JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Article

pubs.acs.org/JACS

Structural Basis for Telomeric G-Quadruplex Targeting by Naphthalene Diimide Ligands

Gavin W. Collie, Rossella Promontorio, Sonja M. Hampel, Marialuisa Micco, Stephen Neidle,* and Gary N. Parkinson*

CRUK Biomolecular Structure Group, The School of Pharmacy, University of London, London, WC1N 1AX, United Kingdom

Supporting Information

ABSTRACT: The folding of the single-stranded 3′ end of the human telomere into G-quadruplex arrangements inhibits the overhang from hybridizing with the RNA template of telomerase and halts telomere maintenance in cancer cells. The ability to thermally stabilize human telomeric DNA as a four-stranded G-quadruplex structure by developing selective small molecule compounds is a therapeutic path to regulating telomerase activity and thereby selectively inhibit cancer cell growth. The development of compounds with the necessary selectivity and affinity to target parallel-stranded G-quadruplex structures has proved particularly challenging to date, relying heavily upon limited structural data. We report here on a structure-based approach to the design of quadruplex-binding ligands to enhance affinity and selectivity for human telomeric DNA. Crystal structures have been determined of complexes between a 22-mer intramolecular human telomeric quadruplex and two potent tetra-substituted naphthalene diimide compounds, functionalized with positively charged N-methyl-piperazine side-chains. These compounds promote parallel-stranded quadruplex topology, binding exclusively to the 3′ surface of each quadruplex. There are significant differences between the complexes in terms of ligand mobility and in the interactions with quadruplex grooves. One of the two ligands is markedly less mobile in the crystal complex and is more quadruplex-stabilizing, forming multiple electrostatic/hydrogen bond contacts with quadruplex phosphate groups. The data presented here provides a structural rationale for the biophysical (effects on quadruplex thermal stabilization) and biological data (inhibition of proliferation in cancer cell lines and evidence of *in vivo* antitumor activity) on compounds in this series and, thus, for the concept of telomere targeting with DNA quadruplex-binding small molecules.

INTRODUCTION

The terminal 2−10 kb of DNA in human chromosomes are composed of tandem G-rich d(TTAGGG) repeats.[1] Telomeric DNA terminates in a single-stranded 3′ overhang (ca. 100−200 nucleotides), which can self-associate through Watson−Crick and Hoogsteen hydrogen bonding to form higher-order structures, termed G-quadruplexes.[2] Several topological arrangements for four-repeat human telomeric quadruplexes have been reported,[3] depending on terminal sequences and environmental factors such as the nature of associated monovalent ions.

Telomeric DNA is associated with the Shelterin nucleoprotein complex, which protects chromosomal ends from end-to-end fusions and DNA damage and also with the enzyme telomerase, which plays a pivotal role in telomere regulation.[4] Telomerase actively extends the length of telomeric DNA by adding d(TTAGGG) repeats to the 3′ terminus, ensuring chromosomal integrity after each cell cycle. Importantly, telomerase is up-regulated in 85% of human cancers, with minimal activity in somatic cells, making telomerase, in principle, an attractive target for therapeutic intervention in cancer. A number of approaches for telomerase inhibition have been developed.[5] A particularly well-studied (albeit indirect) strategy for inhibition[6] involves the use of small molecule compounds to thermally stabilize G-quadruplex structures formed from these telomeric d(TTAGGG) repeats, with the aim of displacing telomerase and/or other key telomeric proteins such as hPOT1.[7] Many small molecules that bind to telomeric quadruplexes have been described.[8] None have been fully evaluated as potential anticancer therapeutics to date, in part because potency and selectivity is often suboptimal yet challenging to improve.[9] *In vivo* antitumor data is mostly restricted to acridine-derived compounds and the natural product telomestatin.[10,11]

The naphthalene diimide and closely related scaffolds have been developed as a suitable starting-point for quadruplex-binding ligands, in part because of their chemical accessibility and large planar surfaces suggestive of favorable quadruplex-binding.[12,13] In particular, we have disclosed a series of tetra-substituted naphthalene diimide-based (ND) compounds with exceptional biophysical and biological characteristics.[13] This recently developed series of ND compounds, with all identical side-chains composed of 3−5 carbons linking a positively charged methyl-piperazine functionality to the naphthalene diimide core, have been shown to significantly thermally

