthalene diimide-mediated gene promoter G-quadruplex stabilization exerts anti-tumor activity in oncogene-addicted human medullary thyroid cancer. *Oncotarget* **2016**, *7*, 49649–49663. [CrossRef]
86. Răsădean, D.M.; Sheng, B.; Dash, J.; Pantoş, G.D. Amino-acid-derived naphthalenediimides as versatile G-quadruplex binders. *Chemistry* **2017**, *23*, 8491–8499. [CrossRef] [PubMed]

87. Street, S.; Chin, D.; Hollingworth, G.; Berry, M.; Morales, J.C.; Galan, M.C. Divalent naphthalene diimide ligands display high selectivity for the human telomeric G-quadruplex in K^+ buffer. *Chemistry* **2017**, *23*, 6953–6958. [CrossRef] [PubMed]
88. Recagni, M.; Greco, M.L.; Milelli, A.; Minarini, A.; Zaffaroni, N.; Folini, M.; Sissi, C. Distinct biological responses of metastatic castration resistant prostate cancer cells upon exposure to G-quadruplex interacting naphthalenediimide derivatives. *Eur. J. Med. Chem.* **2019**, *177*, 401–413. [CrossRef]
89. Pirota, V.; Nadai, M.; Doria, F.; Richter, S.N. Naphthalene diimides as multimodal G-quadruplex-selective ligands. *Molecules* **2019**, *24*, 426. [CrossRef] [PubMed]
90. Vo, T.; Oxenford, S.; Angell, R.; Marchetti, C.; Ohnmacht, S.A.; Wilson, W.D.; Neidle, S. Substituted naphthalenediimide compounds bind selectively to two human quadruplex structures with parallel topology. *ACS Med. Chem. Lett.* **2020**, *11*, 991–999. [CrossRef] [PubMed]
91. Platella, C.; Trajkovski, M.; Doria, F.; Freccero, M.; Plavec, J.; Montesarchio, D. On the interaction of an anticancer trisubstituted naphthalene diimide with G-quadruplexes of different topologies: A structural insight. *Nucleic Acids Res.* **2020**, *48*, 12380–12393. [CrossRef] [PubMed]
92. Hao, X.; Wang, C.; Wang, Y.; Li, C.; Hou, J.; Zhang, F.; Kang, C.; Gao, L. Topological conversion of human telomeric G-quadruplexes from hybrid to parallel form induced by naphthalene diimide ligands. *Int. J. Biol. Macromol.* **2021**, *167*, 1048–1058. [CrossRef]
93. Sanchez-Martin, V.; Schneider, D.A.; Ortiz-Gonzalez, M.; Soriano-Lerma, A.; Linde-Rodriguez, A.; Perez-Carrasco, V.; Gutierrez-Fernandez, J.; Cuadros, M.; González, C.; Soriano, M.; et al. Targeting ribosomal G-quadruplexes with naphthalene-diimides as RNA polymerase I inhibitors for colorectal cancer treatment. *Cell Chem. Biol.* **2021**, *28*, 1590–1601. [CrossRef]
94. Platella, C.; Napolitano, E.; Riccardi, C.; Musumeci, D.; Montesarchio, D. Disentangling the structure−activity relationships of naphthalenediimides as anticancer G-quadruplex-targeting drugs. *J. Med. Chem.* **2021**, *64*, 3578–3603. [CrossRef]
95. Collie, G.W.; Campbell, N.H.; Neidle, S. Loop flexibility in human telomeric quadruplex small-molecule complexes. *Nucleic Acids Res.* **2015**, *43*, 4785–4799. [CrossRef]
96. Chung, W.J.; Heddi, B.; Hamon, F.; Teulade-Fichou, M.-P.; Phan, A.T. Solution structure of a G-quadruplex bound to the bisquinolinium compound Phen-DC(3). *Angew. Chem. Int. Ed. Engl.* **2014**, *53*, 999–1002. [CrossRef]
97. Dai, J.; Carver, M.; Hurley, L.H.; Yang, D. Solution structure of a 2:1 quindoline-c-MYC G-quadruplex: Insights into G-quadruplex-interactive small molecule drug design. *J. Am. Chem. Soc.* **2011**, *133*, 17673–17680. [CrossRef] [PubMed]
98. Papi, F.; Bazzicalupi, C.; Ferraroni, M.; Ciolli, G.; Lombardi, P.; Khan, A.Y.; Kumar, G.S.; Gratteri, P. Pyridine derivative of the natural alkaloid berberine as human telomeric G4-DNA Binder: A solution and solid-state study. *ACS Med. Chem. Lett.* **2020**, *11*, 645–650. [CrossRef] [PubMed]
99. Dickerhoff, J.; Dai, J.; Yang, D. Structural recognition of the MYC promoter G-quadruplex by a quinoline derivative: Insights into molecular targeting of parallel G-quadruplexes. *Nucleic Acids Res.* **2021**, *49*, 5905–5915. [CrossRef]
100. Campbell, N.H.; Smith, D.L.; Reszka, A.P.; Neidle, S.; O'Hagan, D. Fluorine in medicinal chemistry: β-fluorination of peripheral pyrrolidines attached to acridine ligands affects their interactions with G-quadruplex DNA. *Org. Biomol. Chem.* **2011**, *9*, 1328–1331. [CrossRef]
