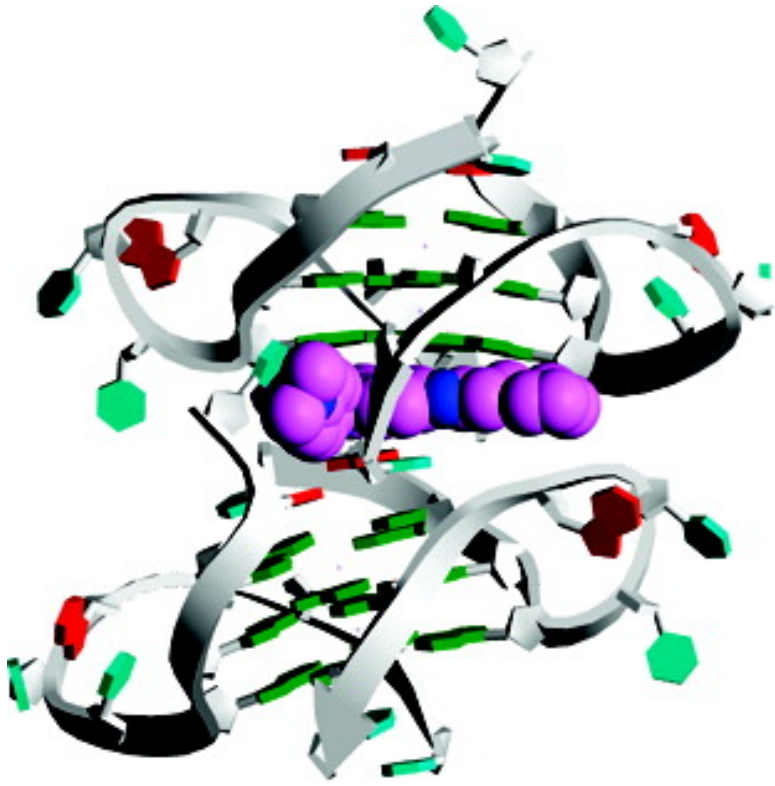
Communication

Structural Basis of DNA Quadruplex Recognition by an Acridine Drug

Nancy H. Campbell, Gary N. Parkinson, Anthony P. Reszka, and Stephen Neidle

J. Am. Chem. Soc., **2008**, 130 (21), 6722-6724 • DOI: 10.1021/ja8016973 • Publication Date (Web): 06 May 2008



J|A|C|S
COMMUNICATIONS

Published on Web 05/06/2008

Structural Basis of DNA Quadruplex Recognition by an Acridine Drug

Nancy H. Campbell, Gary N. Parkinson, Anthony P. Reszka, and Stephen Neidle*

The Cancer Research UK Biomolecular Structure Group, The School of Pharmacy, University of London, 29−39 Brunswick Square, London WC1N 1AX, United Kingdom

Received March 6, 2008; E-mail: stephen.neidle@pharmacy.ac.uk

Cancer cells have evolved mechanisms to maintain telomere length, in most instances by the synthesis of further telomeric DNA repeats using the telomerase enzyme.[1] The consequence of indefinite telomere length maintenance is cellular immortalization, a key step along the pathway to tumorigenesis. The telomerase complex, which maintains telomere length in some 80−85% of cancer cells[2] by virtue of its reverse transcriptase activity, is thus a potential target for therapeutic intervention.[3] The 3′-end of human telomeres, which comprises repeats of the sequence d(TTAGGG), is single-stranded for the final 100−200 nucleotides of its length, and can be induced to form both intra- and intermolecular four-stranded G-quadruplex structures[4] by quadruplex-binding ligands. These can inhibit telomerase activity since quadruplex DNA is not recognized by the single-stranded template of the RNA component of the telomerase complex.[5] Many such ligands have been reported,[6] and empirical SARs have been defined for a number of them, although structural information to date on quadruplex-ligand complexes is limited. The biological consequences of telomere maintenance disruption have been reported for a few ligands and include telomere uncapping and consequent induction of a DNA damage response.[7]

