

See discussions, stats, and author profiles for this publication at: <https://www.researchgate.net/publication/7546705>

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

ARTICLE *in* JOURNAL OF THE AMERICAN CHEMICAL SOCIETY · OCTOBER 2005

Impact Factor: 12.11 · DOI: 10.1021/ja054373h · Source: PubMed

---

CITATIONS

46

---

READS

54

3 AUTHORS, INCLUDING:



[Steven W Millward](#)

University of Texas MD Anderson Cancer Center

13 PUBLICATIONS 295 CITATIONS

SEE PROFILE

## 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

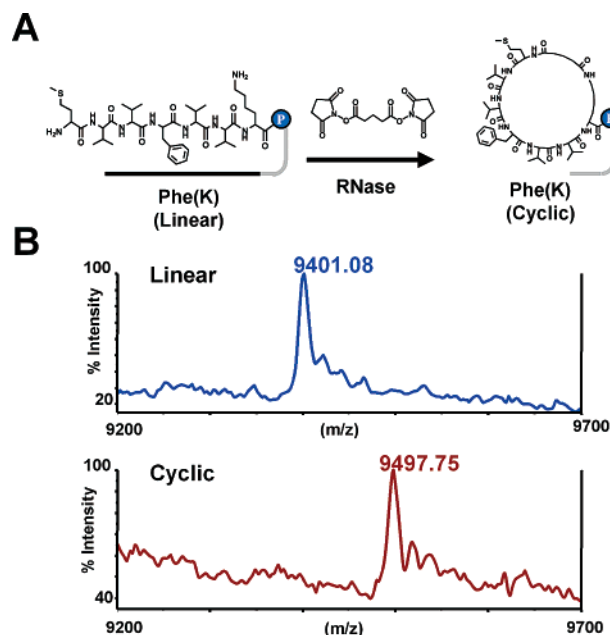
Received July 1, 2005; E-mail: rroberts@its.caltech.edu

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),<sup>1</sup> cyclosporin A (Sandimmune),<sup>2,3</sup> eptifibatide (Integrilin),<sup>4,5</sup> and Vancomycin,<sup>6</sup> exhibit conformational rigidity and potent therapeutic effects. Cyclization typically results in a 10–1000-fold increase in affinity relative to a corresponding linear sequence.<sup>7</sup> Biological display libraries (e.g., phage display) enable natural peptides to be cyclized via a disulfide bond.<sup>8–10</sup> 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.<sup>11</sup> Chemical libraries (e.g., one-bead–one-compound<sup>12</sup>) can incorporate unnatural residues and alternate cyclization chemistries,<sup>13,14</sup> 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<sup>15</sup> is a combinatorial platform combining high complexity ( $>10^{13}$  individual sequences), evolvability, and expanded chemical complexity well beyond the 20 natural amino acids.<sup>16</sup> For example, unnatural chemical functionalities can be incorporated into mRNA display libraries co-translationally<sup>17–19</sup> or post-translationally by chemical modification.<sup>20</sup> 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  $\epsilon$ -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 ( $\Delta m_{\text{observed}} = 96.67$  Da versus  $\Delta m_{\text{predicted}} = 96.02$  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.<sup>15</sup>

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  $\text{NH}_4\text{OH}$  hydrolysis.<sup>21</sup> 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<sup>22</sup> bearing  $\alpha$ -hy-

