[B0012]

An Efficient Synthetic Strategy for the Preparation of Nucleic Acid-Encoded Peptide and Protein Libraries for *In Vitro* Evolution Protocols

Markus Kurz, Guannan Kuang, and Peter A. Lohse*
Phylos, Inc., 128 Spring Street, Lexington, MA 02421

* To whom correspondence should be addressed.
Tel. +1 781 862 6400 ext. 224; Fax +1 781 402 8800; Email: plohse@phylos.com

Abstract

We describe an improved synthetic strategy for the preparation of nucleic acid encoded peptide and protein libraries. A solid-phase format was used to prepare and purify a novel type of mRNA-template for *in vitro* mRNA-protein fusion synthesis. The present protocol simplifies and accelerates the preparation of fusion libraries and should prove most useful for *in vitro* protein evolution procedures which involve repetitive cycles of fusion library preparation and selection.

Keywords

Nucleic acid-encoded libraries; mRNA-protein fusion; PROfusion^{TM}; *in vitro* selection; puromycin; psoralen.

Introduction

Combinatorial library technologies have emerged during the past 15 years as powerful tools for basic research and drug discovery and development. As first developed for peptides, the general concept of combinatorial libraries involves the generation of sequence permutations for a peptide of a given length, in connection with a screening and selection process that enables the isolation and identification of rare, functional peptides in the presence of large numbers of less active or inactive compounds. Since the individual synthesis of such large numbers of compounds ($>10^4$) is impractical, a variety of approaches have been developed to generate and screen mixtures of peptides, proteins and organic compounds. Whereas combinatorial chemistry technologies allow for easy incorporation of nonproteinogenic building blocks and are focused on the preparation of low molecular weight compound libraries for the discovery of drug candidates (1), nucleic acid-encoded libraries like phage display or yeast display are usually employed to discover novel polypeptide binding motifs or to identify protein binding partners within a cell's proteome (2). Common for the biosynthetic approaches are the iterative process of randomization, selection and amplification. While the last step is not readily available for chemical libraries, it is an important advantage of nucleic acid-encoded libraries that allows the handling of minimal amounts of material with maximum diversity, and eliminating virtually any worries about detection limits.

Until recently, the limitations of nucleic acid-encoded library technologies were largely due to necessary *in vivo* steps, which limit library size and provide selection pressure that cannot be controlled by the experimenter. A lately introduced method for the preparation of a minimal genetic unit which consists of a mRNA covalently linked to a polypeptide (PROfusion^{TM}) circumvents the *in vivo* step and allows the selection for peptides and proteins with desired properties totally *in vitro* (3-5). RNA-protein fusion libraries with up to 10^{14} different sequences have been generated and have been successfully used for the isolation of peptide and protein binders to a variety of different targets (6).