The 3,6,9-trisubstituted acridine ligand BRACO-19 (Figure 1) was designed by qualitative molecular modeling, on the assumption that the three substituents would each occupy a groove in a quadruplex.[8a] BRACO-19 inhibits telomerase enzymatic activity, resulting in telomere shortening, and also produces end-to-end chromosomal fusions in cancer cells[8b] as a consequence of quadruplex disruption of uncapping of proteins associated with the single-strand overhang. It shows significant *in vivo* anticancer activity in tumor xenografts, which is associated with telomere uncapping.[9]

We report here the 2.5 Å crystal structure (R factor = 18.3% and R_{free} = 21.3%) of a complex between a bimolecular human telomeric G-quadruplex of sequence d(TAGGGTTAGGGT) and BRACO-19. The crystals grew in the presence of physiological levels of K^+ ions and revealed a parallel-stranded quadruplex arrangement, with the biological unit being two 5′ to 3′ stacked quadruplexes and a well-resolved BRACO19 molecule (Figure 2); the unit packs into infinite columns in the crystal. This topology has been found in the crystal structures of native unimolecular and bimolecular human telomeric quadruplexes,[10] in a complex with the porphyrin TMPyP4,[11] indicating that it is a preferred topology for quadruplex-ligand complexes. Antiparallel structures have been assigned by solution NMR studies[12,13] for human telomeric DNA quadruplexes, with one parallel and two lateral loops. These may not be compatible with drug binding since a G-tetrad platform has to be unrestricted by lateral or diagonal loops for effective drug stacking onto the platform.

Each bimolecular quadruplex in the present structure contains three planar stacked G-tetrads with a BRACO-19 molecule stacking directly onto the 3′ end G-tetrad face. The TTA sequences, which intersperse the G-tracts, form propeller loops that are similar to

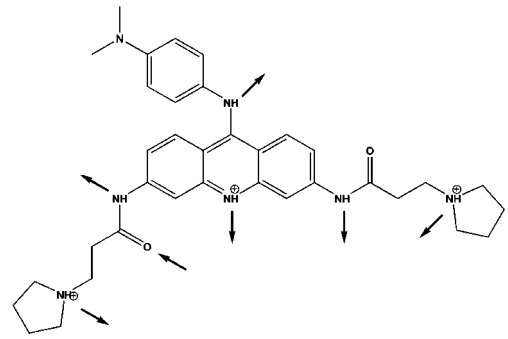


Figure 1. Structure of BRACO-19. Arrows indicate directionality of hydrogen bonding found in the crystal structure reported here.

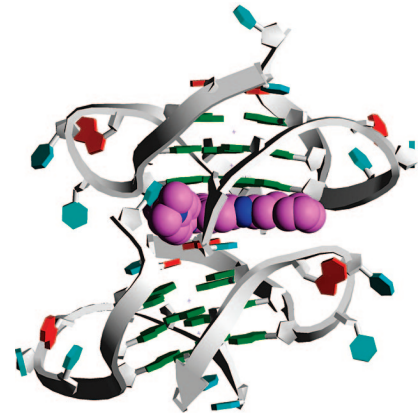


Figure 2. The biological unit in the crystal (PDB id 3CE5). A BRACO-19 molecule (mauve) is shown at the interface of the two quadruplexes in the unit, stacked between a G-quartet (top) and a TATA tetrad (bottom).

those observed in the native crystal structures.[10] BRACO-19 fits into a site bounded on one side by the 3′-end G-tetrad and on the other by a 5′-end TATA tetrad face formed from adenine and thymine bases at the interface between the two 5′ to 3′ stacked bimolecular quadruplexes in the biological unit. The drug is asymmetrically stacked on the G-tetrad at the end of one quadruplex, and there is $\pi-\pi$ overlap with just two guanine bases (Figure 4). The cationic ring nitrogen atom of the acridine ring is in-line with the K^+ ion channel that runs through the quadruplex. (Ions were unequivocally assigned as K^+ in the crystallographic analysis). The other side of the acridine chromophore surface is stacked not onto a second G-tetrad but onto a reverse Watson−Crick A·T base pair that forms part of a TATA quartet formed at the dimer interface. This TATA platform is constructed from the 5′-TA ends of two strands of one quadruplex, and the 3′-thymine from a third strand on the other. The G- and the TATA tetrads are off-set with respect to each other, and inclined by ca. 30° in two directions (Figure 4). The other 3′ end thymine base is flipped into the binding site so

