cDNA – Protein Fusions: Covalent Protein – Gene Conjugates for the In Vitro Selection of Peptides and Proteins

Markus Kurz,[a] Ke Gu,[a] Amal Al-Gawari,[b] and Peter A. Lohse*[a]

We report a method for the synthesis of covalent cDNA-protein fusions for protein display applications. A branched mRNA template was developed which carries a peptidyl acceptor and a reverse transcription primer at the 3′-end. Translation in vitro followed by reverse transcription produced a protein covalently bonded to its encoding cDNA. Both single- and double-stranded cDNA – protein fusions were prepared. cDNA – protein fusions are stable in alkali and resistant to ribonucleases. Their simple preparation and their resistance towards degradation should make cDNA – protein fusions a useful tool for the in vitro selection and evolution of high affinity ligands from large libraries of polypeptides.

KEYWORDS:

combinatorial chemistry · in vitro selection · molecular evolution · protein display

Introduction

The need for target-specific peptide and protein ligands for the discovery of novel therapeutics, diagnostics, medical imaging compounds, and research reagents demands efficient selection technologies that will allow the identification of the polypeptides with the best binding properties from large libraries. The technology for preparing diverse libraries of polypeptides has long existed; however, in the absence of a genetic selection and amplification protocol, the number of individual mutants which could be screened for enhanced ligand – binding properties limited the size of the library and therefore the likelihood of successfully identifying protein binders with novel properties. The emergence of bacteriophage display,[1] the first library technology that allowed the topological linkage of genotype (DNA) to phenotype (expressed protein), enabled the isolation of genes based on the properties of the encoded protein. Thus, large libraries of up to 10^{10} individual mutant clones could be screened for ligand – binding affinity and the selected mutants could be isolated from the mixture. To expand the utility of genetic selection technologies, alternative techniques have been developed, such as technologies based on prokaryotic and eukaryotic organisms as well as in vitro technologies, which allow protein selection and evolution in the absence of a living cell or virus.[2, 3]

In vitro selection technologies combine two key strengths: First, the potential to generate large libraries by obviating a cell transformation step and, second, the construction and screening of libraries independent of their effect on cell viability. Originally described by Matteakis et al.[4] for short peptides, ribosome display involves the preservation of a polypeptide – ribosome – mRNA ternary complex as a genetic unit. Ribosome display has been adapted to the screening of larger proteins, such as single-chain antibodies;[5, 6] an optimized protocol has been described for identifying proteins with improved expression, stability, and affinity.[7] The limitations of this noncovalent display method include the low temperatures, the narrow range of salt concentrations, and pH values under which the selection must be carried out to preserve the integrity of the mRNA – ribosome – peptide complex.

The chemically more robust DNA has also been used to tag protein libraries. For example, peptides that bound specifically to the anti-dynorphin B monoclonal antibody were identified from a library of random peptides which had been linked to the C terminus of the lac repressor.[8] The repressor protein links the peptides to the plasmid encoding them by binding to *lac* operator sequences on the plasmid. To ensure complex formation of the library peptides with their encoding plasmids and to prevent the unwanted *trans* conjugation, the reaction must be carried out inside a cell. Furthermore, the noncovalent nature of the linkage between these peptides and DNA requires gentle reaction conditions for product preparation and for selection, again imposing limits on the scope of the in vitro selection experiments.

A display method based on conjugates between streptavidin-fused polypeptides and their encoding biotinylated DNA in confined compartments was developed recently.[9] To ensure *cis* conjugation, transcription and translation were carried out in

[a] *P. A. Lohse, M. Kurz, K. Gu*
Phylos, Inc., 128 Spring Street, Lexington, MA 02421 (USA)
Fax: (+1) 781-402-8800
E-mail: plohse@phylos.com

[b] *A. Al-Gawari*
Millennium Pharmaceuticals, Inflammation/Immunobiology, 75 Sidney St., Cambridge, MA 02139 (USA)

 © WILEY-VCH-Verlag GmbH, D-69451 Weinheim, 2001 1439-4227/01/02/09 $ 17.50+.50/0 CHEMBIOCHEM **2001**, 2, 666 – 672

emulsion compartments generated from a water in oil emulsion. The resulting protein – DNA conjugates can be recovered from the emulsion and can be subjected to affinity selection based on the properties of the peptide portion. Success of this approach for the efficient identification of polypeptide ligands will largely depend on the extent of *trans* conjugation as a result of diffusion of DNA and polypeptides between aqueous compartments.

