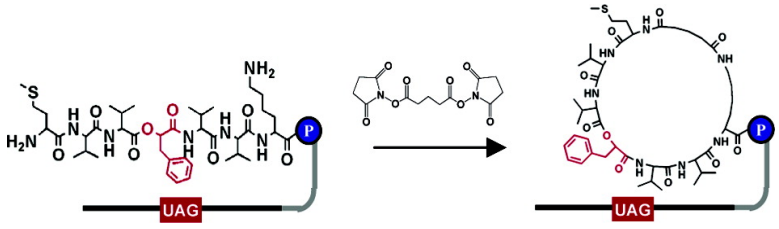
Communication

A General Route for Post-Translational Cyclization of mRNA Display Libraries

Steven W. Millward, Terry T. Takahashi, and Richard W. Roberts

J. Am. Chem. Soc., **2005**, 127 (41), 14142-14143• DOI: 10.1021/ja054373h • Publication Date (Web): 27 September 2005



More About This Article

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Links to the 3 articles that cite this article, as of the time of this article download
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

View the Full Text HTML

J|A|C|S
COMMUNICATIONS

Published on Web 09/27/2005

A General Route for Post-Translational Cyclization of mRNA Display Libraries

Steven W. Millward, Terry T. Takahashi, and Richard W. Roberts*

Division of Chemistry and Chemical Engineering, California Institute of Technology, Pasadena, California 91125

Received July 1, 2005; E-mail: rroberts@its.caltech.edu

Cyclic peptide scaffolds provide excellent starting points for the combinatorial design of high affinity ligands. Natural and synthetic cyclic peptides, including the commercially developed octreotide (Sandostatin),[1] cyclosporin A (Sandimmune),[2,3] eptifibatide (Integrilin),[4,5] and Vancomycin,[6] exhibit conformational rigidity and potent therapeutic effects. Cyclization typically results in a 10−1000-fold increase in affinity relative to a corresponding linear sequence.[7] Biological display libraries (e.g., phage display) enable natural peptides to be cyclized via a disulfide bond.[8−10] While these libraries have high sequence diversity, the disulfide bonds used to cyclize the peptide are substantially less stable than the amide backbone and are reduced inside the cell.[11] Chemical libraries (e.g., one-bead−one-compound[12]) can incorporate unnatural residues and alternate cyclization chemistries,[13,14] but they are limited in complexity and cannot be evolved through iterative selection. A general strategy to create chemically diverse, evolvable, cyclic peptide libraries would, therefore, be a powerful tool for the design of novel, conformationally constrained molecular therapeutics.

mRNA display[15] is a combinatorial platform combining high complexity ($>10^{13}$ individual sequences), evolvability, and expanded chemical complexity well beyond the 20 natural amino acids.[16] For example, unnatural chemical functionalities can be incorporated into mRNA display libraries co-translationally[17−19] or post-translationally by chemical modification.[20] Here, we demonstrate that high-diversity mRNA display libraries can be covalently cyclized via post-translational cross-linking of the N-terminal amine and the ϵ-amine of a lysine side chain.

We first explored cyclization using the Phe(K) template to generate the model peptide fusion shown in Figure 1A. Following purification, the reactive amino groups were cross-linked at pH = 8 with disuccinimidyl glutarate (DSG). Under these conditions, NHS esters react quantitatively with the model peptide fusion (Supporting Information). The RNA was removed, and the peptide−DNA conjugate was purified and analyzed by MALDI-TOF MS (Figure 1B). The DSG-reacted fusion material (red) shows one major peak in this region of the mass spectrum ($[M + H]^+ = 9497.75$) corresponding to the cyclic product ($\Delta m_{observed} = 96.67$ Da versus $\Delta m_{predicted} = 96.02$ Da, relative to the linear starting material). This number is clearly distinguishable from the mass shift predicted for the noncyclic monoacylated product ($\Delta m = 114.03$ Da) or the bismodified product ($\Delta m = 228.06$ Da). Our data also confirm that the peptide component of RNA−peptide fusions can include all of the amino acids encoded in the template.[15]

Having confirmed the cyclization reaction in the model fusion by mass spectrometry, we next sought to quantify its efficiency. It was previously shown that the incorporation of an ester linkage into disulfide loops could be used to map disulfide bond connectivity by NH_4OH hydrolysis.[21] Following this strategy, we constructed a variant of the Phe(K) template containing a UAG amber codon at position 4 (Figure 2A). The resulting UAG(K) template was suppressed with THG73 amber suppressor tRNA[22] bearing α-hydroxy phenylalanine (AhF), resulting in the incorporation of a base-labile ester linkage.[23] The UAG(V) template, which lacks a reactive lysine, was constructed as a negative control. Following fusion formation, ^{35}S-labeled fusions were reacted with DSG as described above, followed by RNase treatment and NH_4OH cleavage of the ester. The resulting products were separated by urea−PAGE and analyzed by autoradiography (Figure 2B).

