SYNTHESIS OF LONG NON-NATURAL SEQUENCE-SPECIFIC HETEROPOLYMERS

Byoung-Chul Lee[†] Ken A. Dill[†‡] and Ronald N. Zuckermann[§]*

Graduate group in Biophysics[†] and Department of Pharmaceutical Chemisty[‡], 600 16th St., University of California, San Francisco, CA 94143 and Chiron Corporation[§], 4560 Horton Street, Emeryville, CA 94608.

Introduction

The only molecules currently known to fold into unique three-dimensional conformations and perform sophisticated functions are biological polymers – proteins and nucleic acids. Our aim is to see if we can create foldable molecules in aqueous solution using a non-biological sequence-specific polymer. This would be a significant step toward the design and synthesis of artificial proteins.

There are many recent efforts to construct synthetic polymers that mimic protein properties[1]. But, it has been difficult to achieve stable secondary structures, a diversity of monomers *and* long chain lengths in the same polymer system. However, the recent invention of peptoid polymers has opened up the opportunity for synthesizing specific sequences of significant length that take advantage of a wide array of possible monomers[2]. Peptoids are N-substituted glycine oligomers where the side chains are appended to the backbone nitrogen. They are prepared by a solid-phase submonomer method where each monomer is added sequentially[3].

The chemical synthesis of a single-chain four-helix bundle presents a major technical hurdle. Such a sequence would need to be approximately 60 monomers in length. This would require the yield of a monomer addition cycle to be greater than 99% to obtain a 55% yield of final product. Because of this technical difficulty, we used a chemical conjugation strategy to link individual 15mer helical units together[4]. Importantly, chemical conjugation allows ligating a variety of helical analogs combinatorially (with either head-to-head or head-to-tail orientations) in order to find an optimal configuration of side-chain packing.

Here we describe the synthesis of a peptoid 60mer containing 4 helical units. To link helices together, we used disulfide linkages (derived from a sulfhydryl group and an activated disulfide) and oxime linkages (derived from an aldehyde and an aminooxy group). We previously developed these chemoselective groups for peptoids because the ligation reactions occur in high yield and are orthogonal to other side chain functionalities[5].

Experimental

Amine submonomers for peptoid synthesis. Commercially available amines were purchased from Aldrich Chemical Co. (Milwaukee, WI), Fluka (Buchs, Switzerland), Acros Organics (Morris Plains, NJ), TCI America (Philadelphia, PA) and Bachem California, Inc. (Torrance, CA).

(S)-N-(1-phenylethyl) glycine (Nspe), (S)-N-(1-cyclohexylethyl) glycine (Nsch), N-(1-phenylmethyl) glycine (Npm), N-(1-cyclohexylmethyl) glycine (Nchm), N-(1-diphenylethyl) glycine (Ndpe), N-(2-carboxyethyl) glycine (Nglu), N-(methoxyethyl) glycine (Nme) and N-(2-nitrophenol) glycine (Nnp) were derived from the amines L(-)-α-methylbenzylamine (Acros), S(+)-1-cyclohexylethylamine (Fluka), benzylamine (Aldrich), cyclohexylmethylamine (Aldrich), 2,2-diphenylethylamine (Aldrich), *t*-butyl-β-alanine (Bachem), 2-methoxyethylamine (TCI America) and 4-amino-2-nitrophenol (Aldrich), respectively. (S)-N-(N'-2-hydroxylethyl alaninamide) glycine (Nsahe), (S)-N-(N'-2-carboxymethyl alaninamide) glycine (Nsacm) and (S)-N-(N'-2-aminoethyl alaninamide) glycine (Nsaae) with suitable protecting groups were synthesized as previously described[6]. N-(2-hydroxyethyl) glycine (Nser) and N-(2-aminoethyl) glycine (Nae) were derived from [(triisopropylsilyl)oxy]ethylamine and N-*t*-BOC-1,2-diaminoethane which were prepared as described[7]. N-(2-anthranilamido-ethyl) glycine (Naae) were derived from 2-*t*-butyloxycarbonylamino-1-aminoethyl-phenylamide that was made by coupling between N-CBZ-ethylene diamine and 2-*t*-butyloxycarbonylamino benzoate and removing the CBZ group with catalytic hydrogenation. The coupling reaction was same as described previously. All other solvents and reagents were obtained from commercial sources and used without further purification.