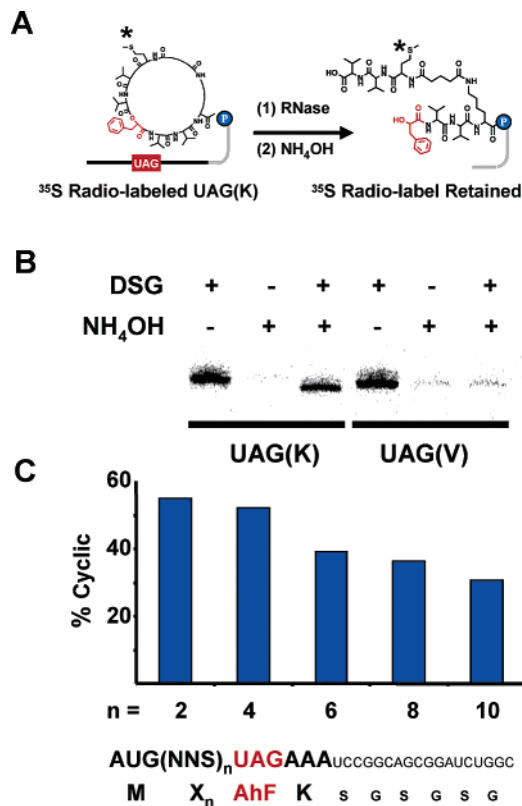


**Figure 1.** (A) Cyclization of mRNA display libraries. RNA (black), DNA linker (gray), puromycin (blue). (B) MALDI-TOF MS of Phe(K) (blue) and DSG-reacted Phe(K) (red).

droxy phenylalanine (AhF), resulting in the incorporation of a base-labile ester linkage.<sup>23</sup> The UAG(V) template, which lacks a reactive lysine, was constructed as a negative control. Following fusion formation, <sup>35</sup>S-labeled fusions were reacted with DSG as described above, followed by RNase treatment and  $\text{NH}_4\text{OH}$  cleavage of the ester. The resulting products were separated by urea–PAGE and analyzed by autoradiography (Figure 2B).

In the absence of a cross-linker, fusions treated with  $\text{NH}_4\text{OH}$  will be cleaved and the N-terminal tripeptide bearing the radio-label 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.<sup>24</sup>



**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<sub>4</sub>OH. (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<sub>4</sub>OH 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  $\alpha$ -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  $\alpha$ -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

## Supporting Information

### Synthesis of Aminoacylated tRNA

The pdCpA dinucleotide (tetrabutylammonium salt) and a sample of  $\alpha$ OH-Phenylalanyl-dCA were obtained as a gift from Neurion Pharmaceuticals (Pasadena, CA). Subsequent preparation of  $\alpha$ OH-Phenylalanyl-dCA was carried out according to the protocols in references 17-18. An example synthesis is described below.

### Synthesis of $\alpha$ OH-Phenylalanine cyanomethyl ester (1)

L-phenyllactic acid (266 mg, 1.6 mmol) was dissolved in 3 mL DMF. To this was added chloroacetonitrile (3 mL, 47.4 mmol) and TEA (651  $\mu$ L, 4.6 mmol). The reaction was allowed to proceed under nitrogen at room temperature overnight. The desired product was purified by flash chromatography (silica gel, 3:7 EtOAc:Hexanes). The final yield (amber oil) was 18.9 mg (68%). Analysis by ESI-MS (Expected  $[M+Na]^+=228.07$ , Observed  $[M+Na]^+=227.8$ )

### Synthesis of $\alpha$ OH-Phenylalanine-dCA (2)

$\alpha$ OH-Phenylalanine cyanomethyl ester (11 mg, 54  $\mu$ mol) was dissolved in 400  $\mu$ L DMF and added to the tetrabutyl ammonium salt of pdCpA (9  $\mu$ mol) in the presence of a catalytic amount of TBA-acetate. The reaction was allowed to proceed under nitrogen for 4 hr. at room temperature. The aminoacylated dinucleotide was purified by RP-HPLC using a gradient from 25 mM  $NH_4OAc$  (pH=4.5) to  $CH_3CN$ . Following lyophilization of the pooled fractions, the product was dissolved in 10 mM HOAc and lyophilized again. The final yield was determined by absorbance at 260 nm. and found to be 97 nmol (1%). The product was analyzed by ESI-MS (Expected  $[M-H]^-=783.16$ , Observed  $[M-H]^-=783.2$ ).

### *In vitro* Transcription of THG73 tRNA

The plasmid harboring the THG73 gene was linearized with FokI and transcribed with T7 RNA polymerase. The transcription product was gel-purified by Urea-PAGE, dissolved in  $dH_2O$ , and quantitated by absorbance at 260 nm.