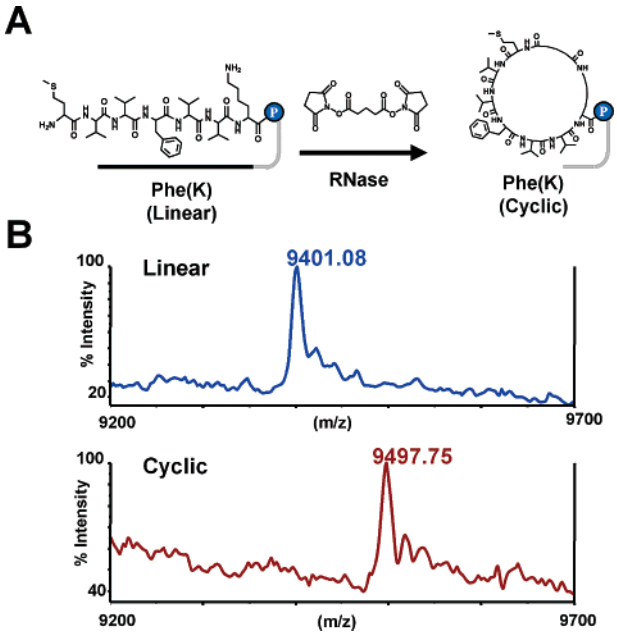


Figure 1. (A) Cyclization of mRNA display libraries. RNA (black), DNA linker (gray), puromycin (blue). (B) MALDI-TOF MS of Phe(K) (blue) and DSG-reacted Phe(K) (red).

In the absence of a cross-linker, fusions treated with NH_4OH will be cleaved and the N-terminal tripeptide bearing the radiolabel will be lost, while fusions that are cyclized will retain the radio-labeled tripeptide after ester hydrolysis. UAG(V) showed no detectable cyclization, demonstrating the selectivity of the cross-linking chemistry. However, 55% of the UAG(K) fusion band intensity remained following DSG reaction and hydrolysis, corresponding to the cyclization efficiency of the model fusion.

Finally, we used this methodology to construct and analyze highly diverse cyclic libraries. Five libraries were assembled (MK2−MK10) having 2, 4, 6, 8, and 10 random positions (X) between the N-terminal methionine and the C-terminal constant region (Figure 2C). Cyclization via the constant lysine will generate macrocycles with 24, 30, 36, 42, and 48 atoms, respectively. In these libraries, cyclization can be followed using AhF incorporation as described above. As can be seen, the cyclization efficiencies range from 55 to 31% and gradually decrease as the number of intervening residues increases. A similar trend has also been reported in disulfide-mediated cyclizations.[24]

10.1021/ja054373h CCC: $30.25

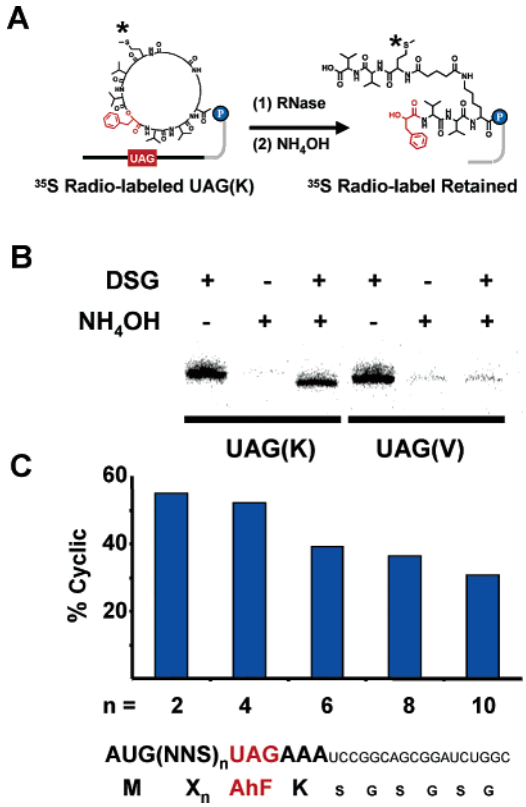


Figure 2. (A) The UAG(K) template contains the UAG amber codon which directs the incorporation of AhF (red). Cyclization prevents the loss of the radio-label after treatment with NH_4OH. (B) Urea−PAGE of hydrolyzed reaction products. (C) Cyclization of MK libraries. The efficiency of cyclization was determined by the extent of radio-label retention following NH_4OH hydrolysis. The bars represent the average of two experimental values (see Supporting Information).

Our macrocycle libraries rely on amine-based cross-linking. In the size range we have examined, the majority of cycles will be bridged between the fixed lysine and the N-terminus, as in Figure 1A. Lysine residues within the NNS cassettes also enable two different types of cycles: (i) those with linkages between a cassette side chain and the fixed lysine, and (ii) those where a cassette lysine is linked to the N-terminus. The second product will not be detected in our cyclization assay (Figure 2C), and therefore, the overall cyclization efficiency we report does not include this contribution (see the Supporting Information for additional discussion).

