Encoded combinatorial libraries for the construction of cyclic peptoid microarrays†

Yong-Uk Kwon[ab] **and Thomas Kodadek***[a]

A "one bead two compound" approach to the synthesis of encoded cyclic peptoid libraries is reported.

Drug-like small molecules (<500 Daltons) generally do not bind well to the relatively shallow surfaces of proteins involved in protein-protein interactions. Thus, in order to develop effective therapeutic agents against these increasingly important targets, it is necessary to develop libraries of compounds able to cover a greater surface area and engage in multiple contacts with the target protein, as well as efficient methods to screen these libraries. With regard to their protein-binding properties, peptides are an attractive class of molecules, but linear peptides have many undesirable features. They are peptidase- and protease-sensitive, relatively cell impermeable and generally form complexes with only modest dissociation constants (K_Ds in the high nM to mid μM range). However, cyclic peptides can exhibit enhanced cell permeability[1,2] and are much less sensitive to enzymatic degradation.[3] Moreover, it is presumed that the conformational restriction imposed by cyclization may generally afford higher binding affinities, though rigorous proof for this idea is lacking.[4,5] Indeed, many naturally occurring cyclic peptides and depsipeptides have been found to display potent biological activities.[6–10] This interest has led to the development of facile methods for the creation of either synthetic[11] or genetically encoded[12,13] libraries of cyclic peptides as potential sources of drug leads.

A limitation of peptide libraries, cyclic or linear, is that only a relatively small number of building blocks are available. Moreover, although cyclic peptides, can be more cell permeable than their linear counterparts, this appears to be dependent on their ability to form intramolecular hydrogen bonds,[2] a property that is likely to vary from compound to compound. Therefore, we became interested in the development of libraries of cyclic peptoids (*N*-substituted oligoglycines)[14,15] as potential protein ligands. Large libraries of linear peptoids with a wide variety of different side chains[16–18] are readily accessible using split and pool methods and efficient protocols with which to screen these libraries for protein binding have

been developed.[18–24] These studies have confirmed that peptoid libraries are indeed rich sources of protein ligands. We have also shown that peptoids are, in general, far more cell permeable than peptides due to their lack of polar main chain N–H bonds.[25] In this report, we describe a strategy that is optimized for the creation of cyclic peptoid microarrays.

A key design issue has to do with the ability to determine the sequence of hits after screening a one bead one compound library. Since cyclic peptides or peptoids lack a free *N*-terminus, Edman sequencing cannot be employed. Moreover, while peptoids, like peptides, can be sequenced by tandem mass spectrometry,[26] cyclic molecules will fragment at multiple positions, complicating interpretation of the MS/MS spectrum severely. This issue has limited the development of synthetic cyclic peptide libraries. Pei and co-workers addressed this problem recently by developing a "two compound, one bead" approach in which each bead contains both a linear and cyclic molecule containing the same peptide sequence.[11] In other words, the linear molecule encodes the cyclic molecule. This was accomplished using the strategy of Lam[27] in which different solvents were employed to segregate beads into two different domains (internal and surface-exposed) to which were attached glutamic acid residues with differentially protected carboxylate side chains. The same peptide chain was then extended from both the internal and external Glu residues. Finally, only the surface-exposed Glu side chains were deprotected, allowing them to be cyclized with the terminal amino group of the peptide. The peptides in the internal layer remained linear and thus served as the encoding strand.

We have developed a different one bead two compound strategy that is tailored to the creation of microarrays, a useful platform for protein fingerprinting and library screening.[28,29] The idea was to employ differential deprotection to create two chains, both of which contain the peptoid of interest, but only one of which contains both a glutamic acid residue to support cyclization as well as a Cys residue to allow specific conjugation of only the cyclic peptoid molecule to a maleimide-activated microscope slide[30] (Fig. 1). The linear molecule would not couple to the slide, but would be present to support tandem MS-based sequencing.

To effect this strategy (see Fig. 2), a 7 : 1 ratio of Fmoc-Cys(Trt)-OH and ivDde-β-Ala-OH was added to β-Ala-primed Rink amide resin. This ratio was optimized empirically to provide enough linear peptoid for tandem MS sequencing from a single bead, but also produce as much cyclic peptoid as possible. After selective deprotection of Fmoc, Fmoc-β-Ala-OH was again attached to Cys followed, after