10.1021/ja8016973 CCC: $40.75

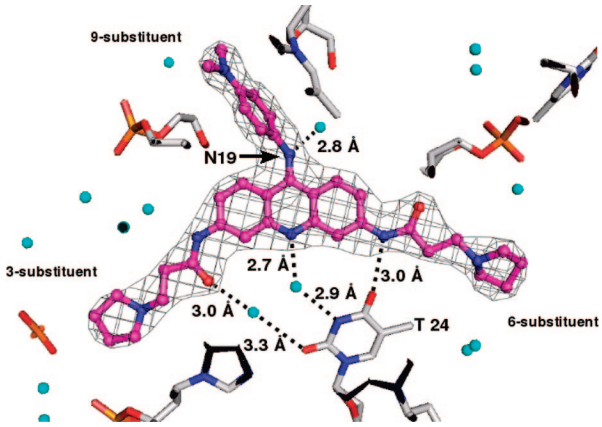


Figure 3. View onto the face of the BRACO-19 molecule in an omit map, showing the hydrogen bonds to water molecules and the flipped-in thymine base T24, with electron density for the drug contoured at 1.0 σ.

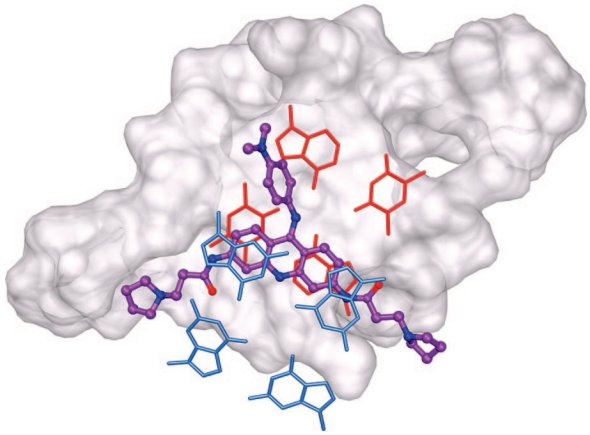


Figure 4. View on the face of the BRACO-19 molecule (mauve), showing the G-tetrad (in blue) above the plane and the TATA quartet (in red) stacked below the drug. The solvent-accessible surface of the lower quadruplex is also shown. This and other figures have been produced by the UCSF CHIMERA package (*J. Comput. Chem.* **2004**, *25*, 1605−1612).

that it is able to interact with the BRACO-19 molecule (Figure 3), an arrangement reminiscent of the *Oxytricha nova* antiparallel bimolecular quadruplex crystal structure complexed with a disubstituted acridine ligand.[14] This thymine thus plays a significant role in ligand binding, with a direct H-bond to a side-chain amide nitrogen atom, and two water-mediated interactions, one with the carbonyl atom of the other side-chain and the other to the central ring nitrogen atom of the acridine (Figure 3).

The 3- and 6-position substituents of BRACO19, comprising flexible side chains with cationic termini each extending into a wide groove, are located on opposite sides of the G-tetrad face, while the 9-position anilino substituent fits into a narrow pocket. These wide shallow grooves, whose walls are formed by the negatively charged phosphate backbones and floors by the edges of the G-tetrads, are filled with water molecules that contact phosphate oxygen atoms and base edges (Figure 5a,b). There are no direct H- bonds from the cationic pyrrolidino ring nitrogen atoms at the ends of the side chains to any of the negatively charged phosphate backbones. Instead, both nitrogen atoms participate in the networks of water molecules in the grooves, and these make contact with phosphate groups (Figure 6). Water molecules also contact the amide groups of the side chains, on the opposite side of the H-bonded thymine base. The phenyl ring of the 9-substituent fits tightly into a narrow hydrophobic pocket at the dimer interface,