Received: October 31, 2011
Published: January 12, 2012

   dx.doi.org/10.1021/ja2102423 | J. Am. Chem.Soc. 2012, 134, 2723−2731

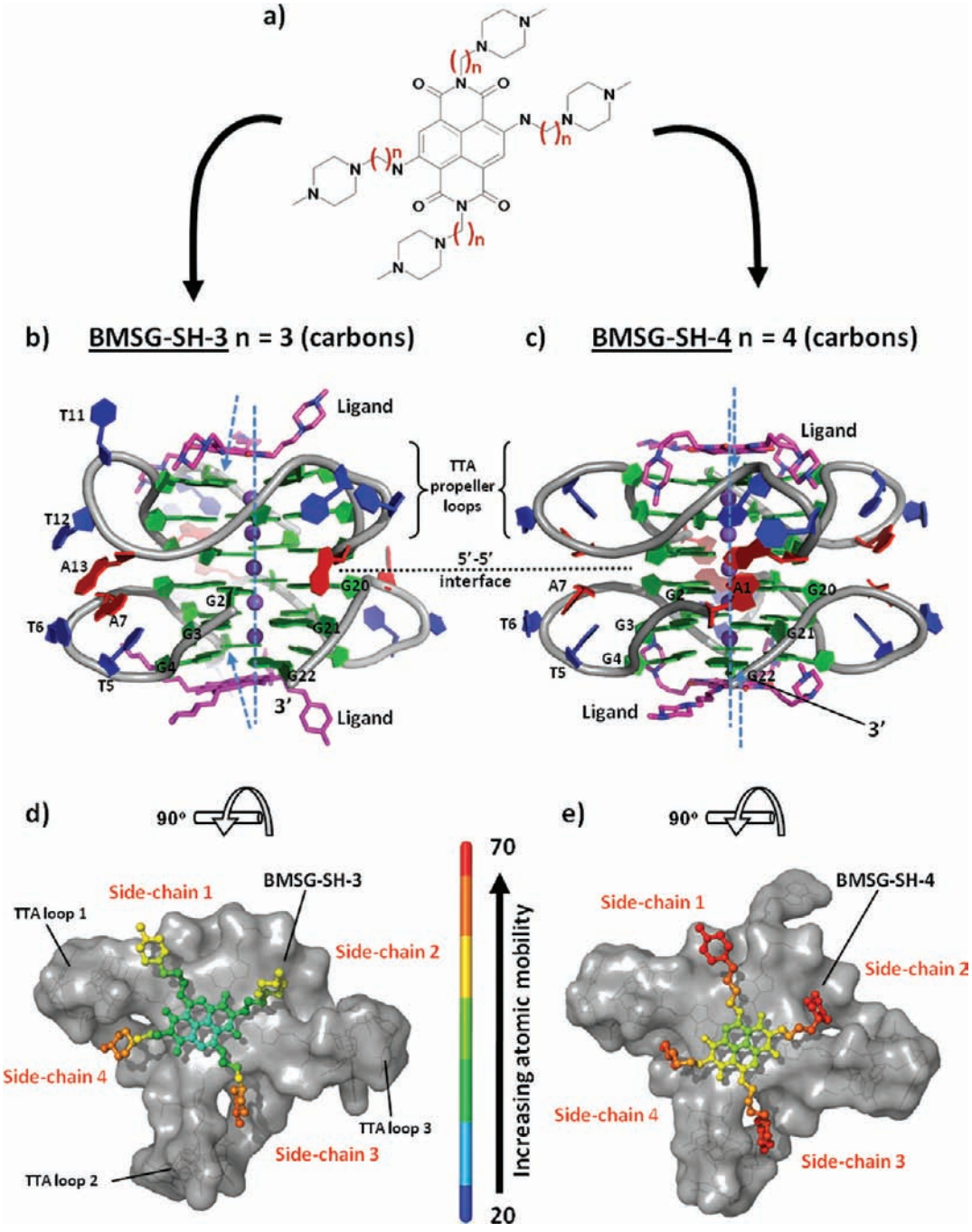

Figure 1. (a) Naphthalene diimide series (n = 2–5). (b, c) Cartoon representations of the BMSG-SH-3/Gtel22 and BMSG-SH-4/Gtel22 complex structures. The end-to-end quadruplex dimers are shown, highlighting packing, topology and ligand binding orientation. The potassium ions are shown as blue spheres. (d, e) Representations, projected onto the plane of the terminal G-quartets, of the Gtel22 quadruplex solvent-accessible binding surfaces with the bound ligands BMSG-SH-3 and BMSG-SH-4 shown colored by crystallographic temperature factors (or B-factors). Temperature factors (in Å^2) provide a measure of atomic mobility in the crystal lattice, and are defined as: $B_i = 8\pi^2 U_i^2$ (where U_i^2 is the mean square displacement of atom i).

stabilize human telomeric G-quadruplex DNA (with exceptionally high ΔT_m values) and are able to selectively inhibit the growth of several cancer cell lines at sub-μM concentrations.[14] They also effectively inhibit telomerase and displace hPOT1 from telomeric DNA, in accord with the hypothesis that these naphthalene diimides act at the cellular level through telomeric quadruplex thermal stabilization, as well as showing significant antitumor activity that is associated with telomerase inhibition, in a xenograft model for pancreatic cancer.[15]

We have approached the challenge of rationally developing improved tetra-substituted naphthalene diimide compounds by means of a structure-based approach employing crystallography. Structural data to date on quadruplex–small molecule complexes suggests a distinct preference for an all-parallel topology for telomeric quadruplexes when complexed with high-affinity heteroaromatic ligands.[16] All except one of these crystal structures are of complexes with bimolecular telomeric quadruplexes: the sole exception is of an ND compound bound

dx.doi.org/10.1021/ja2102423 | J. Am. Chem.Soc. 2012, 134, 2723–2731

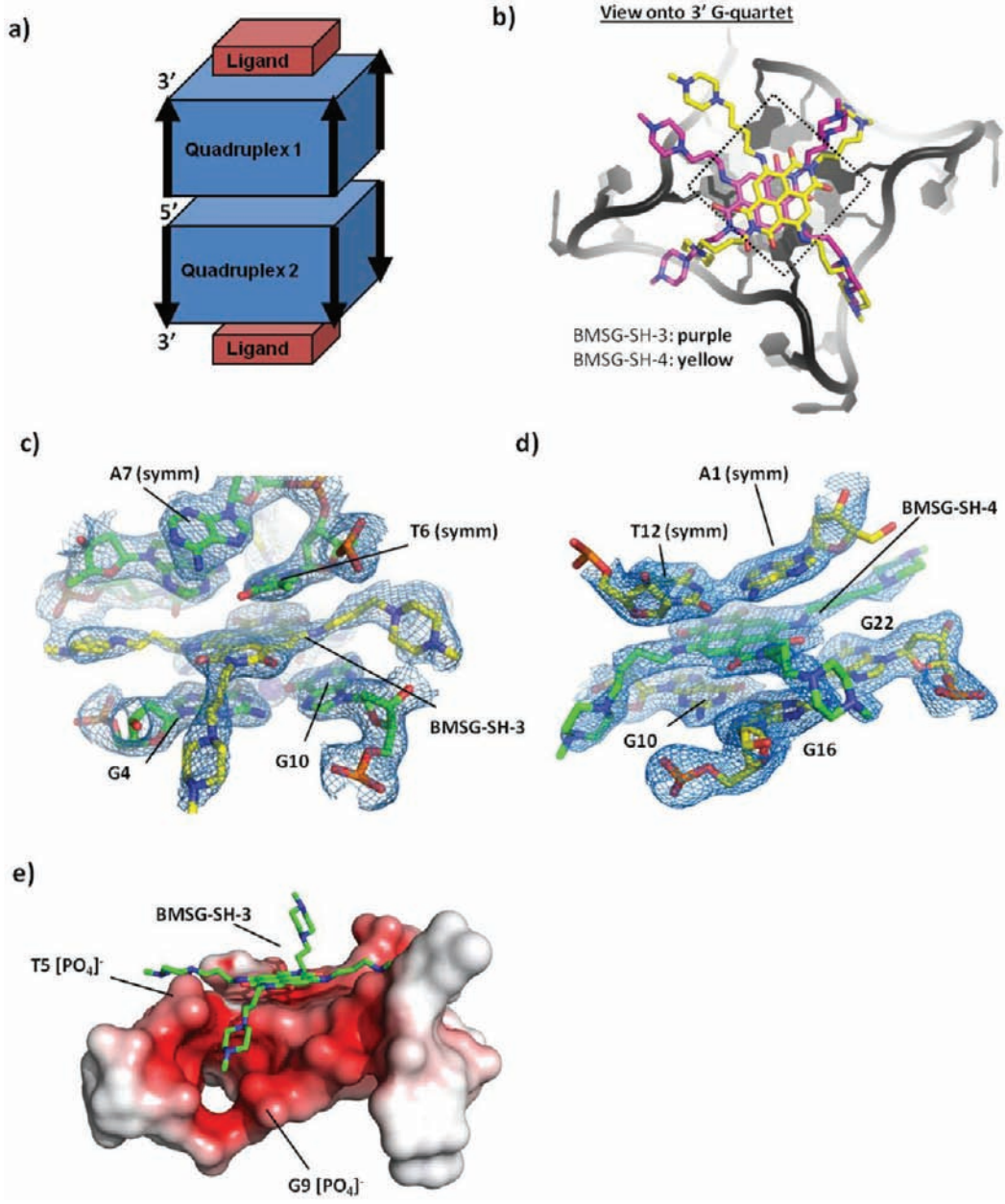

Figure 2. (a) Schematic showing quadruplex associations with 5′−5′ packing interactions and 3′ ligand binding. (b) rmsd overlap of BMSG-SH-3 and -SH4 onto G-quartet plane highlighting the asymmetric positioning of the ligands over the G-quartet. (c, d) Unequivocal placement of both ligands and side chains into the electron density ($2F_o-F_c$, drawn at the 0.9σ level). (e) Surface representation of the charged groove region of the BMSG-SH-3-Gtel22 complex.

