ChemComm

Dynamic Article Links

Cite this: *Chem. Commun.*, 2012, **48**, 8958–8960

www.rsc.org/chemcomm

COMMUNICATION

The role of thermodynamics and kinetics in ligand binding to G-quadruplex DNA†

Ben Corry*[a] and Nicole M. Smith[b]

Received 28th March 2012, Accepted 18th July 2012
DOI: 10.1039/c2cc32240d

Molecular dynamics simulations were used to investigate the binding of four different 2,4,6-triarylpyridines to G-quadruplex DNA. Both the binding free energies, and the kinetics of binding are required to explain the measured degree of ligand induced stabilisation of the compounds, with bulky substituents having the potential to prevent the ligand from reaching the lowest energy binding site.

G-quadruplex DNA (G4-DNA) is a highly dynamic and polymorphic four stranded DNA structure that can form from certain guanine rich (G-rich) sequences, such as those occurring at the 3′-terminus of human telomeric DNA (HTelo) and in the promoter region of certain oncogenes.[1–3] Small molecules that bind and stabilise G4-DNA have been shown to down regulate oncogene expression and induce telomere disruption, hence G4-DNA is an attractive target for selective anti-cancer therapy.[3–7]

We previously reported the G4-DNA stabilisation ability of a series of novel 2,4,6-triarylpyridines with three quadruplex-forming oligonucleotides that mimic: (i) the G-rich HTelo overhang, (ii) a G-rich sequence in the c-kit promoter, and (iii) a G-rich sequence in the K-ras promoter.[8] In terms of the stabilisation ability of these ligands, we found that there exists three categories within this series of compounds: (1) compounds which exhibit high stabilisation of G4-DNA that are comparable or superior in magnitude to the best G4-DNA ligands; (2) compounds which show mid-range stabilisation; and (3) compounds which exhibit low stabilisation of G4-DNA.

Herein, we conducted molecular dynamics simulations to further investigate the molecular basis of binding of these compounds to G4-DNA and the reasons for their different ability to stabilise the quadruplex. Furthermore, we report the first use of metadynamics simulations to examine the kinetics of binding of these triarylpyridines to G4-DNA. Structural information derived from the modelling of G4-DNA–ligand interactions can lead to new and more effective G4-DNA stabilisers.

Simulations focused on the interactions of 4 compounds from our previously reported 2,4,6-triarylpyridine series with the NMR solution structure of the HTelo Repeat, d($AG_3[T_2AG_3]_3$) in Na^+ conditions (PDB 143D).[8,9] The non-polymorphic nature of HTelo in Na^+ conditions and the availability of a high resolution structure makes it an ideal candidate for molecular modelling studies. Results obtained from the simulations were compared to the ligand induced change in melting temperature (ΔT_m) for the F21T G4-DNA sequence, d($G_3[T_2AG_3]_3$), in Na^+ conditions previously measured by us using Förster Resonance Energy Transfer (FRET).[8] F21T is a commonly used sequence to mimic HTelo, and differs by only one nucleotide. The compounds studied were: compound **1** which exhibited high stabilisation of F21T G4-DNA in 100 mM Na^+ buffer (pH 7.2) with a ΔT_m of 31.6 ± 1.8 °C; compounds **2** and **3** which exhibited mid-range stabilisations with ΔT_m values of 10.2 ± 0.2 °C and 10.6 ± 0.0 °C respectively; and compound **4** which exhibited low stabilisation with a ΔT_m of 5.7 ± 0.5 °C (Fig. 1).[8]

As for compounds with acridine based central scaffolds,[10–12] the compounds in the present study bind strongly between the upper G-quartet and the connecting loop, with up to 2 additional adenine nucleotides stacking above them (Fig. 2), although other binding modes cannot be excluded. In this position there is a very large contact area between the extended aromatic regions of the compounds and the guanine and adenine nucleotides that can be expected to assist in stabilising the quadruplex.

The ability of the compounds to stabilise the quadruplex can be influenced by both how strongly they bind as well as

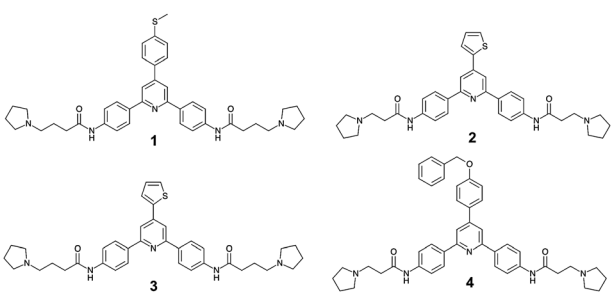

Fig. 1 Schematic representation of compounds studied.