[a] *Division of Translational Research, Departments of Internal Medicine and Molecular Biology, University of Texas Southwestern Medical Center, 5323 Harry Hines Boulevard, Dallas, Texas 75390-9185, USA. E-mail: thomas.kodadek@utsouthwestern.edu; Fax: 1-214-648-4156; Tel: 1-214-648-1239*
[b] *Department of Chemistry and Nano Science, Ewha Womans University, 120-750 Seoul, South Korea*
† Electronic supplementary information (ESI) available: Experimental details for synthesis and microarray construction, supplemental data. See DOI: 10.1039/b812735b
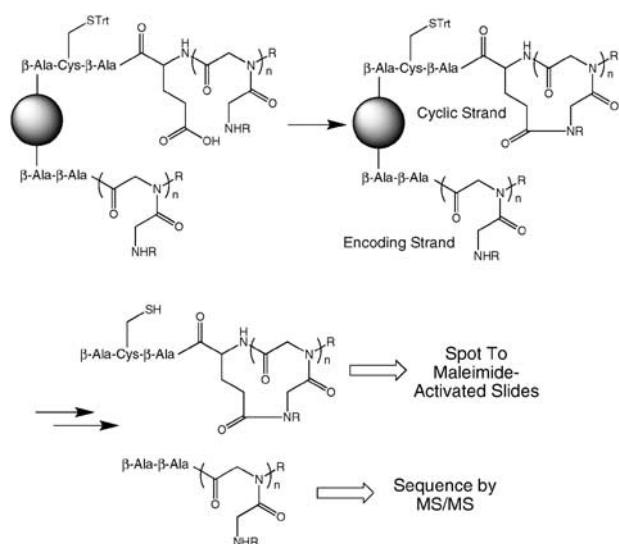
Fig. 1 Schematic view of the general strategy employed to create a library in which each bead carries a cyclic peptoid and an analogous linear encoding strand. Only the cyclic molecule contains a thiol and thus will couple to a maleimide-activated glass slide.
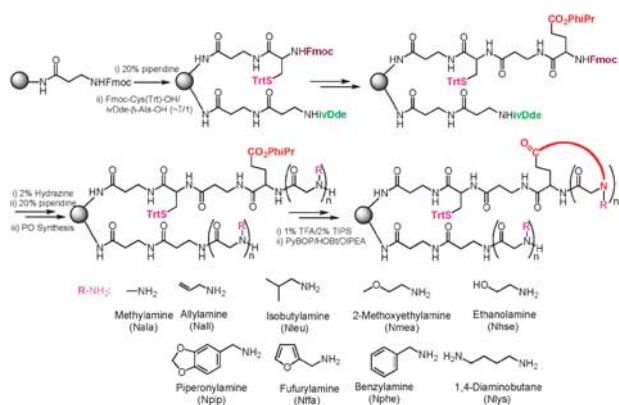

Fig. 2 Synthesis of the encoded cyclic peptoid library *via* the one bead two compound strategy. The amines employed in the sub-monomer peptoid synthesis are shown at the bottom of the figure (one of the amines in 1,4-diaminobutane and a hydroxyl group in ethanolamine were protected).

removal of this Fmoc, by addition of Fmoc-Glu(*O*-2-PhiPr)-OH. At this point, both the Fmoc and ivDde protetecting groups were removed and peptoid synthesis was carried out on both strands. Seven peptoid residues were incorporated using conventional sub-monomer chemistry and the nine amines shown in Fig. 2. The 2-PhiPr protecting group on the Glu side chain was then removed selectively with 1% TFA. Finally, macrocyclization was carried out using the method of Kirshenbaum and colleagues (PyBOP (3 eq), HOBt (3 eq) and DIPEA (10 eq).[15] Note that the linear molecule lacks two residues present in the (presumed) cyclic molecule and thus the mass peaks derived from each can be distinguished easily, facilitating analysis and sequence determination.

To determine the efficacy of this procedure, individual beads were separated and treated with acid to cleave the molecules from the beads, followed by HPLC, MS and MS/MS analysis. Gratifyingly, we found that in almost every case the sequence

of the peptoid on a particular bead could be determined easily by tandem MS analysis of the linear molecule. We therefore turned our attention to the determination of the efficiency of the cyclization reaction. For some of the molecules, mass spectrometry and HPLC analysis showed that cyclization of the Cys-Glu-containing molecule was clearly incomplete, as linear starting material was clearly detectable. This was not surprising, since a general problem in the creation of cyclic libraries is that not all sequences cyclize with equivalent efficiencies.[31] One would presume that the nature of the *N*-terminal residue would have the largest effect on cyclization. Indeed, an analysis of more than 50 peptoids by MS/MS
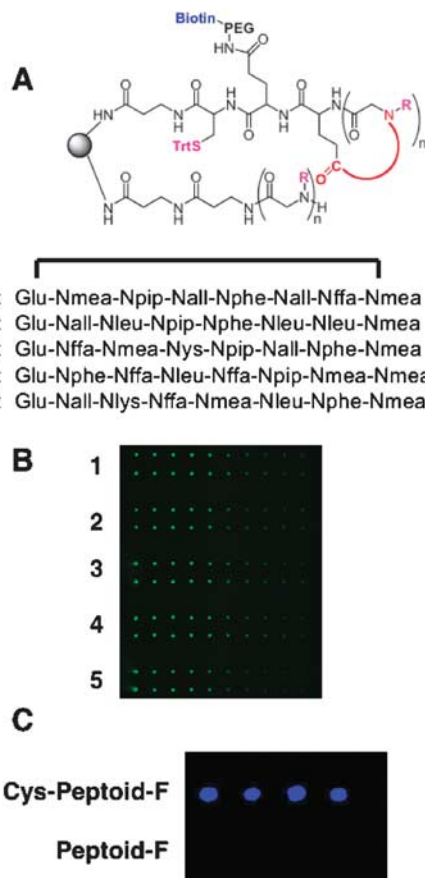