### Ligation to THG73 tRNA

20  $\mu$ g THG73 tRNA (8  $\mu$ L in  $dH_2O$ ) in HEPES (22  $\mu$ L, 10 mM, pH=7.5) was heated to 94  $^{\circ}C$  for 3 min and allowed to cool slowly at room temperature. 8  $\mu$ L  $\alpha$ OH-Phenylalanine-dCA (3 mM in DMSO), 32  $\mu$ L 2.5 X Reaction Buffer (25  $\mu$ L 400 mM HEPES pH=7.5, 10  $\mu$ L 100 mM DTT, 25  $\mu$ L 200 mM  $MgCl_2$ , 3.75  $\mu$ L 10 mM ATP, 10  $\mu$ L

5 mg/mL BSA, 26.25  $\mu$ L dH<sub>2</sub>O), 7  $\mu$ L water, and 4  $\mu$ L T4 RNA Ligase (N.E.B.). After incubation at 37 °C for 1 hr, the reaction was extracted once with an equal volume of phenol (saturated with 300 mM NaOAc, pH=5.0:CHCl<sub>3</sub>:isoamyl alcohol (25:24:1)) and ethanol precipitated. The pellet was washed with 70% EtOH, dried under vacuum, and dissolved in 1 mM NaOAc to a concentration of 2  $\mu$ g/ $\mu$ L.

### Construction of Fusion Templates (Phe(K), UAG(K), UAG(V))

These dsDNA templates were created by PCR using overlapping primers: (Forward primer = **Gen-FP**: 5'-TAATACGACTCACTATAGGGACAATTACTATTTACAATTACA-3') and a unique reverse primer (**Phe(K)-RP**: 5'-CTTAACAACGAAAACAACCATTGTAATTGTAAATAGTAATTG-3', **UAG(K)-RP**: 5'-CTTAACAACCTAAACAACCATTGTAATTGTAAATAGTAATTG-3', **UAG(V)-RP**: 5'...AACAAACCTAAACAACCATTGTAATTGTAAATAGTAATTG...3'). Prior to PCR, all primers were purified by Urea-PAGE. Following PCR, the dsDNA templates were transcribed into mRNA by T7 runoff transcription and purified by Urea-PAGE. The mRNA was ligated to F30P (5'-dA<sub>21</sub>[C<sub>9</sub>]<sub>3</sub>dAdCdC-P; C<sub>9</sub>=triethylene glycol phosphate (Glen Research), P= puromycin (Glen Research)), a flexible DNA linker containing puromycin. 5' phosphorylated F30P was ligated to the mRNA in the presence of an oligonucleotide splint (**Phe(K)-splint**: 5'-TTTTTTTTTTTTTCTTAACAACGAAA-3', **UAG(K)-splint**: 5'-TTTTTTTTTTTTTCTTAACAACCTAA-3', **UAG(V)-splint**: 5'...TTTTTTTTTTTTTAACAACAACCTAA...3'). The ratio of RNA:PF30P:splint was 1:1.2:1.4. 4  $\mu$ L of T4 DNA Ligase (N.E.B.) was added and the reaction was carried out at room temperature for 1 hr. The mRNA-DNA template was gel purified by Urea-PAGE, dissolved in water, and quantitated by absorbance at 260 nm.

### Construction of MK(n) Library Templates (MK0, MK2, MK4, MK6, MK8, MK10)

Antisense ssDNA templates were synthesized at the Keck Oligonucleotide Synthesis Facility (Yale). The sequences for the MK Library ssDNA templates are as follows:

**MK(2)**: 5'-TCCGCTGCCGGATTCTASNNNSNNCATTGTAATTGTAAATAGTAATTG-3'  
**MK(4)**: 5'-TCCGCTGCCGGATTCTASNNNSNNNSNNCATTGTAATTGTAAATAGTAATTG-3'  
**MK(6)**: 5'-TCCGCTGCCGGATTCTASNNNSNNNSNNNSNNCATTGTAATTGTAAATAGTAATTG-3'  
**MK(8)**: 5'-TCCGCTGCCGGATTCTASNNNSNNNSNNNSNNNSNNCATTGTAATTGTAAATAGTAATTG-3'  
**MK(10)**: 5'-TCCGCTGCCGGATTCTASNNNSNNNSNNNSNNNSNNNSNNCATTGTAATTGTAAATAGTAATTG-3'

N=A,T,C,G S=G,C

The reagent bottle used for the “N” positions was made by mixing A:C:G:T in the ratio 3:3:2:2. The reagent bottle

for “S” positions was made by mixing C:G in a 3:2 ratio.