101. Ma, D.L.; Ma, V.P.; Chan, D.S.; Leung, K.H.; Zhong, H.J.; Leung, C.H. In silico screening of quadruplex-binding ligands. *Methods* **2012**, *57*, 106–114. [CrossRef] [PubMed]
102. Cosconati, S.; Marinelli, L.; Trotta, R.; Virno, A.; Mayol, L.; Novellino, E.; Olson, A.J.; Randazzo, A. Tandem application of virtual screening and NMR experiments in the discovery of brand-new DNA quadruplex groove binders. *J. Am. Chem. Soc.* **2009**, *131*, 16336–16337. [CrossRef] [PubMed]
103. Cosconati, S.; Marinelli, L.; Trotta, R.; Virno, A.; De Tito, S.; Romagnoli, R.; Pagano, B.; Limongelli, V.; Giancola, C.; Baraldi, P.G.; et al. Structural and conformational requisites in DNA quadruplex groove binding: Another piece to the puzzle. *J. Am. Chem. Soc.* **2010**, *132*, 6425–6433. [CrossRef] [PubMed]
104. Alcaro, S.; Musetti, C.; Distinto, S.; Casatti, M.; Zagotto, G.; Artese, A.; Parrotta, L.; Moraca, F.; Costa, G.; Ortuso, F.; et al. Identification and characterization of new DNA G-quadruplex binders selected by a combination of ligand and structure-based virtual screening approaches. *J. Med. Chem.* **2013**, *56*, 843–855. [CrossRef]
105. Di Leva, F.S.; Zizza, P.; Cingolani, C.; D'Angelo, C.; Pagano, B.; Amato, J.; Salvati, E.; Sissi, C.; Pinato, O.; Marinelli, L.; et al. Exploring the chemical space of G-quadruplex binders: Discovery of a novel chemotype targeting the human telomeric sequence. *J. Med. Chem.* **2013**, *56*, 9646–9654. [CrossRef] [PubMed]
106. Kang, H.-J.; Park, H.-J. In silico identification of novel ligands for G-quadruplex in the c-MYC promoter. *J. Comp. Aided Mol. Des.* **2015**, *29*, 339–348. [CrossRef]
107. Hou, J.-Q.; Chen, S.-B.; Zan, L.-P.; Ou, T.-M.; Tan, J.H.; Luyt, L.G.; Huang, Z.-S. Identification of a selective G-quadruplex DNA binder using a multistep virtual screening approach. *Chem. Commun.* **2015**, *51*, 198–201. [CrossRef]
108. Monsen, R.C.; Trent, J.O. G-quadruplex virtual drug screening: A review. *Biochimie* **2018**, *152*, 134–148. [CrossRef]
109. Luo, J.; Wei, W.; Waldispühl, J.; Moitessier, N. Challenges and current status of computational methods for docking small molecules to nucleic acids. *Eur. J. Med. Chem.* **2019**, *168*, 414–425. [CrossRef]
110. Ortiz de Luzuriaga, I.; Lopez, X.; Gil, A. Learning to model G-quadruplexes: Current methods and perspectives. *Annu. Rev. Biophys.* **2021**, *50*, 209–243. [CrossRef]

111. Roy, S.; Ali, A.; Bhattacharya, S. Theoretical insight into the library screening approach for binding of intermolecular G-quadruplex RNA and small molecules through docking and molecular dynamics simulation studies. *J. Phys. Chem. B* **2021**, *125*, 5489–5501. [CrossRef]
112. Wei, D.; Parkinson, G.N.; Reszka, A.P.; Neidle, S. Crystal structure of a c-kit promoter quadruplex reveals the structural role of metal ions and water molecules in maintaining loop conformation. *Nucleic Acids Res.* **2012**, *40*, 4691–4700. [CrossRef]
113. Stump, S.; Mou, T.C.; Sprang, S.R.; Natale, N.R.; Beall, H.D. Crystal structure of the major quadruplex formed in the promoter region of the human c-MYC oncogene. *PLoS ONE* **2018**, *13*, e0205584. [CrossRef] [PubMed]
114. Ou, A.; Schmidberger, J.W.; Wilson, K.A.; Evans, C.W.; Hargreaves, J.A.; Grigg, M.; O'Mara, M.L.; Iyer, K.S.; Bond, C.S.; Smith, N.M. High resolution crystal structure of a KRAS promoter G-quadruplex reveals a dimer with extensive poly-A pi-stacking interactions for small-molecule recognition. *Nucleic Acids Res.* **2020**, *48*, 5766–5776. [CrossRef] [PubMed]
115. Li, K.; Yatsunyk, L.; Neidle, S. Water spines and networks in G-quadruplex structures. *Nucleic Acids Res.* **2021**, *49*, 519–528. [CrossRef]
116. Clark, G.R.; Pytel, P.D.; Squire, C.J. The high-resolution crystal structure of a parallel intermolecular DNA G-4 quadruplex/drug complex employing syn glycosyl linkages. *Nucleic Acids Res.* **2012**, *40*, 5731–5738. [CrossRef] [PubMed]