Fig. 3 Attachment of Cys-containing cyclic peptoid to a maleimide-activated glass slide. A. General structure of the cyclic and linear molecules made on each bead before cleavage and deprotection of the thiol side chain. Below: Sequences of the variable regions of five peptoids picked for the spotting experiment. B. Fluorescent image of microarrays in which each of the five peptoids have been spotted onto the activated surface. A DMSO solution of each peptoid (≈ 2 mM) was spotted two times, the solution was diluted three-fold, spotted again, *etc*. After washing and drying, the arrays were hybridized with Cy3-conjugated streptavidin, washed and the slide was scanned with a fluorescence scanner (see ref. 30 for details). The spots are false-colored green. C. The Cys is essential for retention of the peptoid on the microarray. Two peptoids were synthesized. Each had the sequence Fluorescein-Nlys-Nser-Nleu-Nser-Nall-Npip-Nlys-Nlys. One peptoid also contained a C-terminal cysteine, while the other did not. The two peptoids were spotted onto a maleimide-activated glass slide. After washing, the slide was scanned using a fluorescence scanner. The fluorescence intensity is false-colored blue.

revealed that if the *N*-terminal residue was Nmea, the cyclization yield was almost quantitative. Therefore, this residue was made invariant in all subsequent work.

With this information in hand, we prepared ten peptoids of the form: β-Ala-Cys-Glu(Biotin)-cyclo(Glu-X-X-X-X-X-X-Nmea) (see Fig. 3), where biotin-Glu bears a side chain-conjugated biotin and X was derived from one of the amines shown in Fig. 2. The molecules were cleaved from the resin and analyzed by HPLC and tandem MS. In each case, we were able to easily sequence the linear species by tandem mass spectrometry (see Supplementary Material). Moreover, all of the detectable Cys-containing molecules were in the cyclic form (see Supplementary Material).

Serial dilutions of the five peptoids shown in Fig. 3 were spotted robotically onto PEGylated, maleimide-activated glass microscope slides[30] and the slides were then washed rigorously. In order to demonstrate the immobilization of the cyclic peptoids, the slides were incubated with Cy3-labeled Streptavidin and scanned. As expected, the amount of protein captured decreased as the amount of peptoid spotted decreased, confirming that the fluorescence is indeed due to specific capture of the protein by the peptoid (Fig. 3B). To demonstrate that the Cys residue is critical for retention of the cyclic peptoid to the maleimide-derivatized slide, we synthesized two fluorescein-conjugated linear peptoids that were identical except for the presence and absence of Cys. These were spotted onto a slide, which was then scanned after washing. As shown in Fig. 3C, detectable fluorescence was seen only where the Cys-containing peptoid was spotted. This experiment validates an important feature of our strategy, which is that the linear encoding molecule (see Fig. 1) will not be retained on the slide when the mixture of it and the Cys-containing cyclic molecule are spotted onto the slide and thus will not interfere with screening experiments.

In summary, we have developed a new 'one bead two compound' strategy for the creation of encoded cyclic peptoid libraries. This particular scheme is useful for the creation of cyclic peptoid microarrays since only the cyclic peptoid, not the linear encoding molecule, contains a cysteine and thus can be spotted onto a maleimide-activated microscope slide. The use of these cyclic peptoid microarrays in screening experiments will be reported in due course.

Notes and references

1 T. Rezai, J. E. Bock, M. V. Zhou, C. Kalyanaraman, R. S. Lokey and M. P. Jacobson, *J. Am. Chem. Soc.*, 2006, **128**, 14073–14080.

2 T. Rezai, B. Yu, G. L. Millhauser, M. P. Jacobson and R. S. Lokey, *J. Am. Chem. Soc.*, 2006, **128**, 2510–2511.

3 T. Satoh, S. Li, T. M. Friedman, R. Wiaderkiewicz, R. Korngold and Z. Huang, *Biochem. Biophys. Res. Commun.*, 1996, **224**, 438–443.