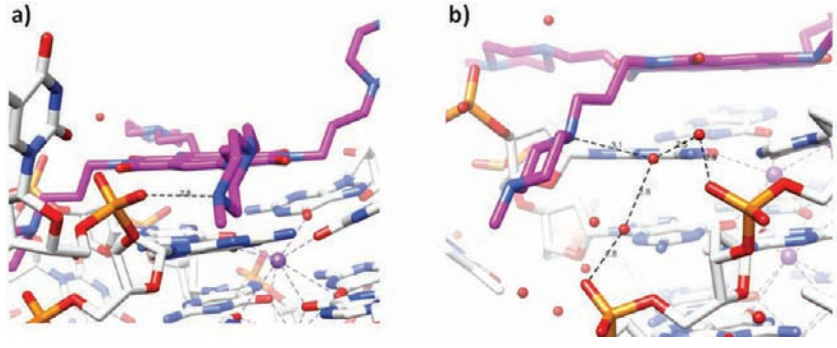

Figure 3. Views of two N-methyl-piperazine groups in the Gtel22-BMSG-SH-3 structure. The ND ligand is colored mauve and hydrogen bonds are shown as dashed lines. Water molecules are shown as red spheres. (a) Side-chain 3; (b) side-chain 4.

dx.doi.org/10.1021/ja2102423 | J. Am. Chem.Soc. 2012, 134, 2723−2731

to a 23-mer.[16d] In this structure, the flanking 5′-thymine is involved in secondary external stacking interactions with an externally bound ND molecule. The present study reports on two crystal structures of ND complexes with the identical sequence to the much-studied 22-mer d(AGGG[TTAGGG]$_3$) intramolecular G-quadruplex.[3a,g] This sequence does not suffer from any potential extraneous features caused by flanking sequences; thus, both complexes have 1:1 stoichiometry and may represent a biologically relevant view of a small-molecule ligand bound to a quadruplex within a human telomere in a high-affinity binding mode. The structures enable rationalizations to be made of the quadruplex affinity and biological activity of the ND compounds and others in the series, as well as indicating directions for their further improvement.

RESULTS

Overall Arrangement of the Quadruplex–Naphthalene Diimide Complexes. The ND molecules in both structures are stacked effectively over the terminal 3′ G-quartet surfaces, with extensive $\pi-\pi$ contacts (Figures 1d,e and 2–3), resulting in a 1:1 quadruplex/ligand stoichiometry. This stoichiometry is in contrast to two previously reported complex crystal structures, involving ND ligands with different terminal groups on the side-chains, both of which display multiple ND-binding sites with telomeric quadruplexes.[16d] The 1:1 stoichiometry is suggestive of a certain degree of specificity in the mode by which these ND compounds interact with telomeric G-quadruplex DNA. The intramolecular quadruplexes in both structures, when bound to the ND compounds BMSG-SH-3 and BMSG-SH-4, have parallel-stranded topology (Figure 1b,c) with three characteristic chain-reversal propeller-type loops, in agreement with all other crystallographically determined telomeric quadruplex native and ligand complex structures reported to date and with the topology determined by circular dichroism methods in solution.[14] This topology is also in accord with a recent NMR structure of the native Gtel22 quadruplex determined in molecular crowding conditions.[3g] Both the ND complexes involve two individual quadruplex molecules forming a dimer, by stacking at the 5′ G-quartet interface, an association which is mediated by a coordinating potassium ion. This dimer arrangement is a common one for both native and ligand-bound human telomeric-quadruplexes.[3a,16a−c] Here, in both structures, many of the ND side-chains are positioned within the G-quadruplex groove regions, interacting through hydrogen bonds, water bridges and electrostatic contacts with the negatively charged phosphate groups, as described in Figure 3. This is clear structural evidence of human G-quadruplex-directed ligand side-chains actively exploiting telomeric quadruplex-groove regions for effective contacts and binding.

Although the overall binding sites for the two ND analogues to the telomeric G-quadruplex are at first sight closely similar, there are several significant differences in the details of the binding (see below). The most striking differences between the complexes are seen in the mobilities of the two ligands, as measured by the crystallographic temperature factors (B-factors, <B>). The temperature-factor representations (Figure 1d,e) show that the four N-methyl-piperazine end-groups of BMSG-SH-3 are significantly less mobile (with a <B> of 51 Å^2) than those of BMSG-SH-4 (<B> = 62 Å^2). There is also a large difference in mobilities for the naphthalene diimide cores in the two structures, even though they are both stacked onto the G-quartet surface (BMSG-SH-3 core <B> = 35 Å^2: BMSG-SH-4 core <B> = 48 Å^2). The greater overall mobility of the BMSG-SH-4 molecule within its binding site would be expected to slightly destabilize its quadruplex binding compared to that of BMSG-SH-3. This is in accord with data from thermal denaturation measurements using a FRET (Fluorescence Resonance Energy Transfer) technique (Table 1), which shows that BMSG-SH-3 has greater quadruplex-stabilizing ability, with a significantly higher ΔT_m value compared to that for BMSG-SH-4 (Table 1).

Table 1. Melting Data for the ND Derivatives, Measured by a FRET Method[25] (°C; esds are ±0.1 °C, for 0.5 μM ligand concentrations)[a]

ligand	BMSG-SH-2	BMSG-SH-3	BMSG-SH-4	BMSG-SH-5
ND derivative side-chain length	n = 2	n = 3	n = 4	n = 5
Gtel22 ΔT_m by FRET	19.0	28.3	24.7	23.8
T-loop ΔT_m by FRET	0.8	1.3	0.1	0.2
IC_{50} for MCF7	1.01	0.17	0.10	0.20
IC_{50} for A549	0.47	0.11	0.07	0.26
IC_{50} for WI38	5.59	9.04	5.50	8.43

[a]Values for the n = 3–5 compounds have been previously reported.[14] Also shown are cell proliferation data using the SRB method and 96 h exposures for two cancer cell lines (MCF7 and A549) and a normal human fibroblast line (WI38). IC_{50} values, in μM, represent the mean of three independent determinations. Estimated standard deviations (ESDs) are ±0.01–0.1 μM. IC_{50} values for the n = 3-5 compounds have been presented previously.[14]

BMSG-SH-3-Quadruplex Binding and Interactions. The differences between the two ligands in terms of their ability to stabilize the telomeric quadruplex can be rationalized further when examining the groove interactions in detail. BMSG-SH-3 stacks effectively over only part of the 3′ G-quartet of the quadruplex and is asymmetrically positioned over the core potassium channel. This is a consequence of side-chain-groove contacts pulling the ligand toward loops 1 and 2 (Figures 1d and 4a). Two of the four side-chains (3 and 4) are positioned deep within quadruplex grooves. These interact strongly with the quadruplex, having close direct contacts between their positively charged methyl-piperazine ring nitrogen atoms and the negatively charged phosphates groups of the quadruplex grooves. Thus, side-chain 3 has a close direct contact with the T11 phosphate group (Figure 3a), and side-chain 4 is positioned equidistant between the T5 and G9 phosphates, sitting stably in this highly electronegative region and interacting with two groove phosphate groups via a network of water molecules (Figures 3b and 2e)—this is the only region of the BMSG-SH-3 molecule that interacts with the quadruplex through water contacts.

