Structure-based design of selective high-affinity telomeric quadruplex-binding ligands†

Caterina Maria Lombardo,*,a Iria Sánchez Martínez,a Shozeb Haider,a Valérie Gabelica,b Edwin De Pauw,b John E. Mosesc and Stephen Neidle*a

Received 30th July 2010, Accepted 6th October 2010
DOI: 10.1039/c0cc02917c

A library of triazole-based telomeric quadruplex-selective ligands has been developed that mimic an established family of tri-substituted acridine-based ligands, using crystal structure data as a starting-point for computer-based design. Binding affinities, estimated by electrospray mass spectrometry, are in accord with the design concept.

Oligonucleotides and nucleic acids containing G-tracts can be organised as G-quadruplexes.1 These are polymorphic tertiary structures, characterised by a hydrophobic core of G-quartets and negatively-charged loops. Quadruplexes show exceptional stability over other conformations in the presence of Na+ or K+ ions. Putative quadruplex sequences have been identified in G-rich genomic sequences,2 with over-representation in telomeres,3 as well as in other genomic regions for example in promoter sequences of a number of proto-oncogenes,4 such as c-myc5a and c-kit5b in 5′ untranslated regions5a and in introns.6a A number of these putative quadruplexes are appealing targets for cancer therapeutics. For instance, inducing the single-stranded telomeric DNA overhang to fold into G-quadruplexes has been shown to inhibit telomererase activity7a and cancer cell growth.7b Such precise targeting of human telomeres is significant since in >80% of cancers telomerase is up-regulated and contributes to the malignant phenotype by maintaining cancer cell immortalization.7c

A considerable number of small organic molecules have been found to stabilise quadruplex DNA structures.8 Many, though not all, are based on polycyclic heteroaromatic cores, with the acridine nucleus being especially well explored.9 However, the selectivity of many of these molecules for G-quadruplexes over duplex DNA is frequently less than what would be therapeutically acceptable, and their polycyclic features can make their druggability a challenge.

The 3,6,9-trisubstituted acridine ligand BRACO-19 (Fig. 1) has high affinity for human telomeric quadruplex DNAs and is a potent inhibitor of the telomererase enzyme.10 It has selective cytotoxic activity against a range of human tumour cell lines and shows antitumour activity against xenograft models.11

The BRACO-19 molecule was designed using qualitative molecular modeling, with the crystal structure of the native parallel human telomeric quadruplex as a template. It was rationalized that each of the three substituents emanating from the acridine core of BRACO-19 would be able to interact with a quadruplex groove.10 This feature would, it was suggested, provide binding selectivity over duplex DNA, which has just two grooves. A more recent crystal structure of a BRACO-19 complex with a bimolecular quadruplex has confirmed the essential correctness of this hypothesis and has also provided a more detailed view of the interactions involved.12 The structure has a parallel-stranded quadruplex arrangement, with the biological unit being two 3′ to 3′ stacked quadruplexes. Each bimolecular quadruplex in this structure contains three planar stacked G-quartets with a BRACO-19 molecule stacking directly onto the 3′ end G-quartet face.

We report here the structure-based design, synthesis and preliminary assessment of a novel series of non-polycyclic trisubstituted ligands, whose affinity and selectivity for telomeric G-quadruplex DNA has been evaluated using electrospray mass spectrometry (ESI-MS).13 The goal has been to design molecules (i) with potentially enhanced selectivity based on the quadruplex concept of selective groove binding, and (ii) that do not have a polycyclic heteroaromatic core, so potentially enhancing drug-like features. We have used the BRACO-19 quadruplex complex crystal structure12 as a starting point for the structure-based design of non-polycyclic mimetics of BRACO-19. The single-stranded overhang of human telomeric DNA is 100–200 nucleotides in length, and in principle several quadruplex structures can be formed along its length. We have previously modeled such a higher-order arrangement using a simple linker between 5′ to 3′ ends in this crystal structure to form a continuous sequence, which does not involve any