More recently, new types of in vitro display technologies, which feature the covalent linkage between phenotype and genotype have emerged. Owing to a robust covalent linkage between the nucleic acid moiety and the encoded protein, these covalent display technologies allow the selection of ligand-binding polypeptides under a wide variety of conditions. A recent review describes the preparation of covalent DNA – polypeptide complexes by employing the *cis*-nicking activity of the replication initiator protein from *E. coli* bacteriophage P2A.[10] The protein is an endonuclease that introduces a single-strand discontiniuity in the DNA. Gel-shift experiments suggested that the 5′-end of the nicked DNA became covalently attached to a tyrosine residue in the active site of P2A.[11, 12] The *cis* activity of the nicking process enables pools of polypeptides that are genetically fused to P2A to be synthesized in vitro so that they become covalently attached to their own coding sequence. Library sizes of $> 10^{12}$ peptide sequences have been constructed and enriched for covalent complexes that display target-binding polypeptides from a background of nonbinding complexes.[10]

A new method for the preparation of a minimal selection unit that consists of a mRNA covalently linked to a polypeptide (PROfusion) was introduced recently.[13–15] mRNA – protein fusion molecules are made by in vitro translation of mRNA – puromycin templates which had been prepared enzymatically[15] or chemically.[16] The library size is limited only by the volume and the efficiency of the translation reaction, and by the efficiency of fusion formation on the ribosome. The covalent bond between the mRNA and the encoded protein conferred the necessary stability in such a way as to enable selection and amplification of these conjugates. At the present time, libraries containing up to 10^{14} different sequences have been generated and have been used for the isolation of specific and tight binders to a variety of protein targets (unpublished data, Phylos, Inc.; for applications of mRNA – protein fusion technology see references[2] and[17]).

We sought to extend the concept of fusions between nucleic acids and proteins to covalent cDNA – protein fusions. We reasoned that a covalent link between the DNA genotype and the encoded protein would confer increased stability to the fusion molecule because DNA is generally more stable towards degradation than RNA. Herein we describe experimental strategies for the preparation of single- and double-stranded cDNA – protein fusions. Our experimental strategy builds on a recently reported, improved method for the preparation of mRNA – peptide fusions.[16] A novel type of mRNA – puromycin conjugate was designed for accelerated synthesis of cDNA fusions. The stability of cDNA – protein fusions and the simplicity of their preparation should be an advantage in in vitro selections, especially, when RNase-free preparation of the target cannot be obtained.

Results

The first step toward the construction of cDNA – protein fusions of Type I was the synthesis of an mRNA – protein fusion precursor according to the procedures described by Liu et al.[15] We chose the short mRNA **1** that contained an open reading frame encoding the Flag epitope[18] and StrepTag II[19] peptide sequence, followed by a short linker encoding for the tripeptide ASA (Figure 1). The 5′-terminus of the mRNA contained the tobacco mosaic virus (TMV) 5′-untranslated region for efficient translation intiation. The 3′-end of this mRNA was then enzymatically ligated to the DNA-linker **3** that contained a 3′-