The originally described method for RNA-protein fusion preparation involved translation of an enzymatically linked mRNA-puromycin conjugate (3). Recently, we developed a simplified method for fusion synthesis from psoralen-crosslinked mRNA-puromycin templates (7). Here we report a further improvement of this method in which the mRNA-puromycin template is prepared in an integrated affinity purification and crosslink formation process on solid phase. This novel synthesis protocol further simplifies mRNA-protein fusion preparation and should prove most useful in an automated fusion synthesis process.
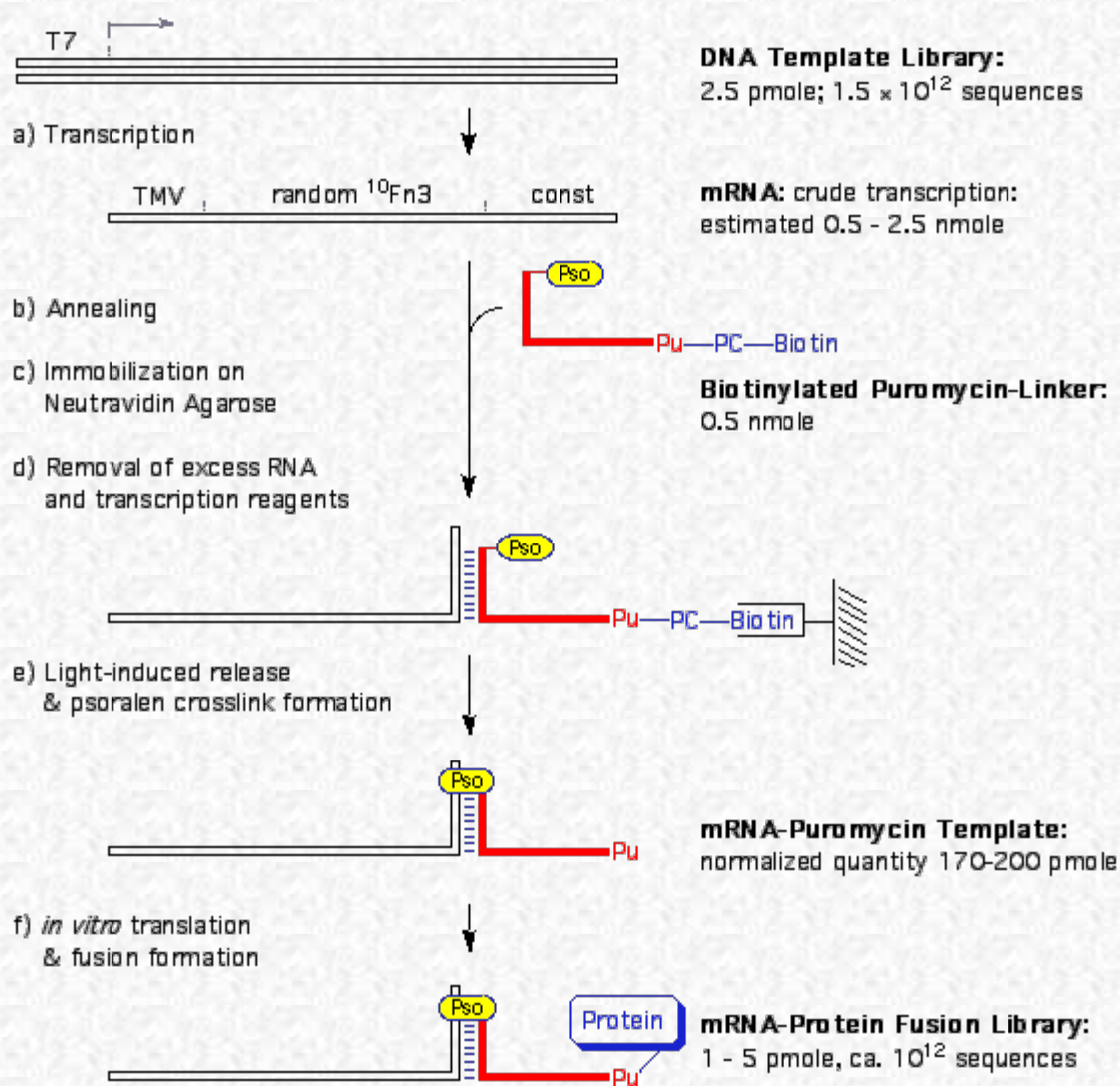

Figure 1: Schematic outline of mRNA-protein fusion synthesis. A double stranded DNA template library was *in vitro* transcribed to produce the mRNA template (a) which was annealed with a limiting amount of biotinylated puromycin-linker (b) and subsequently captured on neutravidin beads (c). After a washing step to remove unbound mRNA (d), the beads were irradiated with UV light (e). Elution yielded the photo-ligated mRNA-puromycin template which was directly subjected to *in vitro* translation to form the mRNA-protein fusions library (f). T7, transcription promoter sequence; TMV, a portion of the tobacco mosaic virus 5'-UTR with good initiation codon context; random ^{10}Fn3, a sequence coding for a randomized 10th fibronection type III domain; const, a constant sequence which codes for the linker hybridization and crosslinking sites as detailed in Fig. 2; Pu, puromycin; Pso, psoralen; PC-biotin, a photo-cleavable biotin-tag described in Fig. 2. METHODS: Freshly transcribed mRNA (0.5 — 2.5 nmol) was hybridized to biotinylated puromycin-linker (0.5 nmol, see Fig. 2) in 300 µl binding buffer (30 mM Tris, pH 7.0, 250 mM NaCl) by heating to 85° C for 30 sec followed by cooling to 4° C in 5 min. 100 µl pre-washed Neutravidin beads (Pierce) were added to the hybridization mixture and incubated for 30 min at 4° C under rocking. Subsequently, the beads were