Peptoid synthesis with N- and C- terminal modifications. Peptoid oligomers were synthesized on an automated peptoid/peptide robot synthesizer and purified as described elsewhere[3]. Rink amide resin (Novabiochem, San Diego, CA) was used for peptoids to generate the C-terminal amide. Sasrin resin coupled with 2,2-dimethyl-1,3-dioxolane-4-methaneamine and cysteamine 2-chlorotrityl resin (Novabiochem, San Diego, CA) were used for peptoids to generate the C-terminal aldehyde and sufhydryl group, respectively.

Fmoc group on Rink amide resin was deprotected with 20% piperidine in DMF before starting the submonomer cycle. Peptoid systhesis on resin was as follows: a 1.2M solution of bromoacetic acid in DMF (1.13 mL in DMF, 1.35 mmol for 50 μmol of resin) and 0.93 eq. of N, N'-diisopropylcarbodiimide (DIC) (0.20 ml, 1.25 mmol for 50 μmol of resin) was added to a resin-bound amine and mixed for 10 min at 35°C during the acylation step of the submonomer cycle. The resin-bound bromide was then displaced with the amine submonomer by adding a 1.5M or 2M solution of the amine (0.85 ml for 50 μmol of resin) in N-methylpyrrolidinone (NMP). This displacement reaction was carried out for 40 min, 60 min, or 90 min at 35°C, depending on the primary amine and the length of the peptoid. In cases when Nnp and Naae were introduced at the first C-terminal residue, the displacement step was extended to 240min for efficient coupling.

When alaninamide glycine derivative submonomers (Nsahe, Nsacm and Nsaae) were used in the peptoid, the molarity of bromoacetic acid in DMF was reduced to 0.4M (0.45 mmol for 50 μmol of resin) with 67 μl of DIC (0.45 mmol) after 7 submonomers were added. In the displacement reaction, the incubation time of the first 10 displacement steps was 60 min, and the remaining 5 displacement steps were extended to 90 min.

With some peptoids, aminohexanoic acid was introduced to give some flexibility for the peptoid ligation region. 0.4M of Fmoc-εAhx-OH (Novabiochem, San Diego, CA) (0.8 mmol) in 2mL of DMF was added to the resin-bound amine with 2 ml of 0.4M hydroxybenzotriazole in DMF and 137 μl of DIC. The reaction mixture was incubated at 35°C for 120min. The Fmoc group was then deprotected with 20% piperidine in DMF for further peptoid synthesis. For the N-terminal modification, aminooxyacetamide and 2-nitrophenylsulfenylmercaptoacetamide were incorporated into the peptoid N-terminus as previously described[5].

Peptoid oligomers were cleaved from the resin with 95:5 trifluoroacetic acid (TFA)/water (v/v) for 50 min at room temperature with gentle stirring. The cleavage solution was filtered and evaporated under a stream of nitrogen to remove the TFA. The crude peptoid product was dissolved in 20% acetonitrile/water and was subjected to further purification by reverse-phase HPLC with a Vydac C4 column (10μm, 22 mm x 250 mm). To generate the C-terminal aldehyde, the crude peptoid after cleavage was treated with 10 ml of 2 mM sodium periodate in 5:1 50 mM sodium phosphate buffer (pH 7.0):acetonitrile for 30 min. The reaction mixture was injected onto an HPLC to stop the reaction and purify the aldehyde product. To generate the C-terminal sulfhydryl group, the peptoid was cleaved from cysteamine 2-chlorotrityl resin with TFA and subjected to HPLC purification.

All reaction mixtures and final products were analyzed by analytical reverse-phase HPLC (5–95% gradient at 0.8 mL/min over 30min at 60°C with a C4 Duragel G C4, 5μm, 50 x 2 mm column) and mass spectrometry (Hewlett Packard Series 1100).