MK library dsDNA was amplified by PCR using the forward primer Gen-FP (5'-TAATACGACTCACTATAGGGACAATTACTATTTACAATTACA-3') and the reverse primer MKLib-RP (5'-GCCAGATCCGCTGCCGGATTTCTA-3'). The mRNA was made by *in vitro* T7 runoff transcription as described above. Following PAGE purification, the mRNA was ligated to F30P via. an oligonucleotide splint (**MKLib-splint**: 5'-TTTTTTTTTTTTTGCCAGATCCGCTG-3'). Following PAGE purification of the ligation reaction, the templates were dissolved in water and quantitated by absorbance at 260 nm.

### **Translation of mRNA-peptide fusions**

Templates were translated in rabbit reticulocyte lysate under standard conditions with a final template concentration of 400 nM. For templates with UAG amber codons (UAG(K), UAG(V), and the MK Libraries), the reaction mixture was supplemented with 2 µg αOH-phenylalanyl-tRNA/25 µL translation volume. <sup>35</sup>S-Methionine was used as the labeling agent. After 1 hr of translation at 30 °C, KOAc and MgCl<sub>2</sub> were added to a final concentration of 600 mM and 50 mM respectively, and the reactions were placed at -20 °C overnight.

### **Fusion Purification**

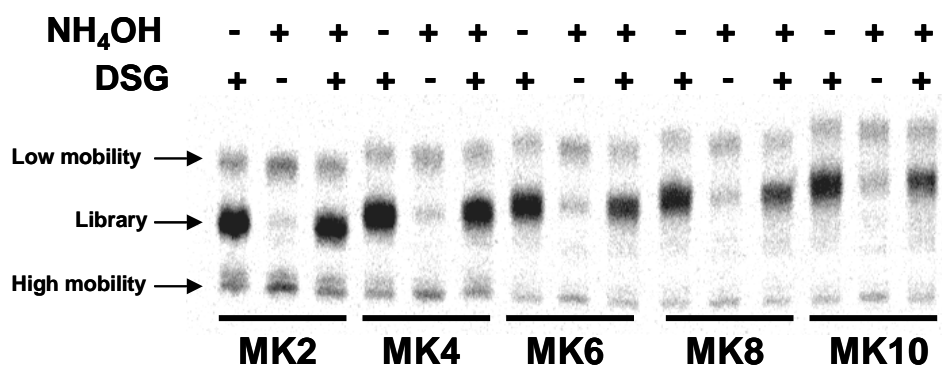
The translation mixture was diluted 1:10 in dT Binding Buffer (10 mg/mL dT cellulose, 1M NaCl, 20 mM Tris, 1 mM EDTA, 0.2% Triton X-100, pH=8) and agitated for 1 hr. at 4 °C. The dT cellulose was filtered and washed with dT Wash Buffer (300 mM NaCl, 20 mM Tris, pH=8). The DNA-peptide conjugates were eluted with 10 mM NH<sub>4</sub>OH and ethanol precipitated in the presence of linear acrylamide (Ambion).

### **Reaction with DSG and Determination of Cyclization Efficiency by NH<sub>4</sub>OH Hydrolysis**

Purified fusions (dissolved in water) were diluted 1:1 in 100 mM phosphate buffer (pH=8). To this was added 1 mg/mL DSG (in DMF) to a final concentration of 600 µM. The final concentration of fusion in these reactions was estimated to be 1-10 nM by scintillation counting of the <sup>35</sup>S radiolabel. The reaction was allowed to proceed for 1 hr at room temperature, treated with RNase Cocktail (Ambion) at room temperature for 2 hr, and diluted 1:1 in 2M NH<sub>4</sub>OH and heated at 65 °C for 30 min to hydrolyze the internal ester linkage. The fusions were treated with RNase to facilitate resolution of the reaction products by Urea-PAGE. The sample was dried *in vacuo*, co-evaporated with water, and redissolved in Urea-PAGE loading buffer (80% formamide, bromophenol blue). The

reaction products were separated on a 15% Urea-PAGE gel at 15W for 90 min. The gel was fixed, dried, and visualized by autoradiography. The fraction cyclized was determined by subtracting the band volume of the mock-reacted sample (-DSG/+NH<sub>4</sub>OH) from that of the DSG reacted sample (+DSG/+NH<sub>4</sub>OH) and dividing the result by the band volume of the control sample (+DSG/-NH<sub>4</sub>OH).