[a] *Research School of Biology, Bldg 134, The Australian National University, Canberra ACT 0200, Australia. E-mail: ben.corry@anu.edu.au; Fax: +61 2 6125 0313; Tel: +61 2 6125 0842*
[b] *Research School of Chemistry, Bldg 35, The Australian National University, Canberra ACT 0200, Australia*

† Electronic supplementary information (ESI) available: Including computational details, a breakdown of the DNA–ligand interaction energies and movies of ligand unbinding. See DOI: 10.1039/c2cc32240d

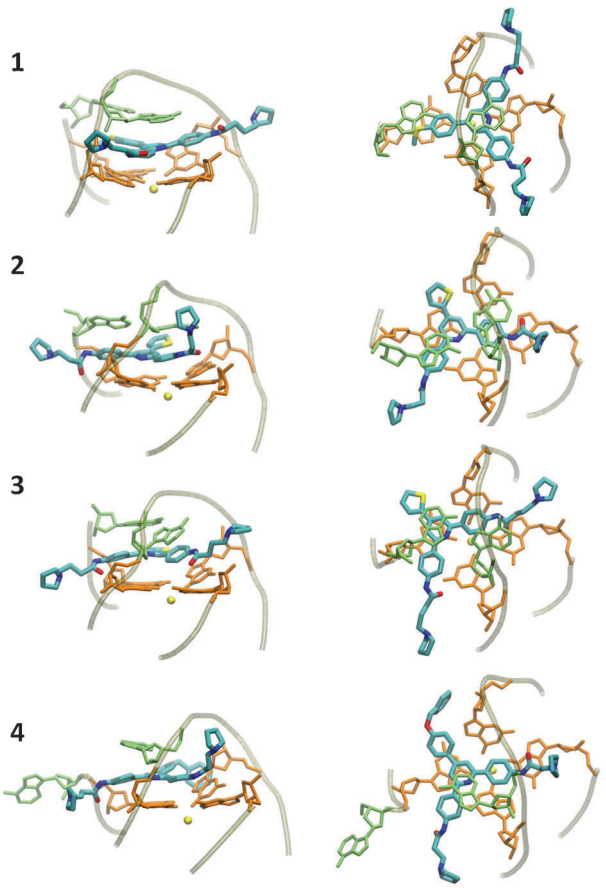

Fig. 2 Location of ligand binding between the upper G-quartet and the connecting loop, viewed at right angles and normal to the plane of the G-quartet for each compound. Snapshots from molecular dynamics simulations are shown with guanine coloured orange, adenine green, the DNA backbone transparent and Na^+ yellow.

how easily they can find this binding position, that is, by both thermodynamic and kinetic factors. Thermodynamic factors can be studied by examining the binding free energy (ΔG) of each compound which is reported in Table 1. Each compound has multiple similar binding modes associated with rotation around the central axis of the G-quartets, leading to the side chains of the ligand extending into different grooves. In general these different binding modes had similar binding energies, with differences less than about 4 kcal mol^{-1}, thus only the mode with the most negative binding energy is shown (Fig. 2 and Table 1). It is clear that **1** binds more strongly than **2** or **3** which in turn bind more strongly than our previously reported triarylpyridine compound lacking the charged side chains, which displayed a simulated ΔG value of −24.6 kcal mol^{-1}.[13]

Table 1 Binding free energies to HTelo, ΔG, for each compound studied (kcal mol^{-1})

Compound	Binding free energy ΔG
1	−42.7
2	−38.6
3	−33.6
4	−46.2

As can be seen in a breakdown of the ligand–DNA interaction energies (Table S1, ESI†), the inclusion of charged side chains notably increases the electrostatic attraction of the compounds to the DNA, with the side chains pointing either up or down to contact the negatively charged DNA backbone of either the loop (above) or strand (below). Although compound **1** contains the same side chains as compound **3**, the electrostatic interactions of the side chains with the DNA are very different in each case. This suggests that the 4-aryl substituent on the central pyridine ring is an important factor in determining the overall positioning of the ligand within the quadruplex. The larger *p*-thiomethylphenyl substituent in **1** allows for greater interaction with the DNA backbone compared to the thiophene substituent of **2** and **3**. This ability of **1** to form strong interactions with three different loops of the DNA (compared to only two for **2** and **3**), along with its more negative binding free energy, may account for its much greater ability to stabilise F21T G4-DNA seen experimentally.[8] Compounds **2** and **3** differ only in their amidoalkylpyrrolidine side chain length, propyl (C_3) for **2** *vs.* butyl (C_4) for **3**. Increasing the length of the side chains, as in **3**, places the positive charges further from the backbone, reducing the electrostatic interaction and consequently the magnitude of the binding energy (Table 1 and Table S1, ESI†). However, despite the difference in binding energy seen in the simulations, compounds **2** and **3** did not differ in their stabilisation ability for the F21T G4-DNA in the melting experiments.[8]