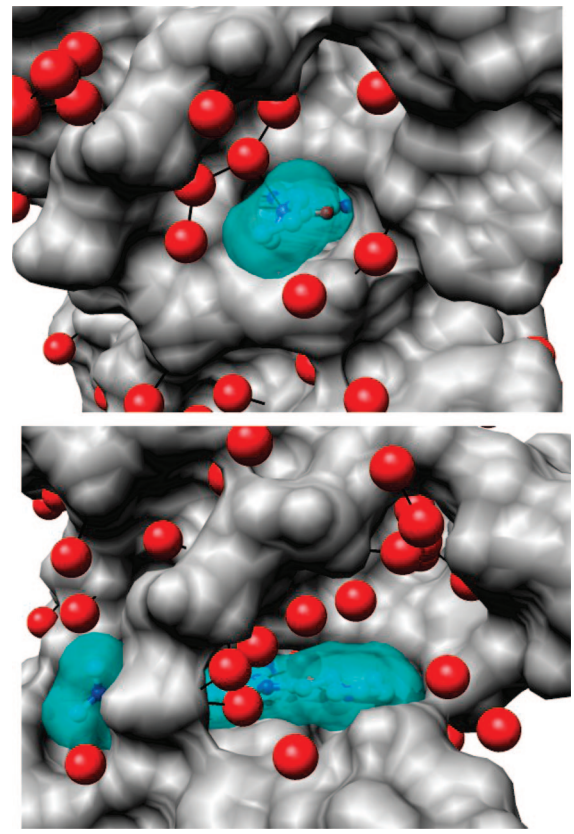


Figure 5. (a,b) Views of the biological unit. The BRACO-19 molecule is shown in a transparent cyan-colored solvent-accessible surface representation with individual atoms shown in ball-and stick form. The solvent-accessible surface of the quadruplex is colored grey. The figures each show one of the two 3,6-substituents protruding into the wide shallow grooves and the arrangement of waters (red spheres) in the grooves. The pyrrolidino nitrogen atoms of each side chain are each H-bonded to a water molecule, which are within water networks. H bonds are shown by black lines.

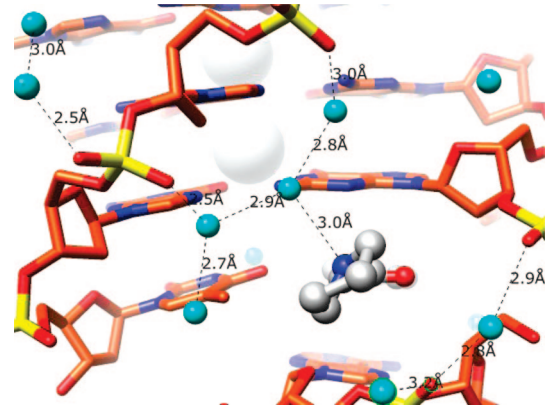


Figure 6. Detailed view of the hydrogen-bonding arrangement involving the water molecule network and the pyrrolidino nitrogen atom at the end of one of the BRACO-19 side chains. This nitrogen is directly hydrogen-bonded to one water molecule, which in turn hydrogen bonds to two others, each of which contacts a phosphate group.

which is formed by a terminal 5′ thymine base from one quadruplex in the biological dimer and the C5′ from the other quadruplex (Figure 7). This pocket is extended into a narrow groove that is filled with water molecules, reminiscent of a DNA duplex minor groove. There is also a H-bond between a water molecule and the nitrogen atom N19 bridging between the acridine core and the phenyl group, although other water molecules contacting this one