Ligation between peptoids. Purified peptoid units from HPLC were used directly for the ligation of peptoids. For the aldehyde-aminooxyacetamide ligation, peptoids were mixed in a 1:1 mixture with 0.1M sodium acetate (pH 4.7). The reaction mixture was incubated for 1 hour at room temperature and then subjected to purification by HPLC. For the disulfide ligation, the sulfhydryl and activated disulfide (2-nitrophenylsulfenylmercaptoacetamide) peptoids were mixed in a 1:1 mixture with 0.1M sodium phosphate (pH 7.0). The reaction mixture was incubated for 1 hour at room temperature and then subjected to purification by HPLC. Final peptoid oligomers were lyophilized, dissolved either in water or buffer (20mM sodium phosphate, pH 7.0) and stored at -70°C. Approximate peptoid concentrations were determined using the extinction coefficients of anthranilic acid (ε=2000 $M^{-1}cm^{-1}$ at 315 nm) and 2-nitrophenol (ε=3500 $M^{-1}cm^{-1}$ at 424 nm) when they are incorporated into peptoids, otherwise peptoids were quantified by weight after lyophilization.

Results and Discussion

The four helices were assembled into a single sequence by first synthesizing four individual helical units, each with terminal ligation groups (Fig. 1, Fig. 2). The N-terminal helix $\mathbf{U^1}$ contained a C-terminal sulfhydryl group and was joined to $\mathbf{U^2}$, which contained an N-terminal disulfide and a C-terminal aldehyde. Similarly, $\mathbf{U^3}$, which contained an N-terminal amino-oxy group and a C-terminal sulfhydryl, was joined to $\mathbf{U^4}$, which contained an activated disulfide. These disulfide products were generated in ~60% and were purified to >95% purity by HPLC (Fig. 3). Then the final 4-helix assembly (Fig. 1b, Fig.2) was prepared by oxime formation between $\mathbf{U^1\_U^2}$ and $\mathbf{U^3\_U^4}$, which proceeded in ~40% yield, and was further purified by HPLC to >90% purity (Fig. 3).

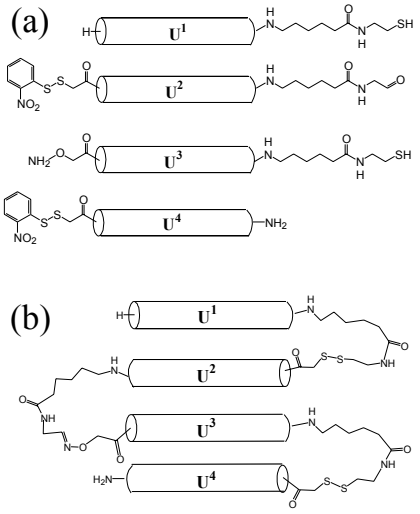

Figure 1. Individual helices were ligated together using disulfide and oxime linkages.

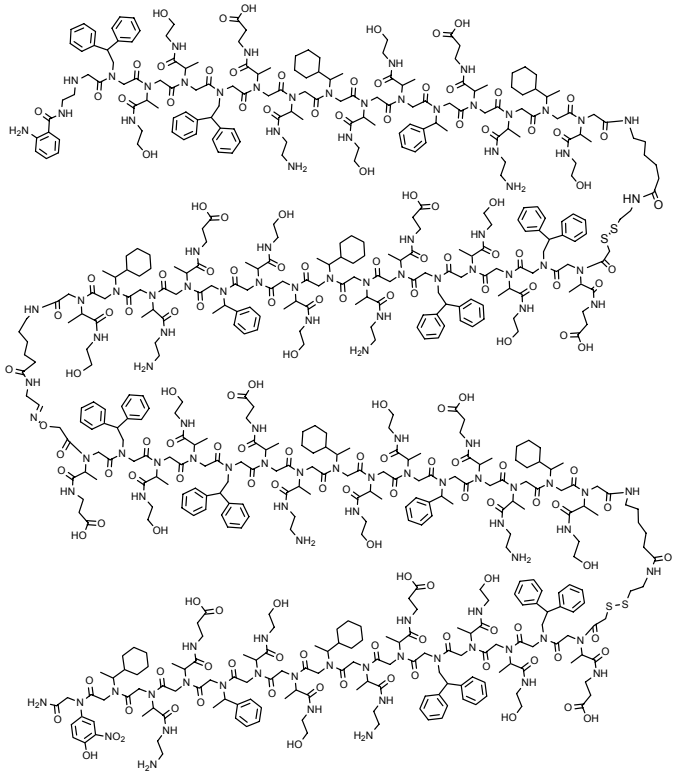

Figure 2. Structure of the 4-helix peptoid containing 60 sequence-specific monomers and 3 linkers.

