J|A|C|S
COMMUNICATIONS

Published on Web 04/17/2003

# G-Quadruplex-Specific Peptide−Hemicyanine Ligands by Partial Combinatorial Selection

James A. Schouten, Sylvain Ladame, Stephen J. Mason, Matthew A. Cooper, and Shankar Balasubramanian*

*University Chemical Laboratory, University of Cambridge, Lensfield Road, Cambridge CB2 1EW, U.K.*

Received November 15, 2002; E-mail: sb10031@cam.ac.uk

The design of ligands that target nucleic acids with specificity is a major goal for chemists. Much of the work in this area has been focused on duplex B-DNA as a target. However, there are a number of nucleic acid targets based on alternative structures. The four-stranded G-quadruplex DNA motif has recently emerged as a biologically important structure.[1] The quadruplex has been linked to mechanisms that relate to a number of disease states,[2] most notably cancer via interfering with telomere maintenance[3] or the regulation of oncogenic promoters.[4] There is now considerable interest in the design of ligands that target G-quadruplex DNA and understanding the molecular basis for quadruplex recognition and discrimination. A number of promising small molecules have appeared in the literature,[5] with two examples of combinatorially selected artificial proteins.[6] A key challenge in the design of quadruplex ligands is to attain specificity for quadruplex over duplex DNA. A promising strategy for targeting nucleic acid or protein structures is to generate peptide−small molecule conjugates.[7] In particular, there is scope to generate ligands with superior recognition properties by combining binding elements that recognize distinct features of the target molecule.[8] In this paper, we describe the generation of novel heterocycle−peptide conjugates that bind telomeric G-quadruplex DNA with near micromolar affinity and high quadruplex/duplex discrimination.

To design quadruplex specific small molecule−peptide conjugates, we have employed a heterocyclic core with potential to bind DNA. A peptide appendage is subsequently added to generate additional interactions. Several carbocyanine dyes bind quadruplex DNA, including DODC (Figure 1A) that has been proposed to bind via a groove interaction.[9] A fragment (**1**) derived from DODC was employed as the heterocycle core. We have assumed that it retains some potential for quadruplex recognition. Candidate peptide fragments were selected from a naive tetrapeptide library (X−X−X−X where X = Arg, Lys, His, Glu, Asp, Asn, Gln, Pro, Phe, Thr, Ser, Val, Gly) synthesized on PEGA resin. The library was screened using the radiolabeled folded human telomeric quadruplex from the sequence $(GTTAGG)_5$, in the presence of competitor DNA (full details are given in the Supporting Information). From a potential pool of >25 000 possible sequences, three tetrapeptides were selected (Figure 1B).

To investigate the binding properties of the selected peptides, we employed surface plasmon resonance (SPR) with immobilized DNA. SPR has proven to be a valuable method for the study of DNA−small molecule interactions.[10] Separate lanes of the SPR chip were each loaded with different target DNA to establish both affinity and selectivity. The two DNA targets of interest we selected were human telomeric G-quadruplex (Htelo), folded from 5′-biotinylated $(GTTAGG)_5$,[11] and a double-stranded control DNA (ds DNA) comprising 5′-biotinylated $(GTTAGG)_5$ hybridized with its complementary sequence.

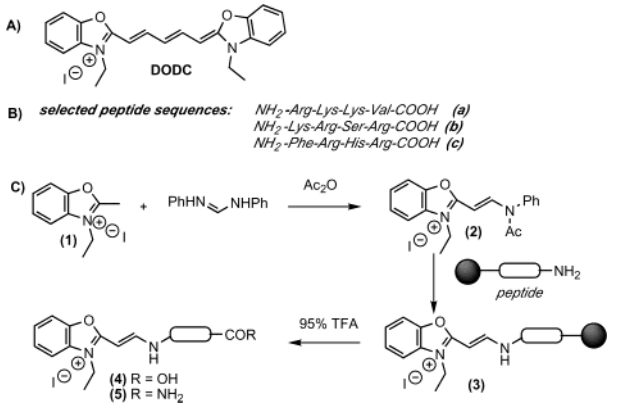


***Figure 1.*** (A) Structure of the cyanine dye DODC; (B) selected peptide sequences; (C) solid-phase synthesis of hemicyanine−peptide conjugates using either polystyrene Wang resin (**4**) or Rink amide resin (**5**).