The cross-linking chemistry we used is efficient, mild, peptide-specific, and compatible with almost any bifunctional NHS cross-linker, affording a facile route for the introduction of chemical diversity incompatible with the translation machinery. The incorporation of α-hydroxy acids gives a general, quantitative method to evaluate any cross-linking or cross-coupling methodologies compatible with mRNA−peptide fusion stability. Additionally, MALDI-TOF MS was found to be a powerful analytical tool for the analysis of mRNA−peptide fusions and their reaction products.

To our knowledge, the MK10 library represents the most diverse cyclic library assembled in a biological display format ($>10^{13}$ sequences) as well as the largest to contain an unnatural amino acid (AhF). We believe these complex, conformationally constrained libraries have enormous potential to generate high affinity ligands to any number of biologically relevant targets.

Acknowledgment. The authors would like to thank Niki Zacharias at Neurion Pharmaceuticals for the generous gift of pdCpA and α-hydroxy phenylalanine-dCA. This work was supported by NIH R01 GM60416, NIH Training Grant T32 GM08501, and the California Institute of Technology.

Supporting Information Available: Experimental procedures and additional data. This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) Pawlikowski, M.; Melen-Mucha, G. *Curr. Opin. Pharmacol.* **2004**, *4*, 608−613.
(2) Fliri, H.; Baumann, G.; Enz, A.; Kallen, J.; Luyten, M.; Mikol, V.; Movva, R.; Quesniaux, V.; Schrier, M.; Walkinshaw, M.; Wenger, R.; Zenke, G.; Zurini, M. *Ann. N.Y. Acad. Sci.* **1993**, *696*, 47−53.
(3) Jin, L.; Harrison, S. C. *Proc. Natl. Acad. Sci. U.S.A.* **2002**, *99*, 13522−13526.
(4) Scarborough, R. M.; Naughton, M. A.; Teng, W.; Rose, J. W.; Phillips, D. R.; Nannizzi, L.; Arfsten, A.; Campbell, A. M.; Charo, I. F. *J. Biol. Chem.* **1993**, *268*, 1066−1073.
(5) Scarborough, R. M.; Gretler, D. D. *J. Med. Chem.* **2000**, *43*, 3453−3473.
(6) Hubbard, B. K.; Walsh, C. T. *Angew. Chem., Int. Ed.* **2003**, *42*, 730−765.
(7) Khan, A. R.; Parrish, J. C.; Fraser, M. E.; Smith, W. W.; Bartlett, P. A.; James, M. N. G. *Biochemistry* **1998**, *37*, 16839−16845.
(8) O'Neil, K. T.; Hoess, R. H.; Jackson, S. A.; Ramachandran, S.; Mousa, S. A.; DeGrado, W. F. *Proteins: Struct., Funct., Genet.* **1992**, *14*, 509−515.
(9) Katz, B. A. *Biochemistry* **1995**, *34*, 15421−15429.
(10) Giebel, L. B.; Cass, R. T.; Milligan, D. L.; Young, D. C.; Arze, R.; Johnson, C. R. *Biochemistry* **1995**, *34*, 15430−15435.
(11) Gilbert, H. F. *Methods Enzymol.* **1995**, *251*, 8−28.
(12) Lam, K. S.; Salmon, S. E.; Hersh, E. M.; Hruby, V. J.; Kazmierski, W. M.; Knapp, R. J. *Nature* **1991**, *354*, 82−84.
(13) Humphrey, J. M.; Chamberlin, R. *Chem. Rev.* **1997**, *97*, 2243−2266.
(14) Davies, J. S. *J. Pept. Sci.* **2003**, *9*, 471−501.
(15) Roberts, R. W.; Szostak, J. W. *Proc. Natl. Acad. Sci. U.S.A.* **1997** *94*, 12297−12302.
(16) Frankel, A.; Li, S.; Starck, S. R.; Roberts, R. W. *Curr. Opin. Struct. Biol.* **2003**, *13*, 506−512.
(17) Li, S.; Millward, S.; Roberts, R. *J. Am. Chem. Soc.* **2002**, *124*, 972−9973.
(18) Frankel, A.; Roberts, R. W. *RNA* **2003**, *9*, 780−786.
(19) Frankel, A.; Millward, S. W.; Roberts, R. W. *Chem. Biol.* **2003**, *10*, 1043−1050.
(20) Li, S.; Roberts, R. W. *Chem. Biol.* **2003**, *10*, 233−239.
(21) England, P. M.; Lester, H. A.; Dougherty, D. A. *Biochemistry* **1999**, *38*, 14409−14415.
(22) Saks, M. E.; Sampson, J. R.; Nowak, M. W.; Kearney, P. C.; Du, F.; Abelson, J. N.; Lester, H. A.; Dougherty, D. A. *J. Biol. Chem.* **1996**, *271*, 23169−23175.
(23) England, P. M.; Lester, H. A.; Dougherty, D. A. *Tetrahedron Lett.* **1999**, *40*, 6189−6192.
(24) Zhang, R.; Snyder, G. H. *Biochemistry* **1991**, *30*, 11343−11348.

JA054373H