JACS
JOURNAL OF THE AMERICAN CHEMICAL SOCIETY

Communication

pubs.acs.org/JACS

Noncovalent Template-Assisted Mimicry of Multiloop Protein Surfaces: Assembling Discontinuous and Functional Domains

Partha S. Ghosh[†] and Andrew D. Hamilton*[,†,‡]

[†]Department of Chemistry, Yale University, New Haven, Connecticut 06520-8107, United States

[‡]Department of Chemistry, Oxford University, Oxford, U.K. OX13TA

Supporting Information

ABSTRACT: We report here a novel noncovalent synthetic strategy for template-assembled *de novo* protein design. In this approach, a peptide was first conjugated with two oligoguanosine strands via click chemistry and the conjugates were then self-assembled in the presence of metal ions. G-quadruplex formation directs two peptide strands to assemble on one face of the scaffold and form an adjacent two loop surface. This approach can be used to rapidly prepare multiple two-loop structures with both homo- and heterosequences.

There is intense current interest in the identification of molecules that bind to protein targets with affinity and specificity similar to antibodies.[1] In considering different strategies to this end, much can be learned from the mechanisms by which antibodies achieve both high variability and complementarity toward their targets. Critical for this are the nature and positioning of multiple peptide loop domains (the hypervariable loops) that comprise the complementary determining region (CDR) and are held in place by covalent (protein sequence) and noncovalent (protein folding and association) bonds within the immunoglobulin light and heavy chains. In antibody–antigen recognition, several loops in the CDR interact with discontinuous epitopes of an antigen (Figure 1a).

In recent years, several approaches have been taken to the identification of synthetic scaffolds that display multiple peptide loops with structural control over their position and potential recognition properties, in direct analogy to antibodies. In a seminal study, Mutter has developed a template assembled synthetic protein approach (TASP) in which as many as three peptide loops are covalently linked across a macrocyclic oligopeptide template (Figure 1b(i)).[2] This strategy has been extended by Fairlie to other macrocyclic scaffolds including linked oxazoles and thiazoles.[3] In an alternative approach, we[4] and others[5] have shown that protein binding agents can be prepared through the covalent attachment of several preformed synthetic peptide loops to a macrocyclic scaffold that is itself covalently or noncovalently formed. Recently in an important development, Heinis and Winter have created phage display libraries of double peptide loop derivatives (bicycles) through the triple alkylation of three defined Cys residues within the randomized peptide sequence.[6] A related strategy based on Cys-containing synthetic peptide libraries has been reported by Timmerman for the preparation of loop structures.[7]

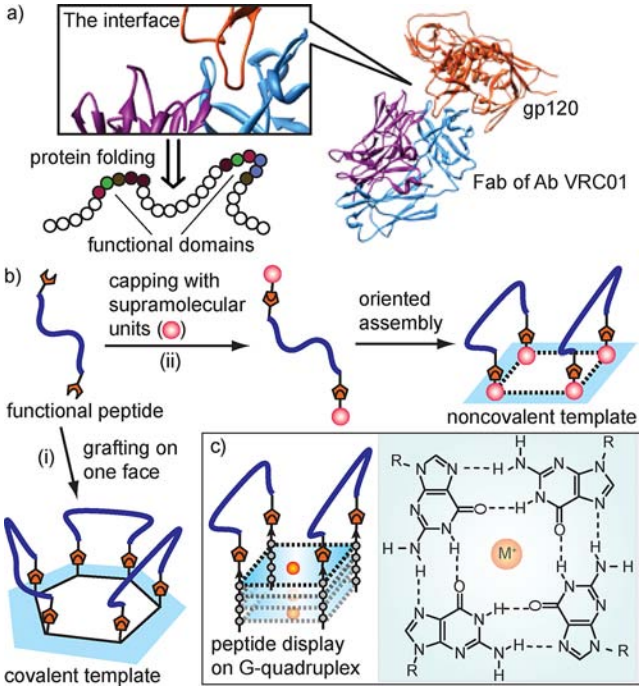


Figure 1. (a) An example of antigen–antibody interactions: gp120 (HIV envelope protein; orange)-VRC01 (antibody for gp120; V_H in blue, V_L in magenta), pdb 3NGB. (b) Covalent (i) and noncovalent (ii) strategies for mimicking multiloop surfaces. (c) Display of two constrained peptides on one face (5′) of G-quadruplex.