***Table 1.*** Dissociation Constants ($K_d$) for Hemicyanine (HC)−Peptide Conjugates for Quadruplex and Duplex DNA As Determined by SPR[a]

| | Htelo $K_d$ ($\mu$M) | ds DNA $K_d$ ($\mu$M) | $K_d$(ds)/$K_d$(Htelo) |
|---|---|---|---|
| $NH_2$−RKKV−COOH | 4850 ± 600 | 11 000 ± 700 | 2.3 |
| $NH_2$−KRSR−COOH | 4150 ± 500 | 11 000 ± 600 | 2.6 |
| $NH_2$−FRHR−COOH | 10 010 ± 3200 | >50 000 | >5 |
| Hemicyanine−Peptide−COOH | | | |
| HC−RKKV | 44.1 ± 3.4 | 111 ± 2.3 | 2.5 |
| HC−KRSR | 48.4 ± 11 | >640[b] | >13.3 |
| HC−FRHR | 48.2 ± 0.2 | >640[b] | >13.3 |
| Hemicyanine−Peptide−$CONH_2$ | | | |
| HC−RKKV | 6.6 ± 0.3 | 35.5 ± 10 | 5.4 |
| HC−KRSR | 7.9 ± 0.2 | 86.4 ± 20 | 10.9 |
| HC−FRHR | 14.7 ± 1.6 | >640[b] | >43.4 |

[a] Errors are based on the standard deviation from three independent measurements. [b]Estimated lower limit for $K_d$ based on binding curves exhibited at 640 $\mu$M analyte concentration.

For the three selected tetrapeptides, the dissociation constants ($K_d$) were found to be in the millimolar range, suggestive of weak binding. However, each peptide exhibited a preference for binding quadruplex over double-stranded DNA (Table 1) with FRHR showing 5-fold selectivity.

We next prepared conjugates of the three selected peptides with a hemicyanine core (Figure 1C). The heterocyclic scaffold (**2**) was obtained from the commercially available 3-ethyl-2-methylbenzoxazolium iodide (**1**) via a published route.[12] Peptides were synthesized on polystyrene resin using either a Wang or a Rink amide linker. Removal of the N-terminal Fmoc protecting group enables the generated primary amine to react with the *N*-phenyl-acetate activated heterocycle (**2**), resulting in the corresponding peptide conjugate (**3**). TFA-mediated side chain deprotection and cleavage gave


**10.1021/ja029356w CCC: $25.00** 

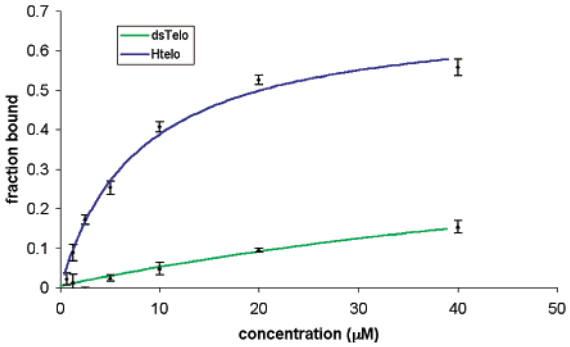


***Figure 2.*** Binding plots (fraction bound versus $C_{free}$) used to determine the $K_d$ value for hemicyanine−KRSR−$CONH_2$ with quadruplex and duplex DNA.

product in very good yields (90%) as either the C-terminal carboxylic acid (**4**) or the carboxamide (**5**) derivative, depending on use of the Wang or Rink amide linker, respectively.[13]

Binding analysis of unconjugated heterocycle (**1**) showed very weak binding to all DNA targets (no detectable binding at the 640 $\mu$M ligand). However, the peptide−hemicyanine conjugates were each found to bind Htelo with a $K_d$ better than 50 $\mu$M (Table 1). The C-terminal carboxylate derivatives (**4**) showed quadruplex binding affinities in the range 44−48 $\mu$M, representing an ∼100-fold increase in affinity as compared to the peptide alone. Analysis of specificity revealed a discrimination factor of 2.5 for the RKKV−$CO_2H$ conjugate, comparable to a value of 2.3 for the peptide alone. Both the KRSR−$CO_2H$ and the FRHR−$CO_2H$ conjugates show a higher discrimination factor of 13.3, which is a significant improvement over the discrimination of the peptides alone.

Initial selection of the tetrapeptides involved C-terminally immobilized peptides that lacked the negatively charged C-terminal carboxylate. This carboxylate may adversely affect nucleic acid binding because of unfavorable charge interactions with the phosphate diester backbone. To address this, we replaced the C-terminal carboxylate group of conjugates of general structure **4** with a neutral carboxamide group to generate three conjugates of general structure **5**. This isosteric replacement by a neutral analogue resulted in a significant enhancement of the binding affinity of all three compounds for G-quadruplex DNA (Table 1). What is particularly striking is that the improvements in binding were also accompanied by an increase in the selectivity of quadruplex over duplex DNA for two out of the three cases (Figure 2). The most promising compound resulting from this study was the FRHR−$CONH_2$ conjugate that was found to bind the human telomeric intramolecular quadruplex with $K_d = 14.7 \pm 1.6$ $\mu$M and a discrimination factor of >40. This level of discrimination is comparable to one of the most selective G-quadruplex ligands reported: a 3,6,9-trisubstituted acridine.[14] At present, we do not fully understand the molecular details of quadruplex recognition by these molecules. The peptides alone bind the intramolecular DNA quadruplex with a relatively high stoichiometry (5:1 for FRHR) suggestive of some nonspecific interactions. However, the conjugates bind the quadruplex with a stoichiometry of 1:1. A UV binding study with DODC shows a red shift, whereas HC−FRHR−$CONH_2$ shows the absence of a red shift (Supporting Information) suggestive of distinct binding modes. On the basis of these studies, we propose a binding mode via loop and groove interactions for these peptide−hemicyanine conjugates without significant $\pi$-stacking with DNA bases.