BMSG-SH-4-Quadruplex Binding and Interactions. The BMSG-SH-4 molecule is positioned more centrally on the terminal G-quartet than BMSG-SH-3, probably as a consequence of the extended length of its side-chains. At first sight a notable feature of the BMSG-SH-4 complex is the positioning of all four ligand side-chains in the quadruplex grooves (Figure 1e). However, contacts between BMSG-SH-4 and groove atoms are weaker and less well-defined compared to those in the BMSG-SH-3 complex, based on the quality of the electron density (Figure 2d) and significantly longer hydrogen-bond distances (Figure S1, Supporting Information). The high values for the temperature factors of the N-methyl-piperazine

dx.doi.org/10.1021/ja2102423 | J. Am. Chem.Soc. 2012, 134, 2723–2731

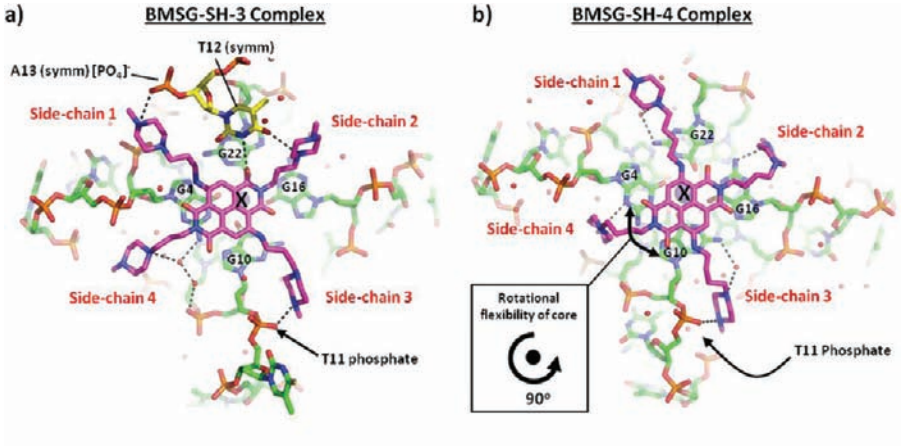

Figure 4. Views onto 3′ G-quartet surfaces of the (a) BMSG-SH-3-Gtel22 and (b) BMSG-SH-4-Gtel22 complexes. The black crosses indicate the center of the potassium channels, with hydrogen bonds shown as black dashes.

groups in BMSG-SH-4 are suggestive of disorder for the side-chains, which would further weaken their groove contacts. Thus, paradoxically, an increase in the number of potential groove contacts available to BMSG-SH-4 results in a reduction in the number of strong specific contacts compared to those involving BMSG-SH-3. It may be concluded that the increased length and mobility of the BMSG-SH-4 side-chains, besides having a greater degree of intrinsic flexibility, are clearly too long for effective fitting into the grooves. Two side-chains have less favorable *syn* dihedral angles—these are required in order for these side-chains to have significant groove contacts.

Additionally, based on the quality and diffuse appearance of the electron density for the core region of BMSG-SH-4, it is likely that several rotational conformations of the ligand are present in the structure (Figure 2d). This is a feature of some tetra-substituted ND-quadruplex binding which has been previously noted,[16d] and arises from the pseudo-4-fold symmetry of these compounds, which allows four orientations of the compound to be accommodated in the crystal lattice, despite an absence of crystallographic 4-fold symmetry. As a consequence of this rotational conformational flexibility, there is a difference in the orientation of the secondary-amine-linked side chains when comparing the BMSG-SH-3 and BMSG-SH-4 complexes (see Figure 2b).

BMSG-SH-2 and BMSG-SH-5 Quadruplex Binding and Biology. Diffraction data for the Gtel22-BMSG-SH-5 complex have revealed similar unit cell dimensions and identical space group to the BMSG-SH-4 complex, suggestive of an analogous packing arrangement and mode of ligand binding. The extended side-chains and core of BMSG-SH-5 would be expected to have even greater mobility than found for BMSG-SH-4 in its Gtel22 complex, consistent with its reduced ΔT_m value. The poor quality of the Gtel22-BMSG-SH-5 complex crystal diffraction data could be indicative of high thermal motion, in accord with this finding.

Simple qualitative molecular modeling studies on the ND analogue with short (n = 2) side chains (Table 1), using the Gtel22 template structure reported here, indicates that these side-chains are significantly too short for effective groove interactions (although the naphthalene diimide core can still stack effectively), suggesting that the overall binding would be reduced compared to the other three analogues. This is in accord with the experimental FRET melting data reported in Table 1.

The effects of all four ligands on the growth of two cancer cell lines and one normal line are summarized in Table 1. These show that the BMSG-SH-3 and-4 compounds are overall the most potent growth inhibitors, followed by BMSG-SH-5; the n = 2 compound is significantly less active in these cancer lines.

DISCUSSION

The crystallographic data presented here highlight three major features that have not been observed to date for quadruplex-ligand complexes: (i) the 1:1 binding of a ND compound to a human intramolecular telomeric G-quadruplex, (ii) the role of positively charged side-chains in quadruplex-groove binding, (iii) the retention of the parallel topology observed in the native Gtel22 crystal structure.[3a] A previous structural study with a distinctly substituted ND ligand had shown in that case the exclusive involvement of loops to generate ND-binding platforms, in addition to the planar ND core binding to both 5′ and 3′ G-quartet, but did not have the expected extensive side-chain-groove interactions.[16d] This is likely to be a consequence of the choice of functional groups on the side-chains rather than chain length *per se*. The present results show that the presence of the N-methyl-piperazine group, when appropriately positioned, is a major determinant of both binding site affinity and direct groove interactions, supporting its future application in other ligand design studies. It is likely that the combination of binding site affinity and direct groove interactions are responsible for the 1:1 stoichiometry observed for these complexes, such stoichiometry suggestive of a level of specificity in the manner by which these methyl-piperazine-modified ND compounds recognize G-quadruplex DNA. Side-chain length, at least for these particular compounds, seems to have an optimal upper limit, as the progression from n = 3 to n = 5 leads to increased conformational flexibility, with n = 5 resulting in significantly poorer quality diffraction data, possibly as a consequence of increased internal motion and thus reduced order in the crystal. Melting (FRET) data shows all three compounds are able to substantially stabilize human telomeric G-quadruplex DNA, although there are real differences between them. There is a clear trend in stabilization, with BMSG-SH-3 being optimal, which correlates well with the crystallographic results reported here.