The Urea-PAGE gel shown below (Figure 1s) is representative of a typical experiment to determine the cyclization efficiencies of the MK libraries (MK2-MK10). The major band represents the full-length desired fusion product and is dependent on the addition of aminoacylated suppressor tRNA. The high and low mobility bands are observed in both the presence and absence of supplemented aminoacyl-tRNA and are not affected by base treatment. The high mobility band is most likely due to internal initiation of translation while the low mobility band may arise from translational frameshifting at the UAG stop codon.



**Figure 1s:** Urea-PAGE of MK Library hydrolysis products. Quantitation of cyclization efficiency is described above.

#### Analysis of Fusions by MALDI-TOF MS.

200 pmol of Phe(K) Template was translated in the absence of <sup>35</sup>S methionine as described above. The translation reactions were dT-purified, ethanol precipitated, and dissolved in 50 mM phosphate buffer (pH=8). These were treated with either DMF (control) or DSG for 1 hr at room temperature. The RNA was then removed by incubation with RNase cocktail (Ambion) at room temperature for 3 hr. Previous experiments showed that removal of the RNA was essential for observation of the mass signal in MALDI-TOF MS. The peptide-DNA fusions were then dT-purified, eluted with 10 mM NH<sub>4</sub>OH, and quenched with 10% HOAc. The samples were concentrated and desalted with a Nanosep 3K centrifugal filtration device (Pall). The final desalted sample was brought to a volume of 10 µL by evaporation. A portion of this sample was incubated for 30 seconds with DOWEX ion exchange resin (previously treated with NH<sub>4</sub>OAc).<sup>1</sup> Approximately 0.15 µL was applied to the matrix spot (3-hydroxypicolinic

acid) and allowed to dry. This was repeated 9 times to increase sample concentration. Methionine oxidation was accomplished by treating the dried sample with 0.25  $\mu\text{L}$  of 0.5%  $\text{H}_2\text{O}_2$  prior to analysis.<sup>2</sup> This was found to quantitatively oxidize the methionine side chain thioether to the sulfoxide form, effectively simplifying the spectra (data not shown).

All MALDI were obtained on an Applied Biosystems Voyager DE PRO MALDI-TOF mass spectrometer operating in linear positive ion mode with delayed extraction. The accelerating voltage was set at 25000 V, grid voltage at 92.5%, and guide wire at 0.15%. The extraction delay time was 325 nsec. A five point external protein calibration (matrix: sinapic acid) was used. The laser power was adjusted to produce the strongest mass signal for the peptide-DNA conjugates. A table of predicted and observed masses for the linear and DSG-reacted fusions as well as the F30P linker is shown below:

<b>Species</b>	<b>Predicted [M+H]<sup>+</sup></b>	<b>Observed [M+H]<sup>+</sup></b>
<b>F30P</b>	8573.72	8582.2
<b>F30P(mono-acylated)</b>	8687.75	8697.27
<b>Phe(K)-F30P(linear)</b>	9392.19	9401.08
<b>Phe(K)-F30P(cyclic)</b>	9488.21	9497.75

**Table 1s:** Predicted and Observed [M+H]<sup>+</sup> values for fusion species described in this experiment. Predicted values represent exact mass calculations (ChemDraw).

### Alternate Cyclization Products

The construction of the libraries described in this manuscript allows the incorporation of lysine in the random region via the AAG codon. The chance of lysine insertion at any of the randomized positions is therefore approximately 1/32. Thus, for the n = 2, 4, 6, 8, and 10 libraries, the fraction of sequences in the library containing one or more internal lysines is 0.062, 0.12, 0.19, 0.25, and 0.31 respectively. Thus, between ~94% (n=2) and ~70% (n=10) of the library members have the potential to form only the desired macrocyclic product.