Most interestingly, compound **4** has a very high binding energy despite the fact that it shows low experimental stabilisation of F21T G4-DNA compared to compounds **1**, **2** and **3**.[8] As seen in Fig. 2, the bulky 4-aryl substituent on the central pyridine ring does not appear to impede binding, the only difference being the lack of interactions with one of the adenine nucleotides. If the difference in stabilisation for **4** cannot be explained by the thermodynamics of binding, it must arise from the ability of the compounds to find the binding site.

To examine the kinetics of binding, metadynamics simulations were conducted to allow the unbinding of each of the four compounds to be simulated. As can be seen in Fig. 3A, the unbinding of **2** takes place in a smooth fashion, with the compound sliding out of the quadruplex before binding on the outside of the molecule with only limited energy barriers on the way. A notable aspect of this process is that the small thiophene substituent at the 4-position of the central pyridine ring is able to swing inside the DNA backbone such that it can occupy the same pocket of the DNA as one of the side chains. From this position, the compound can leave the DNA with little impediment. The energy barrier associated with unbinding of **2** (8 kcal mol^{-1}) is similar to that found for compounds **1** and **3** (10 & 11 kcal mol^{-1} respectively).

In contrast, Fig. 3B shows that **4** has great difficulty in leaving the binding site, with a barrier for unbinding of greater than 20 kcal mol^{-1}. As a consequence, much less of the free energy surface is sampled during the 35 ns. As the benzyl phenyl ether substituent at the 4-position of the central pyridine ring in compound **4** is larger than the thiophene substituent in compound **2**, it cannot swing past the DNA backbone. Thus, the compound must leave by moving over

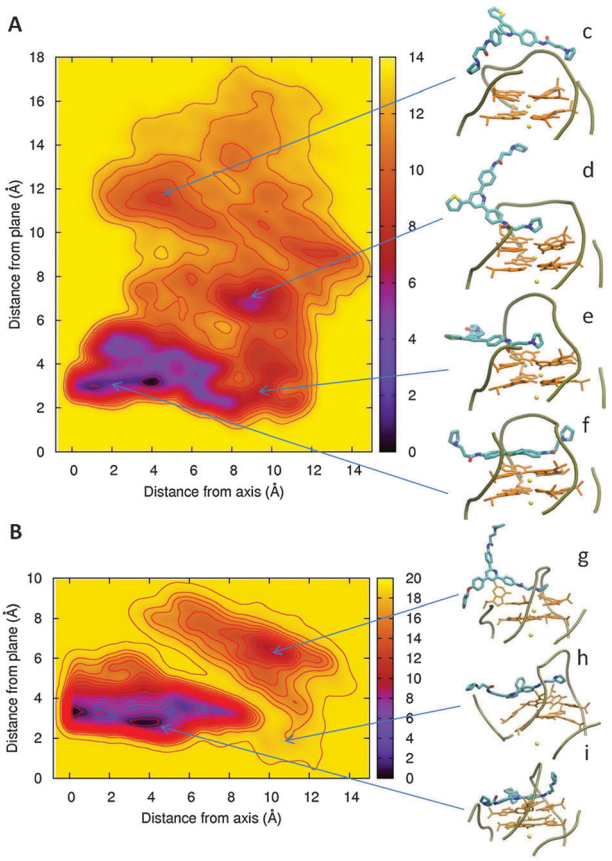

Fig. 3 Free energy surface and transient structures for the unbinding of (A) **2** and (B) **4** from HTelo G4-DNA from metadynamics simulations with free energy plotted as a function of the position of the central pyridine of the compound relative to the distance from the axis passing through the G-quartets and the distance above the plane of the top G-quartet. Contour intervals of 1 kcal mol^{-1} are shown. Colours are as for Fig. 2. Local minimum energy c, d and f and transition state e structures for **2** are shown along with their location on the free energy surface. Local minimum energy g and i and transition state h structures for **4** are shown along with their location on the free energy surface. For **4**, the high energy of the transition state structure h results from the upper G-quartet dissociating to allow the compound to exit the binding site.

one of the end strands. This requires the top G-quartet to break apart, which comes at a large energy cost as seen in the large ridge in the free energy surface between the two minima that represent the compound stacking on the G tetrad and binding on the outside of the G4-DNA. As the same barriers that prevent **4** from leaving will also hinder entry of the compound into the binding site, the simulations suggest that this compound can only enter the binding site by breaking the G-quartet. For this reason it is kinetically more favourable for this compound to bind on the outside of the DNA where it has only a limited stabilising influence on the quadruplex.