Previous telomeric quadruplex-small molecule crystal structures have shown that side-chains can induce changes in loop and terminal flanking residue conformations together with

dx.doi.org/10.1021/ja2102423 | J. Am. Chem.Soc. 2012, 134, 2723−2731

groove dimensions in order to maximize nonbonded interactions.[16] Direct cationic group-phosphate interactions have not been previously observed in these structures, even though they have been widely assumed to take place. The crystal structure of the bimolecular complex with the acridine ligand BRACO-19[16e] has an extreme example of this, with the two pyrrolidine charged side-chains becoming fully enclosed in binding pockets formed by such conformational changes. However, no direct phosphate-cation interactions were observed although there are extensive water-mediated contacts between them. The recent NMR structure of a 2:1 complex between the parallel topology *c-myc* quadruplex and a monosubstituted quindoline compound[17] also shows large conformational changes (in flanking residues) to accommodate the bound molecule. A hydrogen bond between the cationic N-diethyl group terminal side-chain group of the quinolone and an O4 atom of a thymine in this quadruplex has been proposed in this structure although it was not directly observed. It thus appears that direct phosphate-cation contacts are not invariably formed with cationic quadruplex-binding ligands, so their presence is a consequence of a balance of recognition factors, principally the position in which the planar chromophore is held. This is the determinant of where the side-chains are positioned, with the location of the terminal groups depending on their length and flexibility. Comparison with the crystal structure of the BRACO19-quadruplex complex[16e] shows that the three substituents attached to the acridine core of BRACO-19 occupy very similar regions of groove space found here for BMSG-SH-3 (Figure 5). The 9-substituent in BRACO-19 is an

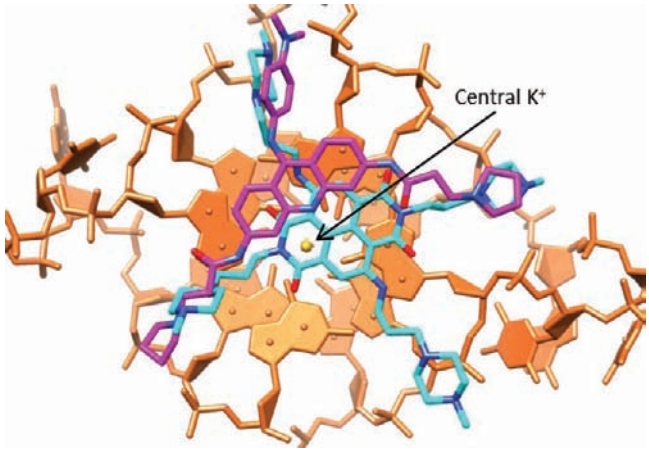

Figure 5. Structural alignment of the native human 12-mer G-quadruplex 3CE5 (orange) in its complex with the trisubstituted acridine compound BRACO19 (colored purple) showing only the naphthalene ligand BMSG-SH-3 (colored cyan) from the 22-mer complex. The overlay has used all the DNA atom positions for alignment.

aniline group, which has not to date been explored in the naphthalene diimide series. It may well be useful in not only being an effective van der Waals groove-binding moiety, with also the potential advantage of modulating their polarity for cell-based studies.

It is notable that the present crystal structures show that the greater mobility of the side-chains in BMSG-SH-4 can be transmitted to the ND core motif, while still retaining their ability to interact (albeit loosely) with all four quadruplex grooves. Side-chain length has been shown to affect flexibility, and consequently the number and type of interactions available to terminal functional groups (here the positively charged N-methyl-piperazine groups). This adaptability may in turn allow multiple conformations of the ND core to be accommodated over the large and open G-quartet binding platform, consistent with the more diffuse electron density maps of the BMSG-SH-4 complex.

All the compounds discussed here are potent inhibitors of cell growth, with particular activity in the two cancer cell lines reported here (Table 1). To what extent are the differences and trends in this data borne out by the two crystal structures and the melting data? Although they altogether suggest that BMSG-SH-4 should be less biologically effective, this is countered by the increased lipophilicity due to the slightly longer side-chains, which improves log P_{calc} by 1.1 units. The greater lipophilicity of the $n = 5$ compound is outweighed by its reduced quadruplex binding, resulting in a significant diminution of biological activity compared to the $n = 3$ and $n = 4$ compounds. The $n = 2$ compound is altogether less active, presumably as a result of the decrease in both quadruplex stabilization and lipophilicity.

The features found in the structures reported here imply that the exploitation of the phosphate and other polar groups in quadruplex groove regions can be an important element of designing telomeric G-quadruplex-directed compounds. It is clear that flexible side-chains can allow multiple groove contacts to be made and the large grooves in this particular quadruplex enable the side-chains to adopt a variety of conformations (see the overlay in Figure 2b). In view of the well-documented conformational variability of telomeric quadruplex loops,[16e,f] rather than actively targeting phosphate groups of the farthest distant loop residues such as A6 and T7, which lie deep within the grooves (which themselves adopt unpredictable conformations), it may be fruitful to target the phosphates which lie closest to the G-quartet, such as those of T5, T11 and T17. These are conformationally constrained by their proximity to the G-quartet scaffold. Indeed, from the two crystal structures reported here, the strongest charge−charge interactions between the quadruplex and both ND compounds are formed between the positively charged piperazine nitrogen and the phosphate group of T11. The use of nonequivalent side-chain lengths (for example, a mix of 3-carbon and 4-carbon side chains linking methyl-piperazine groups to an ND core) may be a route to achieving such G-quadruplex-ligand interactions.

A further strategy for enhancing the affinity of these ND compounds for telomeric quadruplex DNA could be to increase the size of the side-chain terminal groups, even though this would be at the expense of ligand molecular weight. The excess groove spaces, in the cases of the current ligands, contain water molecules (Figures 2e, 3), which can be readily displaced by larger ligand side-chains. The present ND molecules have molecular weights in excess of 850 Da, far higher than conventional small-molecule drugs that obey the Lipinski Rule of Five.[18] Whether they could be administered as orally bioavailable drugs has yet to be determined, but their large molecular size and four cationic charges do not hinder their ready uptake into cancer cells.[13−15] Their size may also hinder their ability to interact with nonquadruplex targets such as enzyme active sites or ion channels—for example, BMSG-SH-3 and -4 do not adversely affect the therapeutically important hERG cardiac ion channel.[15] Whether yet larger molecules would behave in a similar manner remains to be determined, but judicious attention to the hydrophobic/hydrophilic balance