This study suggests that the coupling of combinatorially selected peptides to a known heterocyclic core is a promising strategy for the relatively fast generation of ligands that bind with respectable affinity and high discrimination. The approach has led to a novel class of G-quadruplex specific ligands. Use of other heterocyclic scaffolds to form peptide conjugates is a logical next step. Studies are ongoing to also elucidate the details of molecular recognition by the molecules described in this paper by X-ray crystallography.

**Acknowledgment.** We thank the BBSRC and the E.U. for financial support.

**Supporting Information Available:** Experimental procedures for solid-phase synthesis of peptide library, on-bead screening, half-dye-peptide conjugates synthesis, LC-MS characterization, UV studies, and SPR protocol (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

## References

(1) Simonsson, T. *Biol. Chem.* **2001**, *368*, 621−628.
(2) Arthanari, H.; Bolton, P. H. *Chem. Biol.* **2001**, *8*, 221−230.
(3) Neidle S.; Parkinson, G. *Nat. Rev. Drug Discovery* **2002**, *1*, 384−393.
(4) Siddiqui-Jain, A.; Grand, C. L.; Bearss, D. J.; Hurley, L. H. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 11593−11598.
(5) (a) 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. (b) Mergny, J. L.; Mailliet, P.; Lavelle, F.; Riou, J. F.; Laoui, A.; Helene, C. *Anti-Cancer Drug Des.* **1999**, *14*, 327−339. (c) Mergny, J. L.; Lacroix, L.; Teulade-Fichou, M. P.; Hounsou, C.; Guittat, L.; Hoaurau, M.; Arinondo, P. B.; Vigneron, J. P.; Lenh, J. M.; Riou, J. F.; Garestier, T.; Helene, C. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 3062−3067. (d) Cuesta, J.; Read, M. A.; Neidle, S. *Mini Rev. Med. Chem.* **2003**, *3*, 11−21.
(6) (a) Isalan, M.; Patel, S. D.; Balasubramanian, S.; Choo, Y. *Biochemistry* **2001**, *40*, 830−836. (b) Schaffitzel, C.; Berger, I.; Postberg, J.; Hanes, J.; Lipps, H.; Plückthorn, A. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 8572−8577.
(7) (a) Carlson, C. B.; Beal, P. A. *Org. Lett.* **2000**, *2*, 1465−1468. (b) Guelev, V. M.; Harting, M. T.; Lokey, R. S.; Iverson, B. L. *Chem. Biol.* **2000**, *7*, 1−8. (c) Ladame, S.; Harrison, R. J.; Neidle, S.; Balasubramanian, S. *Org. Lett.* **2002**, *4*, 2509−2512. (d) Carlson, C. B.; Spanggord, R. J.; Beal, P. A. *ChemBioChem* **2002**, *3*, 859−865.
(8) For example, see: (a) Boriack, P. A.; Christianson, D. W.; Wood, J. K.; Whitesides, G. M. *J. Med. Chem.* **1995**, *38*, 2286−291. (b) Maly, D. J.; Choong, I. C.; Ellman, J. A. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 2491−2424.
(9) Chen, Q.; Kuntz, I. D.; Shafer, R. H. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 2635−2639.
(10) (a) Lacy, E. R.; Le, N. M.; Price, C. A.; Lee, M.; Wilson, W. D. *J. Am. Chem. Soc.* **2002**, *124*, 2153−2163. (b) Carrasco, C.; Facompre, M.; Chisholm, J. D.; Van Vranken, D. L.; Wilson, W. D.; Bailly, C. *Nucleic Acids Res.* **2002**, *30*, 1774−1781. (c) Carrasco, C.; Rosu, F.; Gabelica, V.; Houssier, C.; De Pauw, E.; Garbay-Jaureguiberry, C.; Roques, B.; Wilson, W. D.; Chaires, J. B.; Waring, M. J.; Bailly, C. *ChemBioChem* **2002**, *3*, 1235−1241.
(11) G-quadruplex was folded under standard conditions, and its fully folded structure was confirmed by circular dichroism spectroscopy before being used for SPR studies.
(12) US Patent #2185343, 1937.
(13) All compounds were fully characterized by HPLC, mass spectrometry, and amino acid analysis.
(14) (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) Gowan, S. M.; Harrison, R. J.; Patterson L.; Valenti, M.; Read, M. A.; Neidle, S.; Kelland, L. R. *Mol. Pharmacol.* **2002**, *61*, 1154−1162.

JA029356W