washed 3x 100 µl binding buffer followed by centrifugation to remove the liquid phase. The moist beads were then irradiated for 15 min at room temperature with a 25W UV-lamp (Pyrex-filter, wavelength > 300nm). Subsequently, the beads were washed with 100 µl plain water to yield 170 — 200 pmol photo-crosslinked mRNA-puromycin template which was directly used for mRNA-peptide fusion formation in 300 µl rabbit reticulocyte lysate to produce the mRNA-protein fusions library according to reference (7).

Results and Discussion

The preparation of the mRNA template coding for the 10th fibronection type III domain (21 out of 94 amino acids were randomized) began with the *in vitro* transcription from a PCR-DNA library (10^{12} different sequences) according to the method described in reference (5) (Fig. 1). The freshly transcribed mRNA was annealed to a limiting amount of puromycin-linker carrying a photo-cleavable biotin group for affinity purification on solid-phase (Fig. 2). The hybridized mRNA template-linker complex was then incubated with neutravidin agarose beads followed by washing to remove excess mRNA and transcription buffer components. UV-irradiation of the beads leads to the cross-linking of the two strands by psoralen and to simultaneous cleavage of the biotin-tag. The predetermined quantity of pure mRNA-puromycin template could be eluted from the solid phase. Polyacrylamide gel analysis indicated a ligation yield of > 80 % for the mRNA-puromycin template which was directly used for *in vitro* translation and fusion formation in rabbit reticulocyte lysate to produce the mRNA-protein fusion library (7).