After purification, the 4-helix peptoid $\mathbf{U^1\_U^2\_U^3\_U^4}$ (Fig. 2) was analyzed by mass spectrometry (Fig. 4). The observed mass 11906.2 daltons agreed with the calculated mass of 11906.8 daltons.

Conclusions

We have demonstrated that non-natural single chain sequence-specific heteropolymers the length of short proteins are synthetically accessible. We demonstrated the synthesis of a peptoid 60mer, comprised of 4 helical units linked together via 3 linker regions. The ability to synthesize molecules of this type is a significant advance toward the synthesis and screening of functional protein mimetics.

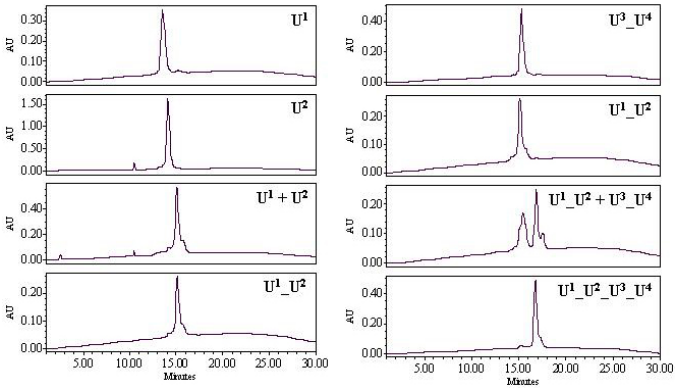

Figure 3. HPLC analysis of the assembly of a 4-helix peptoid conjugate from individual helical 15mer units.

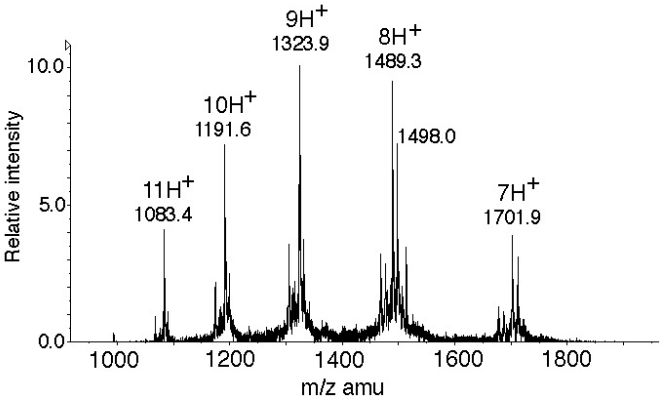

Figure 4. Mass spectrometry analysis of the 4-helix single-chain peptoid. Obsverved mass: 11906.2, calculated: 11906.8.

Acknowledgements. The authors of this paper would like to thank Deb Charych, Michael Connolly and Thomas Horn for valuable assistance.

References

(1) Hill, D. J.; Mio, M. J.; Prince, R. B.; Hughes, T. S.; Moore, J. S. *Chem. Rev.* **2001**, *101*, 3893.
(2) Patch, J. A.; Kirshenbaum, K.; Seurynck, S. L.; Zuckermann, R. N.; Barron, A. E. In *Pseudo-Peptides in Drug Discovery*; Nielsen, P. E., Ed.; Wiley-VCH: Weinheim, 2004, p 1.
(3) Burkoth, T. S.; Fafarman, A. T.; Charych, D. H.; Connolly, M. D.; Zuckermann, R. N. *J. Am. Chem. Soc.* **2003**, *125*, 8841.
(4) Dawson, P. E.; Kent, S. B. H. *Ann. Rev. Biochem.* **2000**, *69*, 923.
(5) Horn, T.; Lee, B.-C.; Dill, K. A.; Zuckermann, R. N. *Bioconj. Chem.* **2004**, *15*, 428.
(6) Burkoth, T. S.; Beausoleil, E.; Kaur, S.; Tang, D.; Cohen, F. E.; Zuckermann, R. N. *Chemistry & Biology* **2002**, *9*, 647.
(7) Zuckermann, R. N.; Martin, E. J.; Spellmeyer, D. C.; Stauber, G. B.; Shoemaker, K. R.; Kerr, J. M.; Figliozzi, G. M.; Goff, D. A.; Siani, M. A.; Simon, R. J.; Banville, S. C.; Brown, E. G.; Wang, L.; Richter, L. S.; Moos, W. H. *J. Med. Chem.* **1994**, *37*, 2678.