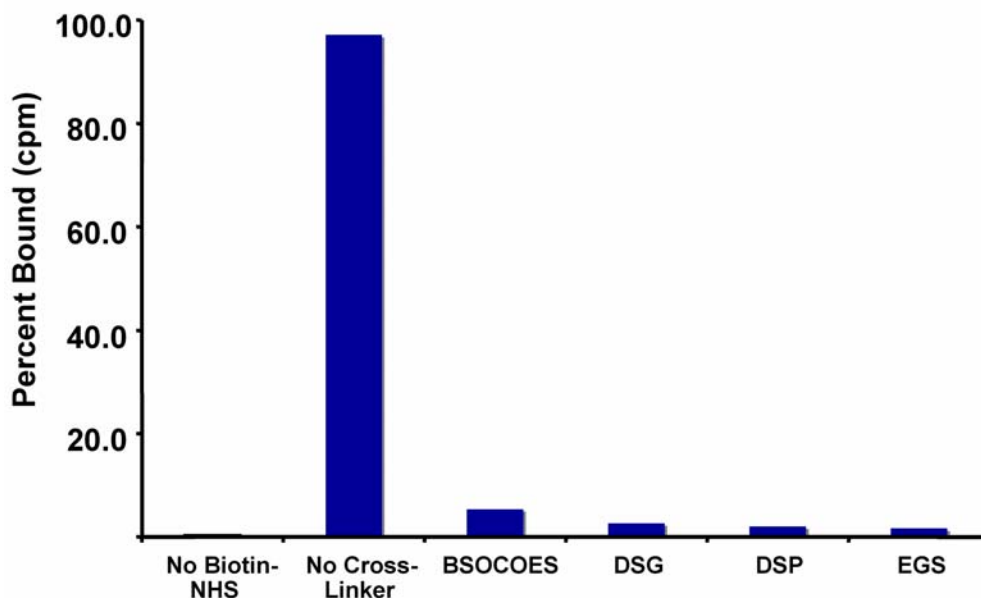
The majority of the remaining library members will have one internal lysine encoded in the randomized cassette region. Library members with one internal lysine have the potential to form 3 types of cyclization products: 1) the intended cyclization product (N-terminus-fixed lysine), 2) fixed lysine-cassette lysine, and 3) cassette lysine-N-terminus. For these molecules, one template sequence corresponds to three different cyclic products, providing



the ability to select functional molecules from each category. After a selection is completed, chemical synthesis can be used to construct each distinct cyclic product and thereby determine which is functional.

### NHS Reactivity with mRNA-peptide Fusions

Figure 2s shows that biotin-NHS reacts nearly quantitatively with the Phe(K) fusions. Reaction with different bis-NHS cross-linkers prior to reaction with biotin-NHS greatly reduces fusion binding to neutravidin agarose, implying that all NHS cross-linkers used here may be used to cyclize fusions.



**Figure 2s:** Biotin-NHS reactivity interference using different NHS-based cross-linkers. dT-purified Phe(K) fusions were first reacted with 2.5 mM cross-linkers (BSOCOES = Bis[2-(succinimidooxycarbonyloxy)ethyl]sulfone, DSG = disuccinimidyl glutarate, DSP = dithiobis[succinimidylpropionate], and EGS = ethylene glycol bis[succinimidylsuccinate]; all cross-linkers were from Pierce Chemical) or with DMSO only (No biotin-NHS and No Cross-linker) in 100 mM phosphate buffer (pH = 7.5), 10% (v/v) DMSO for 3 hours at room temperature. The reactions were then dT purified, ethanol precipitated, and reacted with biotin-NHS or with DMSO only (No Biotin-NHS) for 2 hours at room temperature. The products were again dT purified, eluted in H<sub>2</sub>O, and incubated with 50  $\mu$ L of a 50/50 (v/v) slurry of Neutravidin agarose in 1X PBST (137 mM NaCl, 2.7 mM KCl, 8 mM Na<sub>2</sub>HPO<sub>4</sub>, 2 mM KH<sub>2</sub>PO<sub>4</sub>, pH 7.4, 0.1 % (v/v) Triton X-100) for one hour at 4 °C. The agarose was washed with 1X PBST and the beads counted on a scintillation counter.

### Supplemental Information (References):

- 1) Petersson, E.J.Shahgholi, M., Lester, H.A., Dougherty, D.A. *RNA* **2002**, 8, 542-547.
- 2) Corless, S., Cramer, R. *Rapid. Comm. Mass. Spec.* **2003**, 17(11), 1212-1215.