117. Clark, G.R.; Pytel, P.D.; Squire, C.J.; Neidle, S. Structure of the first parallel DNA quadruplex-drug complex. *J. Am. Chem. Soc.* **2003**, *125*, 4066–4067. [CrossRef]
118. Wei, W.; Luo, J.; Waldispühl, J.; Moitessier, N. Predicting positions of bridging water molecules in nucleic acid−ligand complexes. *J. Chem. Inf. Model.* **2019**, *59*, 2941–2951. [CrossRef] [PubMed]
119. de Beer, S.B.; Vermeulen, N.P.; Oostenbrink, C. The role of water molecules in computational drug design. *Curr. Top. Med. Chem.* **2010**, *10*, 55–66. [CrossRef]
120. Samways, M.L.; Taylor, R.D.; Bruce Macdonald, H.E.; Essex, J.W. Water molecules at protein-drug interfaces: Computational prediction and analysis methods. *Chem. Soc. Rev.* **2021**, *50*, 9104–9120. [CrossRef] [PubMed]
121. Nittinger, E.; Schneider, N.; Lamge, G.; Rarey, M. Evidence of water molecules—A statistical evaluation of water molecules based on electron density. *J. Chem. Inf. Model.* **2015**, *55*, 771–783. [CrossRef]
122. Ladbury, J.E. Just add water! The effect of water on the specificity of protein-ligand binding sites and its potential application to drug design. *Chem. Biol.* **1996**, *3*, 973–980. [CrossRef]
123. Dunitz, J.D. The entropic cost of bound water in crystals and biomolecules. *Science* **1994**, *264*, 670. [CrossRef]
124. Parkinson, G.N.; Cuenca, F.; Neidle, S. Topology conservation and loop flexibility in quadruplex-drug recognition: Crystal structures of inter- and intramolecular telomeric DNA quadruplex-drug complexes. *J. Mol. Biol.* **2008**, *381*, 1145–1156. [CrossRef]
125. Goddard, T.D.; Huang, C.C.; Meng, E.C.; Pettersen, E.F.; Couch, G.S.; Morris, J.H.; Ferrin, T.E. UCSF ChimeraX: Meeting modern challenges in visualization and analysis: UCSF chimera visualization system. *Protein Sci.* **2018**, *27*, 14–25. [CrossRef] [PubMed]
126. Parkinson, G.N.; Lee, M.P.; Neidle, S. Crystal structure of parallel quadruplexes from human telomeric DNA. *Nature* **2002**, *417*, 876–880. [CrossRef]
127. Neidle, S. Challenges in developing small-molecule quadruplex therapeutics. *Ann. Rep. Med. Chem.* **2020**, *54*, 517–546.
128. Molnár, O.R.; Végh, A.; Somkuti, J.; Smeller, L. Characterization of a G-quadruplex from hepatitis B virus and its stabilization by binding TMPyP4, BRACO19 and PhenDC3. *Sci. Rep.* **2021**, *11*, 23243. [CrossRef] [PubMed]
129. Frasson, I.; Nadai, M.; Richter, S.N. Conserved G-quadruplexes regulate the immediate early promoters of human *alphaherpesviruses*. *Molecules* **2019**, *24*, 2375. [CrossRef] [PubMed]
130. Butovskaya, E.; Soldà, P.; Scalabrin, M.; Nadai, M.; Richter, S.N. HIV-1 nucleocapsid protein unfolds stable RNA G-quadruplexes in the viral genome and is inhibited by G-quadruplex ligands. *ACS Infect. Dis.* **2019**, *5*, 2127–2135. [CrossRef]
131. Rudling, A.; Orro, A.; Carlsson, J. Prediction of ordered water molecules in protein binding sites from molecular dynamics simulations: The impact of ligand binding on hydration networks. *J. Chem. Inf. Model.* **2018**, *58*, 350–361. [CrossRef]
132. Giambaşu, G.M.; Case, D.A.; York, D.M. Predicting site-binding modes of ions and water to nucleic acids using molecular solvation theory. *J. Am. Chem. Soc.* **2019**, *141*, 2435–2445. [CrossRef] [PubMed]
133. Rappe, A.K.; Casewit, C.J.; Colwell, K.S.; Goddard, W.A., III; Skiff, W.M. UFF, a full periodic table force field for molecular mechanics and molecular dynamics simulations. *J. Am. Chem. Soc.* **1992**, *114*, 10024–10039. [CrossRef]
134. 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. G4LDB 2.2: A database for discovering and studying G-quadruplex and i-motif ligands. *Nucleic Acids Res.* **2021**, gkab952. [CrossRef] [PubMed]
135. Trott, O.; Olson, A.J. Software news and update Autodock vina: Improving the speed and accuracy of docking with a new scoring function, efficient optimization, and multithreading. *J. Comput. Chem.* **2010**, *31*, 455–461. [PubMed]