dx.doi.org/10.1021/ja2102423 | J. Am. Chem.Soc. 2012, 134, 2723−2731

Table 2. Crystallographic Statistics for Three DNA G-Quadruplex ND Complexes

	data collection		
	complex 1	complex 2	complex 3
Sequence	d(AGGG[TTAGGG]$_3$)	d(AGGG[TTAGGG]$_3$)	d(AGGG[TTAGGG]$_3$)
Ligand	BMSG-SH-3 (n = 3)	BMSG-SH-4 (n = 4)	BMSG-SH-5 (n = 5)
Space group	$P3_121$	$P6$	$P6$
Unit cell dimensions			
a, b, c (Å)	50.87, 50.87, 52.46	63.03, 63.03, 42.39	64.25, 64.25, 43.1
α, β, γ (deg)	90.00, 90.00, 120.00	90.00, 90.00, 120.00	90.00, 90.00, 120.00
Resolution (Å)	26.23−2.17 (2.23−2.17)	54.59−1.94 (1.99−1.94)	Not Integrated
R_{int} (%) overall	3.3 (59.0)	6.0 (66.0)	
I/σ	32.1 (3.5)	15.2 (2.0)	
Completeness (%)	99.5 (100.0)	99.6 (100.0)	
Redundancy	9.5 (9.7)	8.9 (7.0)	
Refinement			
Resolution (Å)	10.30 − 2.30	10.20 − 2.10	
Reflections	3467	5372	
R_{work}/R_{free} (%)	25.8/28.8	24.5/28.2	
No. of atoms	541	572	
Ions	2.5	2.5	
Water	29	38	
Overall B-factor (Å^2)	45.26	40.69	
rms deviations			
Bond-lengths (Å)	0.007	0.008	
Bond-angles (deg)	1.048	1.623	
PDB ID	3SC8	3T5E	

[a]Values in brackets refer to highest resolution shells.

of further groups may ensure that cellular uptake at least is unaffected.

METHODS AND MATERIALS

Synthesis of BMSG-SH-2, 2,7-Bis(2-(4-methylpiperazin-1-yl)ethyl)-4,9-bis((2-(4-methylpiperazin-1-yl)ethyl)amino)-benzo[lmn] [3,8]phenanthroline-1,3,6,8(2H,7H)-tetraone. 2,6-Dibromo-1,4,5,8-naphthalenetetracarboxylic acid[13] (0.2 g, 0.47 mmol), 1-(2-aminoethyl)-4-methylpiperazine (0.5 mL, 3.5 mmol), and NMP (2 mL) were reacted together according to the published procedure.[13,14] The final product was obtained as a blue semisolid (0.03 g, 8% yield) following HPLC purification using a published procedure.[14] 1H NMR (400 MHz, CDCl3, TMS): δ 9.52 (t, 2H, J = 5.4 Hz), 8.41 (br s, 2H), 8.14 (s, 2H), 4.76−4.92 (m, 8H), 4.34 (t, 4H, J = 6.8 Hz), 3.59 (q, 4H, J = 5.2 Hz, J = 6 Hz), 2.76−2.86 (m, 32H), 2.54 (s, 6H), 2.51 (s, 6H). ^{13}C NMR (100 MHz, CDCl3, TMS): δ 167.4, 165.9, 149.0, 125.8, 121.4, 118.7, 102.2, 56.1, 55.2, 54.0, 53.8, 51.2, 51.1, 44.3, 44.1, 40.2, 37.2. HRMS (ES$^+$) calculated for (M + 3H)$^{3+}$ $C_{42}H_{64}N_{12}O_4$ 803.541, found 803.5395.

Crystallography. The DNA G-quadruplex forming sequence d(AGGG[TTAGGG]$_3$) was initially dissolved in water to a final concentration of 2.5 mM single-stranded DNA (ssDNA). Buffer and salts were then added to the DNA to final concentrations of 2 mM ssDNA, 50 mM potassium chloride and 20 mM potassium cacodylate (pH 6.5). The buffered DNA was then annealed, which involved heating the sample to 85 °C in a heat block for 5 min followed by slow cooling to room temperature overnight. The ligands BMSG-SH-3, -4, and -5 were synthesized as described previously[14] and purified by reverse-phase HPLC. Stock solutions of each in water were mixed separately with the DNA 22-mer at equimolar ratios, and crystals grown in standard hanging drops. Crystals of the BMSG-SH-3-DNA complex were grown at 12 °C in a drop containing 20% PEG400, 100 mM lithium sulfate and 50 mM sodium cacodylate (pH 6.5). Crystals of the BMSG-SH-4-DNA complex were grown at 10 °C in a drop containing 15% PEG400, 300 mM potassium bromide, and 50 mM sodium cacodylate (pH 6.5). Crystals of the BMSG-SH-5-DNA complex were grown at 10 °C in a drop containing 20% PEG400, 300 mM potassium bromide, and 50 mM sodium cacodylate (pH 6.5). Crystals of the BMSG-SH-4-DNA complex were cryoprotected in a solution composed of 15% glycerol, 7.5% PEG400, 150 mM potassium bromide and 25 mM sodium cacodylate (pH 6.5) prior to X-ray diffraction experiments. Crystals of all three complexes were flash-frozen in liquid nitrogen and data collected at the Diamond Light Source synchrotron (beamlines I04−1, I03 and I24 for the BMSG-SH-3, -4, and -5 complexes respectively). All data sets were processed and scaled using the XDS, SCALA and XIA2 programs.[19] The crystal structure of the 22-mer BMSG-SH-3 complex was solved by molecular replacement using the PHASER program,[20] with the native 22-mer telomeric quadruplex crystal structure 1KF1 as a search model.[3a] Model building and refinement were performed using the COOT[21] and REFMAC5[22] programs. Initial $2F_o-F_c$ maps showed clear electron density for the core G-quartets and potassium ions, as well as residual density for the loops and other regions omitted from the initial search model. A large region of electron density was visible in both F_o-F_c and $2F_o-F_c$ maps above the 3′ G-quartet, into which the naphthalene compound could be readily fitted. The structure was refined to a resolution of 2.3 Å, with final R_{work} and R_{free} values of 25.8 and 28.8% respectively. The 22-mer BMSG-SH-4 complex structure was solved, built and refined as described above. All 22 residues of the DNA could be readily fitted into electron density; however, electron density for the ligand (above the 3′ G-quartet) was more diffuse than for the BMSG-SH-3 complex. The naphthalene diimide core of BMSG-SH-4 could be readily identified and refined but electron density for the majority of the side-chains was much weaker. The structure was refined to a resolution of 2.1 Å, with final R_{work} and R_{free} values of 24.5 and 28.2% respectively. Diffraction data collected for the 22-mer BMSG-SH-5 complex were of poor quality and could not be integrated. However the space group and unit cell constants could be confidently determined as: $P6$ with $a = b$ = 64.25 Å, c = 43.1 Å, $\alpha = \beta$ = 90°, γ = 120°. Data collection and refinement statistics for the ND-quadruplex complexes are shown in Table 2. Structure factors and coordinates for the BMSG-SH-3 and BMSG-SH-4 complexes have been deposited in the Protein Data Bank with accession codes 3SC8 and 3T5E respectively. Programs CHIMERA[23] and PYMOL[24] were used for structure drawing.