† Electronic supplementary information (ESI) available: Details of synthetic chemistry, molecular modelling, electrospray mass spectrometry and results of ESI-MS competition experiments. See DOI: 10.1039/c0cc02917c
perturbation of the ligand binding site. The resulting 45-mer bis-quadruplex ligand sandwich complex modeled structure has been used here as the template for initial qualitative modelling followed by in silico docking and binding energy calculations for plausible ligands. We started with the hypothesis that such ligands would require three substituents as in BRACO-19 itself, and have explored their length and size. Ligand positions were explored computationally at the interface between the two quadruplexes, stacked on the 3' and 5' terminal G-quartet surfaces. Ligands were derived from a previously devised and extended with one-pot click reaction formation of three series of bis-triazole quadruplex-binding ligands which contain two alkylamine side-chain arms linked to a benzene core through 1,4-triazoles. These had been prepared using a click chemistry approach, which was also used here to add an additional triazole ‘side arm’ to the benzene core.

Initial qualitative modeling with the 45-mer suggested that addition of the third alkylamine side-chain (Fig. 1) could effectively mimic the key structural features of the BRACO-19 structure. This was followed by molecular dynamics simulations to ascertain whether the resulting complex was conformationally stable. The resulting structure (Fig. 1) indicates that the arrangement is structurally sound, with the central and three attached phenyl rings constituting the core of the ligand, all being involved in π-π stacking with G-quartet guanine bases in the binding site. The three alkylamine arms are each positioned in a groove structure, analogous to disposition of the side-chains in the BRACO-19 complex crystal structure. Circular dichroism studies on the binding of compound to a human telomeric quadruplex sequence show that the parallel form is induced (see the ESI†). Docking studies were also performed on an alternative polymorph of the human telomeric quadruplex, one of the (3 + 1) hybrid anti-parallel structures, but all low-energy arrangements involved stereochemically unacceptable buckling of ligand, DNA or both. These in silico trials were then abandoned.

The synthesis of a small library of the trisubstituted click ligands used a convergent click chemistry approach. A series of azide building blocks with focussed structural diversity was initially synthesised, then clicked onto a central core, the commercially available trialkyne 1,3,5-triethynylbenzene (Scheme 1). The diversity of building blocks was designed in order to establish structure–activity relationships. They all have the common features of a phenyl ring (providing an aromatic surface) and a series of basic side-chains. These were built onto a starting 2-nitroaniline or 3-nitroaniline in a maximum of four steps. Briefly, the required nitroaniline was acylated, then one-pot substituted with diverse amines. The nitro group was readily reduced with H2, then the azide was synthesised from the resulting aniline with one-pot diazotisation and azide substitution. The 1,3,5-triethynylbenzene was then extended with one-pot click reaction formation of three triazole rings via Cu(i) catalysed Huisgen 1,4-dipolar cycloaddition. Catalytic Cu(i) was formed in situ and a second catalyst, bathophenanthroline dithiocarbamoyl acid disodium salt hydrate, was necessary to achieve complete trisubstitution. The reaction was complete after only 15 minutes of microwave irradiation; an excess of the required amine in the reaction mixture was necessary to avoid elmination of the amine for side-chains with n ≥ 2. The resulting compounds all have an extended aromatic surface. Initial attempts to study DNA binding using techniques to measure the elevation in melting temperature as a result of ligand were unsuccessful since it appears that the compounds are unstable when UV irradiated at the elevated temperatures required in melting experiments. Their quadruplex and duplex DNA binding abilities were therefore assessed by electrospray mass spectrometry.

Binding of each ligand to the 22-mer human telomeric G-quadruplex sequence d[AGGG(TTAGGG)]3 (tel22) was examined at 1 : 1 and 2 : 1 drug : DNA ratios (Table 1, see also Fig. S1 and accompanying text, ESI†). All the Kd values reported refer to the 1 : 1 complex (Q + L). The meta-substituted compounds are consistently the strongest binders, with compounds 5, 8, 12 and 15 having lower Kd values than that of BRACO-19 itself. All of these have short side chains (n = 1, 2) with either pyrrolidine or diethylamine basic groups. These compounds have consistently higher affinities than the corresponding para compounds, with those having pyrrolidine and diethylamine end-groups being superior to the pyrrolidine derivatives and 13. Compounds 5 (n = 1) and 15 (n = 2) are more active than the related compound (n = 3), which suggests that shorter side chains result in superior quadruplex binding.