We found the above protocol to be useful for efficient synthesis of mRNA-peptide and mRNA-protein fusion libraries from a number of template libraries different in length and complexity. Efficient mRNA affinity purification on neutravidin agarose carrying the biotinylated puromycin-linker replaced the laborious mRNA gel purification step used in previous protocols (3,5,7). Crosslink-formation of the hybridized mRNA on solid-phase and simultaneous biotin-tag cleavage yielded pure mRNA-puromycin template for direct use in an *in vitro* translation system. The ease with which the amount of synthesized mRNA-puromycin template can be controlled by hybridization with limiting amounts of linker allowed us to keep the reaction volumes and stoichiometry of the subsequent fusion synthesis constant and independent from the yield of the preceding transcription reaction. This feature of the present protocol should prove most useful in an automated *in vitro* protein evolution process which involves repetitive cycles of fusion library synthesis and selection and where manual quantification and adjustments only lead to a time-lag in the production process.
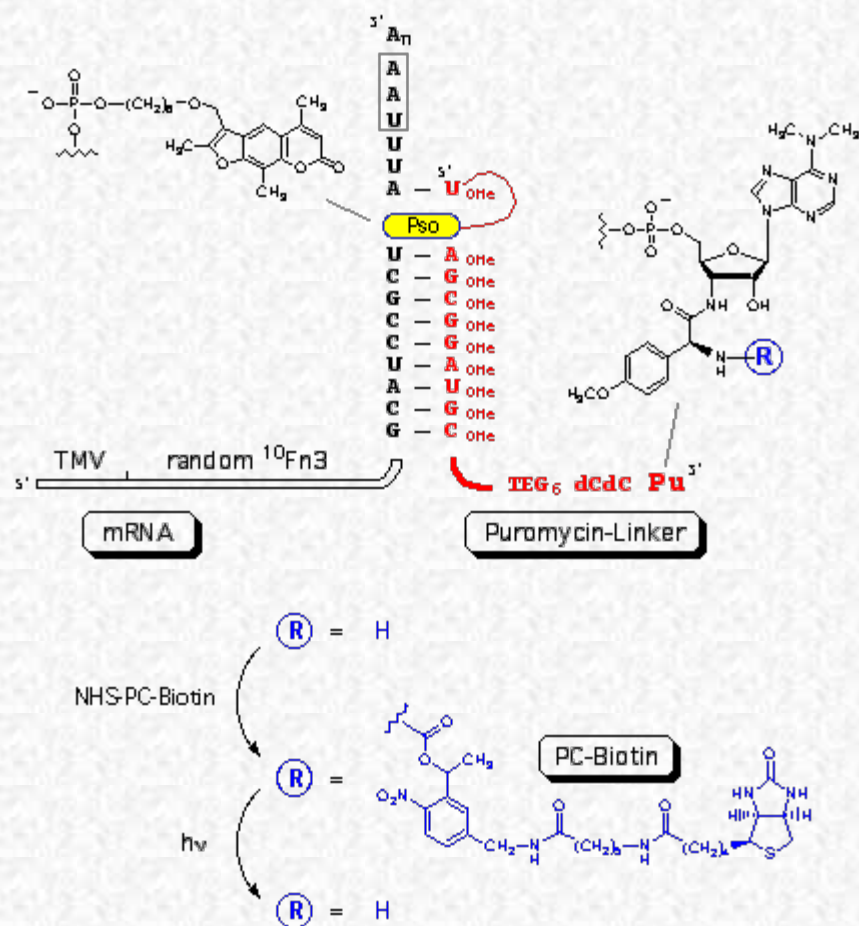
Figure 2: Sequence design of the mRNA template 3'-end and the puromycin-linker. The mRNA template 3'-constant region codes for a linker hybridization sequence, a UAA stop codon and a dA_{18} tag sequence. The biotinylated puromycin-linker sequence shown in red carries a 5-terminal psoralen moiety (Pso), followed by a hybridization region and a flexible tether with puromycin (Pu) at the 3'-end and carrying a photo-cleavable biotin-tag. A 10 base complementary region between mRNA 3'-end and the photo-linker was found to be sufficient to achieve efficient hybridization and photo-crosslink formation. The use of 2'-OMe RNA in the linker hybridization sequence was intended to increase the stability of the 10 base-pair duplex; also, it protected the mRNA from cleavage by any RNase H activity during translation in rabbit reticulocyte lysate. The hydroxymethyl trioxsalen moiety (psoralen) was tethered to the 5'-phosphate through a hexamethylene spacer (8) (psoralen C6, Glen Research). METHODS: The puromycin-linker was prepared according to Kurz et al. (7) and biotinylated by carbamate bond formation between the puromycin amino group (50 μM) and the photo-cleavable biotin-reagent (9) (NHS-PC-Biotin, 5 mM, EZ-Link^{TM}-Biotin, Pierce Chemicals) in 25% DMSO/water for 2 h at room temperature followed by NaCl/EtOH precipitation. Photo-cleavage of the 1-(2-nitrophenyl)ethyl carbamate moiety restored the primary amino group on the 3'-terminal puromycin which acts as the peptidyl-acceptor in the fusion formation process (see Fig. 1).

References

1. Felder, E. (1994) *Chimia*, **48,** 531-541.

2. Shusta, E. V., VanAntwerp, J. and Wittrup, K. D. (1999) *Curr Opin Biotechnol*, **10,** 17-122.

3. Roberts, R. W. and Szostak, J. W. (1997) *Proc Natl Acad Sci U S A*, **94,** 12297-302.

4. Nemoto, N., Miyamoto-Sato, E., Husimi, Y. and Yanagawa, H. (1997) *FEBS Lett*, **414,** 405-8.

5. Liu, R., Barrick, J. E., Szostak, J. W. and Roberts, R. W. (2000) *Methods in Enzymology*, **318,** 268-93.

6. Roberts, R. W. (1999) *Curr Opin Chem Biol*, **3,** 268-73.

7. Kurz, M., Gu, K. and Lohse, P. A. (2000) *Nucleic Acids Res*, **in press**.

8. Pieles, U. and Englisch, U. (1989) *Nucleic Acids Res*, **17,** 285-99.

9. Olejnik, J., Sonar, S., Krzymanska-Olejnik, E. and Rothschild, K. J. (1995) *Proc Natl Acad Sci U S A*, **92,** 7590-7594.

All comments on this poster should be sent by e-mail to (mailto:ecsoc@listserv.arizona.edu) ecsoc@listserv.arizona.edu with **B0012** as the message subject of your e-mail.