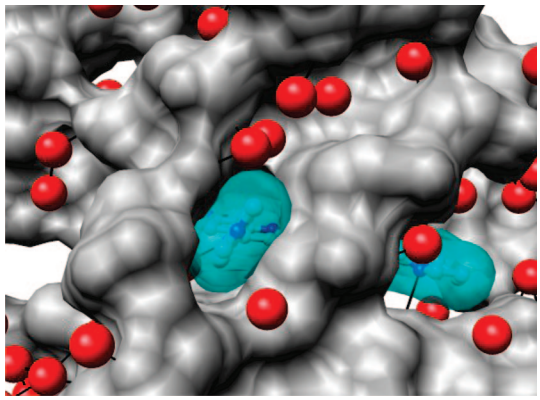


Figure 7. View of the 9-position aniline group (in cyan) tightly bound in one end of a narrow channel, with the upper part of the channel occupied by a number of water molecules. Colors are as in Figure 5.

are not well resolved. Thus of the eight donor−acceptor substituents in BRACO-19, seven participate in H-bonding, of which six are to water molecules rather than directly to quadruplex substituents.

This crystal structure provides a rationale for the quadruplex binding and biological behavior of a number of BRACO-19 analogues. Extending the length of the 3- and 6- side chains by $-(CH_2)-$ groups results in a progressive decrease in quadruplex binding; the binding constant for BRACO-19 itself is 3.1×10^7 M^{-1}, whereas it is 6×10^5 M^{-1} for the analogue with three more $-(CH_2)-$ units in each side chain.[15] The structure shows that the increasingly hydrophobic side chains when extended would no longer be able to participate in a water H-bond network in the groove since the charged end-group would extend too far out of the groove region. The alternative, of the hydrophobic chains packing into the grooves, is also unlikely given the wide yet shallow nature of these grooves. Increasing the size of the end-group from the five-membered pyrrolidino ring to a six-membered piperidino or similar ring does not have a deleterious effect on binding. The crystal structure also shows that the side chains are both in narrow regions at the entrance to the grooves, which would exclude the binding of larger ring systems, in agreement with experiment.[15−17]

The crystal structure explains the key role of the 9-position aniline substituent, which increases quadruplex affinity 10-fold. The narrow cross-section of the binding pocket accommodates a planar aromatic ring, but would also fit other hydrophobic groups, such as a linear $-CH_2)_n-$ chain, in accord with experiment. A substituent attached to one of the anilino amine nitrogen atoms can confer additional stability.[15−17] For example the analogue with a $-(C{=}O)-(CH_2)_3-$pyrrolidino has a binding constant of 7.7×10^7 M^{-1}. Qualitative modeling of this substituent into the crystal structure shows that it lies snugly in the groove and suggests that substituents longer than this would not fit.

The development of many G-quadruplex ligands has to date assumed that stacking to a G-tetrad is an adequate description of the binding. This crystal structure shows that this is not the case, with a structurally complex ligand such as BRACO-19 inducing the formation of a quadruplex binding site with a high degree of three-dimensional complexity. Although the details of the many stabilizing interactions involving water molecules are probably specific to acridines and BRACO-19 analogues, the active involvement of loop and other nucleotides in helping to define and shape the binding site are likely to be general features of ligand binding to telomeric and other quadruplexes. This has been observed in our recently determined structures of two quadruplex complexes with a naphthalene diimide ligand (PBD nos. 3CDM, 3CC0), which have the TTA loops altering their conformation to form part of the binding site platforms for this molecule. The BRACO-19 complex, with its all-parallel topology and 3′ to 5′ packing of repeating units in the crystal structure, is a direct model for the drug-bound single-strand overhang and suggests that the drug may promote the formation of a highly compacted secondary structural arrangement that would effectively compete with telomerase uncapping and with other single-strand binding proteins such as hPOT1.

Acknowledgment. This work is supported by Cancer Research UK (Programme Grant No. C129/A4489 to S.N.).

Supporting Information Available: Crystallographic detail. This material is available free of charge via the Internet at http://pubs.acs.org. Coordinates and structure factors are available from the PDB as entry no. 3CE5.