ESI-MS has also been used to assess quadruplex:duplex selectivity in a competition experiment between three sequences. The telomeric 22-mer (tel22), the parallel intermolecular G-quadruplex sequence (dTGG)4 and a duplex DNA, d(GCGGAATTCGCG), each at 5 μM, were injected with the ligand at 10 μM. Almost all the compounds showed selectivity for G-quadruplex sequences and did not bind to the duplex (Fig. S5, ESI†).

Several of the compounds have higher affinity for the tetramolecular TG4T quadruplex than for the telomeric quadruplex.
or for duplex DNA. By examining the ESI-MS spectra (see Fig. S1 and accompanying text, ESI†), we noted that, in addition to binding to tel22 to form 1 : 1 complexes, all ligands with para substituents were causing partial dimerization of the (dTG4T)4 quadruplex. However, compounds 5, 8, 12 and 15 have high affinity for the telomeric sequence, even in the presence of (dTG4T)4. The correlation with the \( K_d \) values is shown in Fig. S3 (ESI†). These are also the compounds that have the highest affinity for the telomeric quadruplex/duplex selectivity ratio > 1000 for ligand 1, for duplex binding of >3000 \( \mu \)M, for C.M.L.), the FNRS (research associate position to V.G.), the ESI-measured dissociation constants for compounds 5–15 with a human telomeric 22-mer quadruplex. NR2 groups are defined in Scheme 1.

Table 1

<table>
<thead>
<tr>
<th>Compound</th>
<th>Substitution pattern</th>
<th>( n )</th>
<th>NR2</th>
<th>( K_d/\mu M )</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>meta-</td>
<td>1</td>
<td>pyr</td>
<td>6.0 ± 0.6</td>
</tr>
<tr>
<td>6</td>
<td>para-</td>
<td>2</td>
<td>pyr</td>
<td>70 ± 50</td>
</tr>
<tr>
<td>7</td>
<td>meta-</td>
<td>3</td>
<td>pyr</td>
<td>35 ± 9</td>
</tr>
<tr>
<td>8</td>
<td>meta-</td>
<td>2</td>
<td>dieth</td>
<td>5.0 ± 1.0</td>
</tr>
<tr>
<td>9</td>
<td>para-</td>
<td>2</td>
<td>dieth</td>
<td>100 ± 19</td>
</tr>
<tr>
<td>10</td>
<td>meta-</td>
<td>1</td>
<td>pyr</td>
<td>49 ± 26</td>
</tr>
<tr>
<td>11</td>
<td>meta-</td>
<td>1</td>
<td>pip</td>
<td>250 ± 130</td>
</tr>
<tr>
<td>12</td>
<td>meta-</td>
<td>1</td>
<td>dieth</td>
<td>3.0 ± 0.6</td>
</tr>
<tr>
<td>13</td>
<td>para-</td>
<td>1</td>
<td>pip</td>
<td>77 ± 12</td>
</tr>
<tr>
<td>14</td>
<td>para-</td>
<td>1</td>
<td>dieth</td>
<td>32 ± 5</td>
</tr>
<tr>
<td>15</td>
<td>meta-</td>
<td>2</td>
<td>pyr</td>
<td>4.9 ± 1.3</td>
</tr>
<tr>
<td>BRACO-19</td>
<td>n/a</td>
<td>n/a</td>
<td>n/a</td>
<td>7.9 ± 1.4</td>
</tr>
</tbody>
</table>

Abbreviations: (pyr) pyrrolidino, (dieth) diethyl amino, (pip) piperidino. BRACO-19 has been used as a reference with data taken from ref. 20. Esds from two ligand concentrations (5 and 10 \( \mu \)M) and three voltage settings each.

Notes and references

15. S. M. Haider and S. Neidle, *Methods Mol. Biol.*, 2010, 608, 17. See also the ESI†.