Each of these strategies involves covalently constraining the chosen peptide sequence through a series of synthetic modifications and thus lacks the direct and powerful simplicity of the antibody whereby the loops form through noncovalently stabilized protein folding. We sought a synthetic solution in which functionalization of a peptide at either end with supramolecular capping groups, that can themselves non-covalently interact intra- and intermolecularly, might lead to the self-assembly of multiple adjacent peptide loops (as in Figure 1b(ii)). In this way, a noncovalent template would direct oriented assembly of peptide fragments, simultaneously enforcing their conformational restrictions. A suitable supramolecular fragment for directed folding is represented by a strand of oligonucleotides, which depending on their sequence

Received: June 3, 2012
Published: July 27, 2012

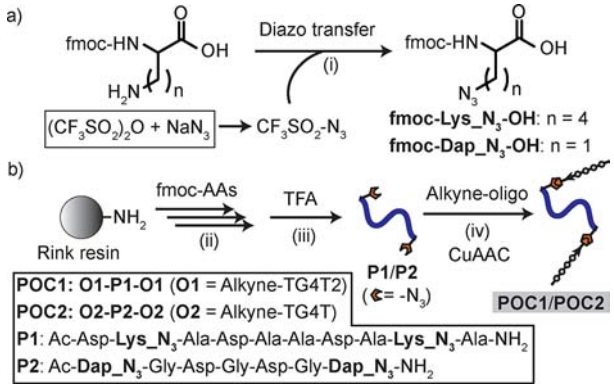
   dx.doi.org/10.1021/ja305360q | *J. Am. Chem. Soc.* 2012, 134, 13208–13211

can form a range of different aggregate structures and stoichiometries.[8] In this paper, we show a first example of this strategy in which two oligoguanosine strands are attached to both N- and C-termini of a peptide through their 5′ ends. Upon addition of metal ions, formation of parallel G-quadruplexes should take place spontaneously in aqueous media via synergistic H-bonding, metal−ligand and aromatic stacking interactions (Figure 1c). The oriented assembly will thus be induced to display two loops from the same face of the template. The noncovalent synthetic strategy can be further exploited to create a combination of multiloop structures.

The key challenge is the ready preparation of linear oligomeric components composed of oligoG−peptide−oligoG strands. The chemistry of conventional peptide solid-phase synthesis is incompatible with that of oligonucleotides, which makes difficult the linear synthesis of peptide−oligonucleotide conjugates (POCs) on a single support. Instead a fragment conjugation approach was developed involving copper(I)-catalyzed azide−alkyne [3 + 2] cycloaddition (CuAAC) as a ligation method. The CuAAC proceeds rapidly with high yield in an aqueous medium.[9] This bioorthogonal click reaction can be exploited for any peptide sequence requiring no protecting groups on either the loop or the template. The optimal strategy involved reacting two alkyne functionalized oligonucleotides with a peptide strand containing two azide groups near the C- and N-termini.

We first synthesized Fmoc-protected azido amino acids (**AA_N$_3$s**) from the corresponding amino acids by Cu(II)-catalyzed diazo transfer reaction (Scheme 1a). These were then

Scheme 1. Synthesis of OligoG−Peptide−OligoG Conjugates: (a) Fmoc-azide Amino Acids (AA_N$_3$s) via diazo transfer reaction, (b) peptides with AA_N$_3$s near the Termini by SPPS and Peptide−Oligo Conjugation by CuAAC[a]



[a](i) $NaHCO_3$, $CuSO_4$, H_2O/MeOH/CH_2Cl_2, rt, ON; (ii) (a) HOBt, HBTU, DIPEA, DMF, rt, 1 h; (b) piperidine, DMF, rt, 15 min; (iii) H_2O, rt, 4 h (iv) 1:10:5 $CuSO_4$:NaAsc:THPTA, 2:1 TEAA:^{t}BuOH, rt, 3 h.

incorporated into model peptides (**P1** and **P2**) by Fmoc-based solid-phase chemistry. The peptides differ in molecular weight, location of the ligation points (**AA_N$_3$s**) and the distance between two **N$_3$-AAs**. **P1** contains aspartic acid and alanine with the **Lys_N$_3$s** separated by six residues. **P2** consists of five aspartic acid and glycine residues with two **Dap_N$_3$s** at the termini.