4 D. G. Udugamasooriya and M. R. Spaller, *Biopolymers*, 2008, **89**, 653–667.

5 S. F. Martin, *Pure Appl. Chem.*, 2007, **79**, 193–200.

6 S. Ho, N. Clipstone, L. Timmermann, J. Northrop, I. Graef, D. Fiorentino, J. Nourse and G. R. Crabtree, *Clin. Immunol. Immunopathol.*, 1996, **80**, S40–S45.

7 S. Banerjee, Z. Wang, M. Mohammad, F. H. Sarkar and R. M. Mohammad, *J. Nat. Prod.*, 2008, **71**, 492–496.

8 E. Lech-Maranda, E. Robak, A. Korycka and T. Robak, *Mini-Rev. Med. Chem.*, 2007, **7**, 1062–1069.

9 M. Fouladi, *Cancer Invest.*, 2006, **24**, 521–527.

10 Y. Hamada and T. Shioiri, *Chem. Rev.*, 2005, **105**, 4441–4482.

11 S. H. Joo, Q. Xiao, Y. Ling, B. Gopishetty and D. Pei, *J. Am. Chem. Soc.*, 2006, **128**, 13000–13009.

12 C. P. Scott, E. Abel-Santos, M. Wall, D. C. Wahnon and S. J. Benkovic, *Proc. Natl. Acad. Sci. U. S. A.*, 1999, **96**, 13638–13643.

13 N. Venkatesh, S.-H. Im, M. Balass, S. Fuchs and E. Katchalski-Katzir, *Proc. Natl. Acad. Sci. U. S. A.*, 2000, **97**, 761–766.

14 R. J. Simon, R. S. Kania, R. N. Zuckermann, V. D. Huebner, D. A. Jewell, S. Banville, S. Ng, L. Wang, S. Rosenberg and C. K. Marlowe and *et al.*, *Proc. Natl. Acad. Sci. U. S. A.*, 1992, **89**, 9367–9371.

15 S. B. Shin, B. Yoo, L. J. Todaro and K. Kirshenbaum, *J. Am. Chem. Soc.*, 2007, **129**, 3218–3225.

16 G. M. Figliozzi, R. Goldsmith, S. C. Ng, S. C. Banville and R. N. Zuckermann, *Methods Enzymol.*, 1996, **267**, 437–447.

17 T. Horn, B. C. Lee, K. A. Dill and R. N. Zuckermann, *Bioconjugate Chem.*, 2004, **15**, 428–435.

18 P. G. Alluri, M. M. Reddy, K. Bacchawat-Sikder, H. J. Olivos and T. Kodadek, *J. Am. Chem. Soc.*, 2003, **125**, 13995–14004.

19 P. G. Alluri, B. Liu, Y. P., X. Xiao and T. Kodadek, *Mol. BioSyst.*, 2006, **2**, 568–579.

20 X. Xiao, P. Yu, H.-S. Lim, D. Sikder and T. Kodadek, *J. Comb. Chem.*, 2007, **9**, 592–600.

21 M. M. Reddy, K. Bachhawatt-Sikder and T. Kodadek, *Chem. Biol.*, 2004, **11**, 1127–1137.

22 H.-S. Lim, C. T. Archer and T. Kodadek, *J. Am. Chem. Soc.*, 2007, **129**, 7750–7751.

23 R. N. Zuckermann, E. J. Martin, D. C. Spellmeyer, G. B. Stauber, K. R. Shoemaker, J. M. Kerr, G. M. Figliozzi, D. A. Goff, M. A. Siani and R. J. Simon and *et al.*, *J. Med. Chem.*, 1994, **37**, 2678–2685.

24 D. G. Udugamasooriya, S. P. Dineen, R. A. Brekken and T. Kodadek, *J. Am. Chem. Soc.*, 2008, **130**, 5744–5752.

25 Y. U. Kwon and T. Kodadek, *J. Am. Chem. Soc.*, 2007, **129**, 1508–1509.

26 M. G. Paulick, K. M. Hart, K. M. Brinner, M. Tjandra, D. H. Charych and R. N. Zuckermann, *J. Comb. Chem.*, 2006, **8**, 417–426.

27 R. Liu, J. Marik and K. S. Lam, *J. Am. Chem. Soc.*, 2002, **124**, 7678–7680.

28 G. MacBeath, A. N. Koehler and S. L. Schreiber, *J. Am. Chem. Soc.*, 1999, **121**, 7967–7968.

29 M. Uttamchandani, D. P. Walsh, S. Q. Yao and Y.-T. Chang, *Curr. Opin. Chem. Biol.*, 2005, **9**, 4–13.

30 M. M. Reddy and T. Kodadek, *Proc. Natl. Acad. Sci. U. S. A.*, 2005, **102**, 12672–12677.

31 S. Li, N. Marthandan, D. Bowerman, H. R. Garner and T. Kodadek, *Chem. Commun.*, 2005, 581–583.