References

(1) Xu, L.; Blackburn, E. H. *Mol. Cell* **2007**, *28*, 315–327.
(2) Kim, N. W.; Piatyszek, M. A.; Prowse, K. R.; Harley, C. B.; West, M. D.; Ho, P. L. C.; Coviello, G. M.; Wright, W. E.; Weinrich, R.; Shay, J. W. *Science* **1994**, *266*, 2011–2015.
(3) (a) Shay, J. W.; Wright, W. E. *Nat. Rev. Drug Discovery* **2006**, *5*, 577. (b) Oganesian, L.; Bryan, T. M. *Bioessays* **2007**, *29*, 155–165.
(4) Burge, S.; Parkinson, G. N.; Hazel, P.; Todd, A. K.; Neidle, S. *Nucleic Acids Res.* **2006**, *34*, 5402–5415.
(5) (a) Zahler, A. M.; Williamson, J. R.; Cech, T. R.; Prescott, D. M. *Nature* **1991**, *350*, 718–720. (b) 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.
(6) For a recent survey see: De Cian, A.; Lacroix, L.; Douarre, C.; Temime-Smaali, N.; Trentesaux, C.; Riou, J.-F.; Mergny, J.-L. *Biochimie* **2008**, *90*, 131–155.
(7) For example: (a) Gunaratnam, M.; Greciano, O.; Martins, C.; Reszka, A. P.; Schultes, C. M.; Morjani, H.; Riou, J.-F.; Neidle, S. *Biochem. Pharmacol.* **2007**, *74*, 679–689. (b) Brassart, B.; Gomez, D.; De Cian, A.; Paterski, R.; Montagnac, A.; Qui, K. H.; Temime-Smaali, N.; Trentesaux, C.; Mergny, J.-L.; Gueritte, F.; Riou, J.-F. *Mol. Pharmacol.* **2007**, *72*, 631–640.
(8) (a) Read, M.; Harrison, R. J; Romagnoli, B.; Tanious, F. A.; Gowan, S. H.; Reszka, A. P.; Wilson, W. D.; Kelland, L. R.; Neidle, S. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 4844–4849. (b) Incles, C. M.; Schultes, C. M.; Kempski, H.; Koehler, H.; Kelland, L. R.; Neidle, S *Mol. Cancer Ther.* **2004**, *3*, 1201–1206.
(9) 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.
(10) Parkinson, G. N.; Lee, M. P. H.; Neidle, S. *Nature* **2002**, *417*, 876–880.
(11) Parkinson, G. N.; Ghosh, R.; Neidle, S. *Biochemistry* **2007**, *46*, 2390–2397.
(12) Phan, A. T.; Kuryavyi, V.; Luu, K. N.; Patel, D. J. *Nucleic Acids Res.* **2007**, *35*, 6517–6525.
(13) (a) Phan, A. T.; Modi, Y. S.; Patel, D. J. *J. Am. Chem. Soc.* **2004**, *126*, 8710–8716. (b) Ambrus, A.; Chen, D.; Dai, J.; Jones, R. A.; Yang, D. *Biochemistry* **2005**, *44*, 2048–2058. (c) Ambrus, A.; Chen, D.; Dai, J.; Bialis, T.; Jones, R. A.; Yang, D. *Nucleic Acids Res.* **2006**, *34*, 2723–2735.
(14) Haider, S. M.; Parkinson, G. N.; Neidle, S. *J. Mol. Biol.* **2003**, *326*, 117–125.
(15) Moore, M. J.; Schultes, C. M.; Cuesta, J.; Cuenca, F.; Gunaratnam, M.; Tanious, F. A.; Wilson, W. D.; Neidle, S. *J. Med. Chem.* **2006**, *49*, 582–599.
(16) Schultes, C. M.; Guyen, B.; Cuesta, J.; Neidle, S. *Bioorg. Med. Chem. Lett.* **2004**, *14*, 4347–4351.
(17) Harrison, R. J.; Cuesta, J.; Chessari, G.; Read, M. A.; Basra, S. K.; Reszka, A. P.; Morrell, J.; Gowan, S. M.; Incles, C. M.; Tanious, F. A.; Wilson, W. D.; Kelland, L. R.; Neidle, S. *J. Med. Chem.* **2003**, *46*, 4463–4476.

JA8016973