The higher molecular weight peptide **P1** (>**P2**) was conjugated with higher molecular weight oligonucleotide **O1** (>**O2**) to give sufficient mass difference between the conjugates (**O1−P1−O1** ≫ **O2−P2−O2**, vide infra). The desired double click product was not obtained using $CuSO_4$/sodium ascorbate (NaAsc) mixture with 1 equiv Cu/azide, as reported.[10] However, using tris(3-hydroxypropyltriazolylmethyl)amine (THPTA), a water-soluble Cu(I)-stabilizing ligand, efficient conjugation was achieved {$CuSO_4$/NaAsc/THPTA (1:10:5; 10 equiv Cu/azide) in 2:1 TEAA/^{t}BuOH (TEAA = triethylamine acetate buffer, 200 mM, pH 7.2)}.[11] The POCs were purified by RP-HPLC and characterized by MALDI-TOF.

The POCs were then self-assembled in metal ion (K^+ or Na^+) containing buffer (10 mM Tris-HCl, 80 mM salt, pH 7.4) by first heating at 90 °C for 10 min, followed by slow cooling to room temperature and finally incubation at 4 °C for an additional 48 h.

The structure of the annealed POCs was characterized by nondenaturing polyacrylamide gel electrophoresis (PAGE) and circular dichroism (CD). As depicted in Figure 2a, in the presence of metal ions, the POCs could reasonably assemble

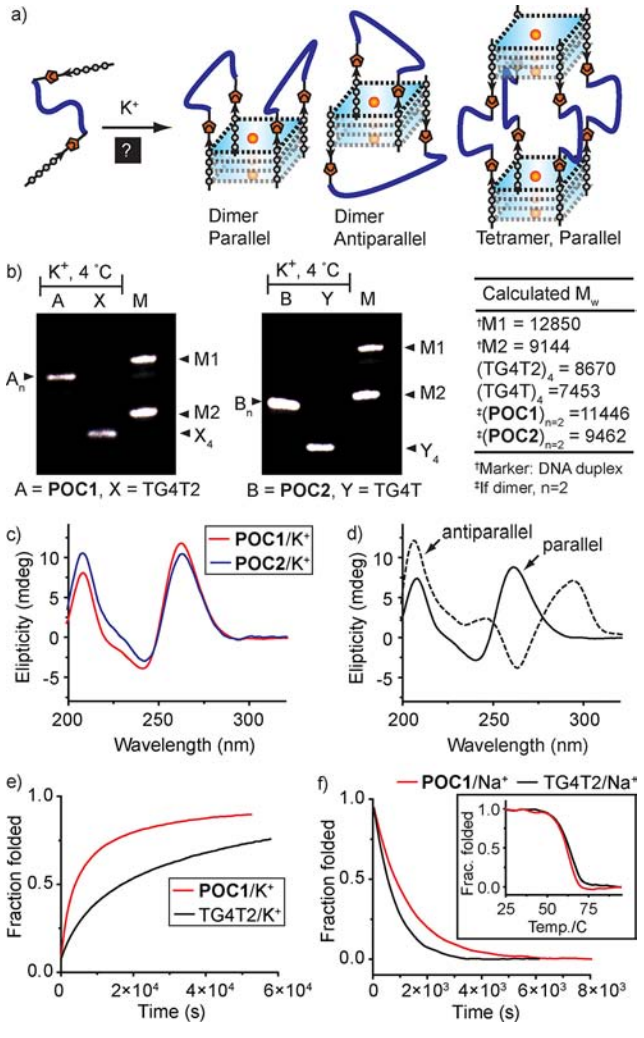


Figure 2. (a) Alternative folded structures of self-assembled POCs. (b) PAGE images showing POCs are forming dimeric complexes according to their electromobility shifts. (c and d) CD spectra of the POC complexes and known parallel and antiparallel G-quadruplexes. (e) Kinetics of G-quadruplex formation of **POC1** and TG4T2. (f) Dissociation kinetics of the complexes at their $T_{1/2}$, inset shows their melting profile.