Sulforhodamine B Short-term Cytotoxicity Assay. Human cancer cell lines, breast (MCF7), lung (A549), and normal human lung fibroblast lines (WI-38) were purchased from American Type Cell Culture (ATCC). Cell lines were maintained in appropriate medium supplemented with 10% fetal bovine serum (Invitrogen, U.K.), 2 mM L-glutamine (Invitrogen, Netherlands), and other components as specified by the suppliers. All cell lines were maintained at 37 °C, 5% CO_2 and routinely passaged. Short-term growth inhibition was measured using the SRB assay as described previously.[13b] Briefly, cells were seeded at appropriate densities into the wells of 96-well plates in their corresponding medium and incubated overnight to allow the cells to attach. Subsequently cells were exposed to freshly made solutions of ligands and incubated for a further 96 h. Following this the cells were fixed with ice cold trichloroacetic acid (TCA) (10%, w/v) for 30 min and stained with 0.4% SRB dissolved in 1% acetic acid for 15 min. All incubations were carried out at room temperature except for TCA fixation which was at 4 °C. The IC_{50} value, the concentration required to inhibit cell growth by 50%, was determined from the mean absorbance at 540 nm for each drug concentration expressed as a percentage of the control untreated well absorbance. All determinations were performed in triplicate.

FRET Assay. The ability of BMSG-SH-3 to -5 to stabilize the human telomeric G-quadruplex DNA sequence 5′-FAM-d(GGG-[TTAGGG]$_3$)-TAMRA-3′ and the duplex sequence 5′-FAM-dTA-TAGCTATA-HEG-TATAGCTATA-TAMRA-3′ (HEG linker: [(−CH_2−CH_2−O−)$_6$]) was investigated using a fluorescence resonance energy transfer (FRET) assay modified to be used as a high-throughput screen in a 96-well format.[25] The labeled oligonucleotides had attached the donor fluorophore FAM, 6-carboxyfluorescein, and the acceptor fluorophore TAMRA, 6-carboxytetramethylrhodamine. The FRET probe sequences were diluted from stock to the correct concentration (400 nM) in a 60 mM potassium cacodylate buffer (pH 7.4) and then annealed by heating to 85 °C for 10 min, followed by cooling to room temperature in the heating block. The compound was stored as a 10 mM stock solution in DMSO; final solutions (at 2× concentration) were prepared using 10 mM HCl in the initial 1:10 dilution, after which 60 mM potassium cacodylate buffer (pH 7.4) was used in all subsequent steps. The maximum HCl concentration in the reaction volume (at a ligand concentration of 20 μM) is thus 200 μM, well within the range of the buffer used. Relevant controls were also performed to check for interference with the assay. Ninety-six-well plates (MJ Research, Waltham, MA) were prepared by aliquoting 50 μL of the annealed DNA into each well, followed by 50 μL of the compound solutions. Measurements were made on a DNA Engine Opticon (MJ Research) with excitation at 450−495 nm and detection at 515−545 nm. Fluorescence readings were taken at intervals of 0.5 °C in the range 30−100 °C, with a constant temperature being maintained for 30 s prior to each reading to ensure a stable value. Final analysis of the data was carried out using a script written in the program Origin 7.0 (OriginLab Corp., Northampton, MA). The advanced curve-fitting function in Origin 7.0 was used for calculation of ΔT_m values. All determinations were performed in triplicate or better. Esds in ΔT_m are ±0.1 °C.

■ ASSOCIATED CONTENT

Supporting Information

Additional descriptive figures, and complete ref 10. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*stephen.neidle@pharmacy.ac.uk; gary.parkinson@pharmacy.ac.uk

■ ACKNOWLEDGMENTS

We are grateful to Cancer Research U.K. for support (Programme grant No. C129/A4489), the School of Pharmacy for a research studentship to G.C. and the Diamond Light Source for access to Synchrotron facilities. We thank Dr. Philip Brown for a grant to S.N. and the University of Bologna for a Fellowship to M.M.

■ REFERENCES

(1) de Lange, T., Lundblad, V., Blackburn, E. H., Eds. *Telomeres*, 2nd ed.; Cold Spring Harbor Press: New York, 2006.

(2) Davis, J. T. *Angew. Chem., Int. Ed.* **2004**, *43*, 668. Burge, S.; Parkinson, G. N.; Hazel, P.; Todd, A. K.; Neidle, S. *Nucleic Acids Res.* **2006**, *34*, 5402. Phan, A. T.; Kuryavyi, V.; Luu, K. N.; Patel, D. J. *Nucleic Acids Res.* **2007**, *35*, 6517. Phan, A. T. *FEBS J.* **2010**, *277*, 1107−1117.

(3) (a) Parkinson, G. N.; Lee, M. P. H; Neidle, S. *Nature* **2002**, *417*, 876−880. (b) Phan, A. T.; Modi, Y. S.; Patel, D. J. *J. Am. Chem. Soc.* **2004**, *126*, 8710−8716. (c) Ambrus, A.; Chen, D.; Dai, J.; Jones, R. A.; Yang, D. *Biochemistry* **2005**, *44*, 2048−2058. (d) Ambrus, A.; Chen, D.; Dai, J.; Bialis, T.; Jones, R. A.; Yang, D. *Nucleic Acids Res.* **2006**, *34*, 2723−2735. (e) Dai, J.; Carver, M.; Punchihewa, M.; Jones, R. A.; Yang, D. *Nucleic Acids Res.* **2007**, *35*, 4927−4740. (f) Lim, K. W.; Amrane, S.; Bouaziz, S.; Xu, W.; Mu, Y.; Patel, D. J.; Luu, K. N.; Phan, A. T. *J. Am. Chem. Soc.* **2009**, *131*, 4301−4309. (g) Heddi, B.; Phan, A. T. . *J. Am. Chem. Soc.* **2011**, *133*, 9824−9833.

(4) de Lange, T. *Genes Dev.* **2005**, *19*, 2100−2110. de Lange, T. *Cold Spring Harbor Symp. Quant. Biol.* **2010**, *75*, 167−177.

(5) See, for example: Kelland, L. R. *Clin. Cancer Res.* **2007**, *13*, 4960−4963. De Cian, A.; Lacroix, L.; Douarre, C.; Temime-Smaali, N.; Trentesaux, C.; Riou, J.-F.; Mergny, J.-L. *Biochimie* **2008**, *90*, 131−155. Shay, J. W.; Wright, W. E. *Cancer Cell* **2010**, *2*, 257−265. Ouellette, M. M.; Wright, W. E.; Shay, J. W. *J. Cell. Mol. Med.* **2011**, *15*, 1433−1442.

(6) Zahler, A. M.; Williamson, J. R.; Cech, T. R.; Prescott, D. M. *Nature* **1991**, *350*, 718−750. Sun, D.; Thompson, B.; Cathers, B. E.; Salazar, M.; Kerwin, S. M.; Trent, J. O.; Jenkins, T. C.; Neidle, S.; Hurley, L. H. *J. Med. Chem.* **1997**, *40*, 2113−2116. De Cian, A.; Lacroix, L.; Douarre, C.; Temime-Smaali, N.; Trentesaux, C.; Riou, J.-F.; Mergny, J.-L. *Biochimie* **2008**, *90*, 131−155. Wang, Q.; Liu, J.-Q.; Chen, Z.; Zheng, K.-W.; Chen, C.-Y.; Hao, Y.-H.; Tan, Z. *Nucleic Acids Res.* **2011**, *39*, 6229−6237.