The metadynamics simulations also suggest possible binding pathways for **2** to HTelo G4-DNA, as highlighted in Fig. 3. Initial binding is likely to occur on the outside of the quadruplex with the two side chains clamping over the top or side of the DNA (c). From here, one of the side chains can swing into the centre of the quadruplex above the top G-quartet (d). Finally, the rest of the molecule can follow the chain through the transition state (e) into the binding site and stack above the guanine nucleotides (f). Movies showing the unbinding of **2** and **4** are included in the ESI.†

We have predicted through molecular dynamics simulations that the presence of a large 4-aryl substituent on the central pyridine ring of the 2,4,6-triarylpyridines, as in compounds **1** and **4**, allows for stronger interactions of these ligands with G4-DNA. This is not only due to the greater interaction of the larger 4-aryl substituents with the DNA backbone, but also due to the ability of these larger 4-aryl substituents to better position these ligands within the G4-DNA resulting in greater interactions with the charged side chains and larger binding energies. However, a more negative binding free energy (ΔG), does not necessarily imply greater stabilisation (ΔT_m) as seen for compound **4**. Using metadynamics simulations to model the unbinding of these compounds from G4-DNA, we suggest that the ability of the compounds to stabilise G4-DNA also depends on the kinetics of binding, *i.e.* the ability of the ligand to enter the G4-DNA binding site. Hence, in the case of compound **4**, large 4-aryl substituents impede the ability of the compound to enter the G4-DNA binding site, resulting in the compound binding weakly on the outside of the DNA and where it provides less stabilisation. In order to design more potent G4-DNA ligands which demonstrate both tight binding and high stabilisation, it is crucial that we consider how substituents on the ligands influence both the thermodynamics and kinetics of binding.

The authors graciously acknowledge support of this work by the Australian Research Council (ARC). This work was also supported by an award under the Merit Allocation Scheme of the National Computing Infrastructure at the ANU and additional computer time from iVEC.

Notes and references

1 S. Burge, G. N. Parkinson, P. Hazel, A. K. Todd and S. Neidle, *Nucleic Acids Res.*, 2006, **34**, 5402–5415.
2 Y. Qin and L. H. Hurley, *Biochimie*, 2008, **90**, 1149–1171.
3 A. De Cian, L. Lacroix, C. Douarre, N. Temime-Smaali, C. Trentesaux, J.-F. Riou and J.-L. Mergny, *Biochimie*, 2008, **90**, 131–155.
4 A. Siddiqui-Jain, C. L. Grand, D. J. Bearss and L. H. Hurley, *Proc. Natl. Acad. Sci. U. S. A.*, 2002, **99**, 11593–11598.
5 A. De Cian, G. Cristofari, P. Reichenbach, E. De Lemos, D. Monchaud, M.-P. Teulade-Fichou, K. Shin-ya, L. Lacroix, J. Lingner and J.-L. Mergny, *Proc. Natl. Acad. Sci. U. S. A.*, 2007, **104**, 17347–17352.
6 D. Gomez, *et al.*, *Cancer Res.*, 2006, **66**, 6908–6912.
7 T.-M. Ou, *et al.*, *J. Med. Chem.*, 2007, **50**, 1465–1474.
8 N. M. Smith, G. Labrunie, B. Corry, P. L. T. Tran, M. Norret, M. Djavaheri-Mergny, C. L. Raston and J.-L. Mergny, *Org. Biomol. Chem.*, 2011, **9**, 6154.
9 Y. Wang and D. Patel, *Structure*, 1993, **1**, 263–282.
10 J. E. Redman, J. M. Granadino-Roldán, J. A. Schouten, S. Ladame, A. P. Reszka, S. Neidle and S. Balasubramanian, *Org. Biomol. Chem.*, 2009, **7**, 76–84.
11 N. H. Campbell, M. Patel, A. B. Tofa, R. Ghosh, G. N. Parkinson and S. Neidle, *Biochemistry*, 2009, **48**, 1675–1680.
12 N. H. Campbell, G. N. Parkinson, A. P. Reszka and S. Neidle, *J. Am. Chem. Soc.*, 2008, **130**, 6722–6724.
13 N. M. Smith, B. Corry, K. S. Iyer, M. Norret and C. L. Raston, *Lab Chip*, 2009, **9**, 2021–2025.