dx.doi.org/10.1021/ja305360q | J. Am. Chem. Soc. 2012, 134, 13208−13211

into two dimeric or one tetrameric forms. To assess the molecularity of the association, we performed a PAGE assay using short oligonucleotide duplexes (15-mer and 21-mer) as molecular weight markers and known tetramers ($(TG4T2)_4/K^+$) and $(TG4T)_4/K^+$) as a positive control. The gel mobility shift of POCs/K^+ indicated that the POCs assembled into a bimolecular quadruplex structure (Figure 2b).

Dimeric assembly of the POCs can take two quadruplex topologies depending on the orientation of the oligonucleotide strands in a parallel or antiparallel direction (see Figure 2a). To determine the orientation, we compared the circular dichroism spectra of the assembled POCs with those of known parallel and antiparallel quadruplexes ($(TG4T2)_4/K^+$) and $(TG4T4G4T)_2/Na^+$, respectively).[12] We observed the characteristic peaks of a parallel DNA quadruplex for both **POC1** and **2** in the presence of K^+ (maxima at 263 and 203 nm and a minimum at 242 nm, Figure 2c,d).

To better understand the assembly, we studied the kinetics of POC association (at 4 °C). That **POC1** formed the G-quadruplex core faster than control TG4T2 is likely due to an initial intramolecular association of two oligoguanosine strands attached to the peptide followed by their bimolecular assembly to form the quadruplex core (Figure 2e). We also studied dissociation behavior of the assembled structures, in Na^+-solution as the K^+-complexes did not denature even at high temperature. The melting profile of the assembled structures with **POC1** was similar to those with $(TG4T2)_4/Na^+$ giving further support for dimeric association of **POC1** forming four G-tetrads as opposed to tetrameric complexes with eight G-tetrads and hence a higher expected melting temperature (Figure 2f inset). By comparing the dissociation kinetics at a particular temperature (60 °C), we found that the presence of the loops slowed down the melting of the quadruplex core (Figure 2f).

Figure 3a shows a molecular model of the double peptide loop structure obtained by keeping the quadruplex core fixed

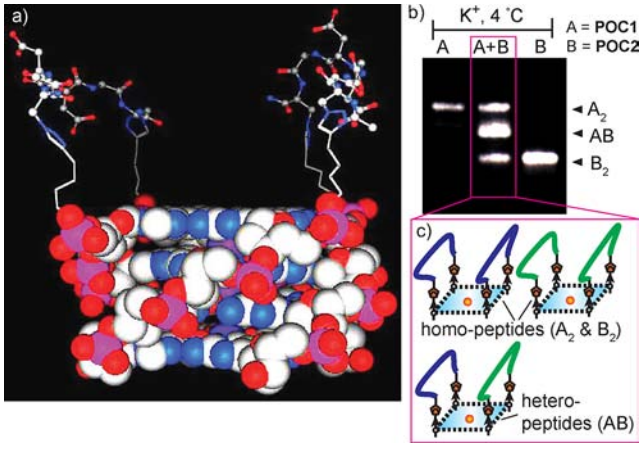


Figure 3. (a) Molecular model of the assembled structures. (b) PAGE of the complexes after coassembly of **POC1** and **POC2** in presence of K^+. (c) Schematic illustration of the structures formed (a statistical mixture) during the coassembly.

and then energy minimizing the peptide fragment. The result shows a potential binding surface defined by the sequence and conformations of the peptide loops as well as their relative distance across the G-quadruplex surface.

This approach can provide a direct route to combinatorial libraries by multicomponent self-assembly.[13] For example, by combining two from a library of n POCs, comprising different peptide sequences, it is possible to formulate a diverse collection of double loop structures with n homo- and $n(n-1)/2$ heterosequences (total $n(n+1)/2$). To demonstrate this, we co-incubated equimolar ratios of **POC1** and **POC2** in K^+-containing buffer. The annealed solution was then analyzed by PAGE using the homoassemblies ($\mathbf{POC1}_2$ and $\mathbf{POC2}_2$) as a reference. As shown in Figure 3b, three bands appeared during co-assembly of the two POCs (middle lane), which correspond to their statistical mixture (Figure 3c).