(7) Zaug, A. J.; Podell, E. R.; Cech, T. R. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 10864−10869. Gomez, D.; O'Donohue, M. F.; Wenner, T.; Douarre, C.; Macadré, J.; Koebel, P.; Giraud-Panis, M.-J.; Kaplan, H.; Kolkes, A.; Shin-Ya, K.; Riou, J.-F. *Cancer Res.* **2006**, *66*, 7908−7912. Rodriguez, R.; Müller, S.; Yeoman, J. A.; Trentesaux, C.; Riou, J.-F.; Balasubramanian, S. *J. Am. Chem. Soc.* **2008**, *130*, 15758−15759.

(8) See for example: Monchaud, D.; Teulade-Fichou, M.-P. *Org. Biomol. Chem.* **2008**, *6*, 627−636. Arola, A.; Vilar, R. *Curr. Top. Med. Chem.* **2008**, *8*, 1405−1415. Neidle, S. *Therapeutic Applications of Quadruplex Nucleic Acids*; Academic Press: San Diego, 2011.

(9) Neidle, S. *FEBS J.* **2010**, *277*, 1118−1125.

(10) Gowan, S. M.; Harrison, J. R.; Patterson, L.; Valenti, M.; Read, M. A.; Neidle, S.; Kelland, L. R. *Mol. Pharmacol.* **2002**, *61*, 1154−1162. Burger, A. M.; Dai, F.; Schultes, C. M.; Reszka, A. P.; Moore, M. J.; Double, J. A.; Neidle, S. *Cancer Res.* **2005**, *65*, 1489−1496. Leonetti, C.; Scarsella, M.; Riggio, G.; Rizzo, A.; Salvati, E.; D'Incalci, M.; Staszewsky, L.; Frapolli, R.; Stevens, M. F.; Stoppacciaro, A.; Mottolese, M.; Antoniani, B.; Gilson, E.; Zupi, G.; Biroccio, A. *Clin. Cancer Res.* **2008**, *14*, 7284−7291. Salvati, E.; et al. *Clin. Invest.* **2007**, *117*, 3236−3247.

(11) Tauchi, T.; Shin-ya, K.; Sashido, G.; Sumi, M.; Okabe, S.; Ohyashiki, J. H.; Ohyashiki, K. *Oncogene* **2006**, *25*, 5719−5725.

(12) Di Antonio, M.; Doria, F.; Richter, S. N.; Bertipaglia, C.; Mella, M.; Sissi, C.; Palumbo, M.; Freccero, M. *J. Am. Chem. Soc.* **2009**, *131*, 13132−13141. Nadai, M.; Doria, F.; Di Antonio, M.; Sattin, G.; Germani, L.; Percivalle, C.; Palumbo, M.; Richter, S. N.; Freccero, M.

Biochimie **2011**, *93*, 1328−1340. Peduto, A.; Pagano, B.; Petronzi, C.; Massa, A.; Esposito, V.; Virgilio, A.; Paduano, F.; Trapasso, F.; Fiorito, F.; Florio, S.; Giancola, C.; Galeone, A.; Filosa, R. *Bioorg. Med. Chem.* **2011**, *19*, 6419−6429.
(13) (a) Cuenca, F.; Greciano, O.; Gunaratnam, M.; Haider, S.; Munnur, D.; Nanjunda, R.; Wilson, W. D.; Neidle, S. *Bioorg. Med. Chem. Lett.* **2010**, *18*, 1668−1673. (b) Gunaratnam, M.; Swank, S.; Haider, S. M.; Galesa, K.; Reszka, A. P.; Beltran, M.; Cuenca, F.; Fletcher, J. A.; Neidle, S. *J. Med. Chem.* **2009**, *52*, 3774−3783.
(14) Hampel, S. M.; Sidibe, A.; Gunaratnam, M.; Riou, J.-F.; Neidle, S. *Bioorg. Med. Chem. Lett.* **2010**, *20*, 6459−6463.
(15) Gunaratnam, M.; de la Fuente, M.; Hampel, S. M.; Todd, A. K.; Reszka, A. P.; Schatzlein, A.; Neidle, S. *Bioorg. Med. Chem.* **2011**, *19*, 7151−7157.
(16) (a) Haider, S. M.; Parkinson, G. N.; Neidle, S. *J. Mol. Biol.* **2003**, *326*, 117−125. (b) Campbell, N. H.; Patel, M.; Tofa, A. B.; Ghosh, R.; Parkinson, G. N.; Neidle, S. *Biochemistry* **2009**, *48*, 1675−1680. (c) Parkinson, G. N.; Ghosh, R.; Neidle, S. *Biochemistry* **2007**, *46*, 2390−2397. (d) Parkinson, G. N.; Cuenca, F.; Neidle, S. *J. Mol. Biol.* **2008**, *381*, 1145−1156. (e) Campbell, N. H.; Parkinson, G. N.; Reszka, A. P.; Neidle, S. *J. Am. Chem. Soc.* **2008**, *130*, 6722−6724. (f) Collie, G. W.; Sparapani, S.; Parkinson, G. N.; Neidle, S. *J. Am. Chem. Soc.* **2011**, *133*, 2721−2728.
(17) DaiJ.CarverM.HurleyL. H.YangD.J. Am. Chem. Soc.2011
(18) Lipinski, C. A.; Lombardo, F.; Dominy, B. W.; Feeney, P. J. *Adv. Drug Delivery Rev.* **2001**, *46*, 3−26.
(19) Collaborative Computational Project Number 4. *Acta Crystallogr.* **1994**, *D50*, 760−763.
(20) McCoy, A. J.; Grosse-Kunstleve, R. W.; Adams, P. D.; Winn, M. D.; Storoni, L. C.; Read, R. J. *J. Appl. Crystallogr.* **2007**, *40*, 658−674.
(21) Emsley, P.; Lohkamp, B.; Scott, W. G.; Cowtan, K. *Acta Crystallogr.* **2010**, *D66*, 486−501.
(22) Murshudov, G. N.; Vagin, A. A.; Dodson, E. J. *Acta Crystallogr.* **1997**, *D53*, 240−255.
(23) Pettersen, E. F.; Goddard, T. D.; Huang, C. C.; Couch, G. S.; Greenblatt, D. M.; Meng, E. C.; Ferrin, T. E. *J. Comput. Chem.* **2004**, *25*, 1605−1612.
(24) DeLano, W. L. *The PyMOL Molecular Graphics System*; DeLano Scientific LLC: Palo Alto, CA, 2008; http://www.pymol.org
(25) Schultes, C. M.; Guyen, B.; Cuesta, J.; Neidle, S. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 4347−4351.