In conclusion, by using an improved peptide–oligonucleotide click conjugation methodology, we have developed a self-organizing structure that through G-quadruplex formation positions two peptide loops on one surface, in direct analogy to antibody binding domains. The multicomponent non-covalent synthesis allows facile formation of homo- and heterocombinations.

■ ASSOCIATED CONTENT

Supporting Information

Synthetic method of POC, experimental protocols and supplementary figures. This material is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

andrew.hamilton@admin.ox.ac.uk

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was financially supported by the National Science Foundation (CHE0750357) and Yale University. We also thank Dr. Ian Jones for his assistance.

■ REFERENCES

(1) (a) Williams, B. A. R.; Diehnelt, C. W.; Belcher, P.; Greving, M.; Woodbury, N. W.; Johnston, S. A.; Chaput, J. C. *J. Am. Chem. Soc.* **2009**, *131*, 17233–17241. (b) Villen, J.; Rodriguez-Mias, R. A.; Nunez, J. I.; Giralt, E.; Sobrino, F.; Andreu, D. *Chem. Biol.* **2006**, *13*, 815–823.

(2) Mutter, M.; Dumy, P.; Garrouste, P.; Lehmann, C.; Mathieu, M.; Peggion, C.; Peluso, S.; Razaname, A.; Tuchscherer, G. *Angew. Chem., Int. Ed.* **1996**, *35*, 1482–1485.

(3) Singh, Y.; Stoermer, M. J.; Lucke, A. J.; Guthrie, T.; Fairlie, D. P. *J. Am. Chem. Soc.* **2005**, *127*, 6563–6572.

(4) (a) Hamuro, Y.; Calama, M. C.; Park, H. S.; Hamilton, A. D. *Angew. Chem., Int. Ed.* **1997**, *36*, 2680–2683. (b) Park, H. S.; Lin, Q.; Hamilton, A. D. *J. Am. Chem. Soc.* **1999**, *121*, 8–13. (c) Cai, J. F.; Rosenzweig, B. A.; Hamilton, A. D. *Chem.—Eur. J.* **2009**, *15*, 328–332.

(5) (a) Chamorro, C.; Kruijtzer, J. A. W.; Farsaraki, M.; Balzarini, J.; Liskamp, R. M. J. *Chem. Commun.* **2009**, 821–823. (b) Pfeifer, M. E.; Robinson, J. A. *Chem. Commun.* **1998**, 1977–1978.

(6) Heinis, C.; Rutherford, T.; Freund, S.; Winter, G. *Nat. Chem. Biol.* **2009**, *5*, 502–507.

(7) Timmerman, P.; Beld, J.; Puijk, W. C.; Meloen, R. H. *ChemBioChem* **2005**, *6*, 821–824.

(8) Rosenzweig, B. A.; Ross, N. T.; Adler, M. J.; Hamilton, A. D. *J. Am. Chem. Soc.* **2010**, *132*, 6749–6754.

(9) (a) Kolb, H. C.; Finn, M. G.; Sharpless, K. B. *Angew. Chem., Int. Ed.* **2001**, *40*, 2004–2021. (b) Beatty, K. E.; Tirrell, D. A. *Bioorg. Med. Chem. Lett.* **2008**, *18*, 5995–5999.

(10) Gogoi, K.; Mane, M. V.; Kunte, S. S.; Kumar, V. A. *Nucleic Acids Res.* **2007**, *35*, e139.

(11) (a) Hong, V.; Presolski, S. I.; Ma, C.; Finn, M. G. *Angew. Chem., Int. Ed.* **2009**, *48*, 9879−9883. (b) Ghosh, P. S.; Hamilton, A. D. *Chem.—Eur. J.* **2012**, *18*, 2361−2365.
(12) (a) Lu, M.; Guo, Q.; Kallenbach, N. R. *Biochemistry* **1992**, *31*, 2455−2459. (b) Miyoshi, D.; Nakao, A.; Sugimoto, N. *Biochemistry* **2002**, *41*, 15017−15024.
(13) Roy, L.; Case, M. A. *J. Am. Chem. Soc.* **2010**, *132*, 8894−8896.