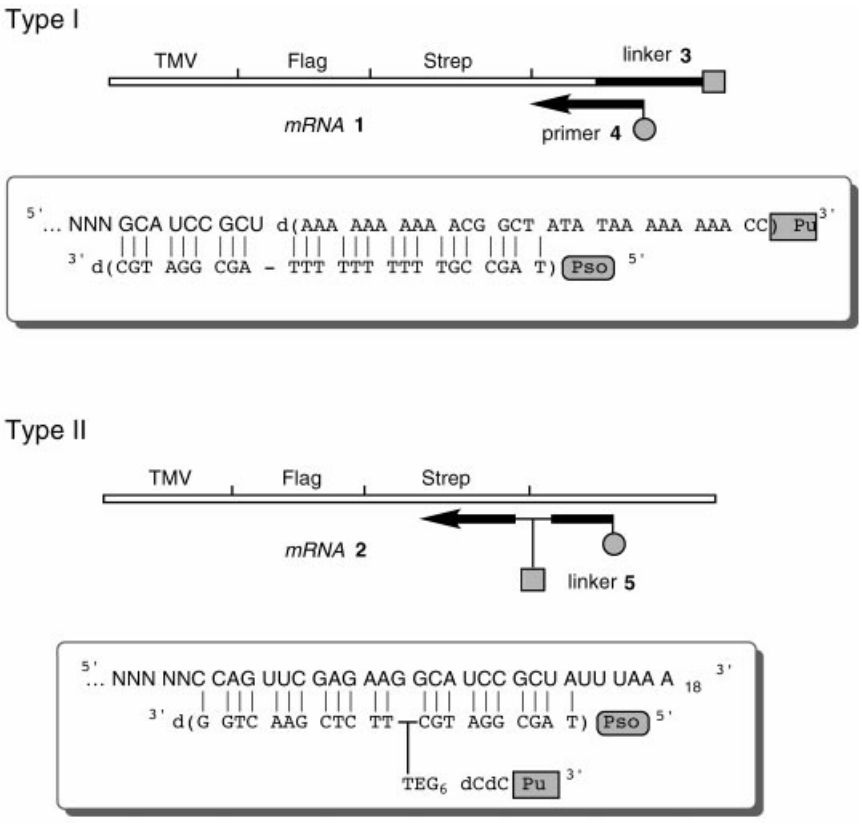

Figure 1. *Design of mRNA templates and oligonucleotides used for the preparation of cDNA – peptide fusions (Type I & Type II). The mRNAs contained a promoter sequence derived from the tobacco mosaic virus (TMV), followed by the open reading frame which encoded for a Flag epitope and the StrepTag II sequence. Nucleotide sequences are shown for linkers and primers. The putative psoralen photo-crosslink site is underlined. DNA and RNA are represented with filled and open lines respectively. Psoralen (Pso) and puromycin (Pu) are represented with a gray circle and a gray rectangle, respectively.*

puromycin and a d(TA) psoralen photo-crosslinking site (Figure 2).

In vitro translation of template **8** was carried out at 30° C for 30 min in rabbit reticulocyte lysate. Upon addition of potassium and magnesium salts and incubation for 1 h at room temperature, mRNA – protein fusion **9** was formed. Subsequent purification on oligo-dT cellulose allowed separation of the fusion product **9** from buffer components and free peptide. Ligated mRNA **8** was copurified with the fusion product **9**; therefore, all subsequent reaction steps affected both fused and unfused mRNA. Unfused mRNA **8** did not carry the ^{35}S-radiolabel and was invisible to gel analysis by autoradiography. Psoralen-modified primer **4** was added in excess and the reaction mix was irradiated with $\lambda > 300$ nm light for 15 min. A pyrex filter (300 nm cutoff) was used during irradiation to minimize photodamage of the mRNA and photocleavage of the crosslink due to short wavelength UV light. Because the gel shift of the crosslinked fusion product **10** was indistinguishable from the shift of **9** (Figure 2, lanes 1 and 2), the photo-crosslinking yield was determined indirectly after RNase H digestion. Different gel shifts of the digestion products **11** and **12** indicated high yield for the psoralen photo-crosslinking reaction (Figure 2 lanes 3 and 4). Subsequent reverse transcription of **10** yielded the mRNA/cDNA double-stranded fusion **14**. As expected, the noncrosslinked control sample produced only unmodified mRNA – protein fusion **9** upon electrophoresis on a denaturing polyacrylamide – urea gel (Figure 2, lane 5). A new product band was identified for product **14**, where dissociation of the nucleic acid duplex was prevented by the interstrand psoralen crosslink (Figure 2, lane 6). Finally, treatment of **14** with RNase H removed the RNA portion and furnished the single-stranded cDNA – protein fusion **15** which migrated as expected with a similar shift as the ssRNA – fusion **9** (Figure 2, lane 8). RNase H treatment of the noncrosslinked control **13** led to complete dissociation of the linker – peptide fragment **11** from the cDNA (lane 7). Double-standed cDNA fusion **16** was obtained by elongation of primer **7** using reverse transciptase.

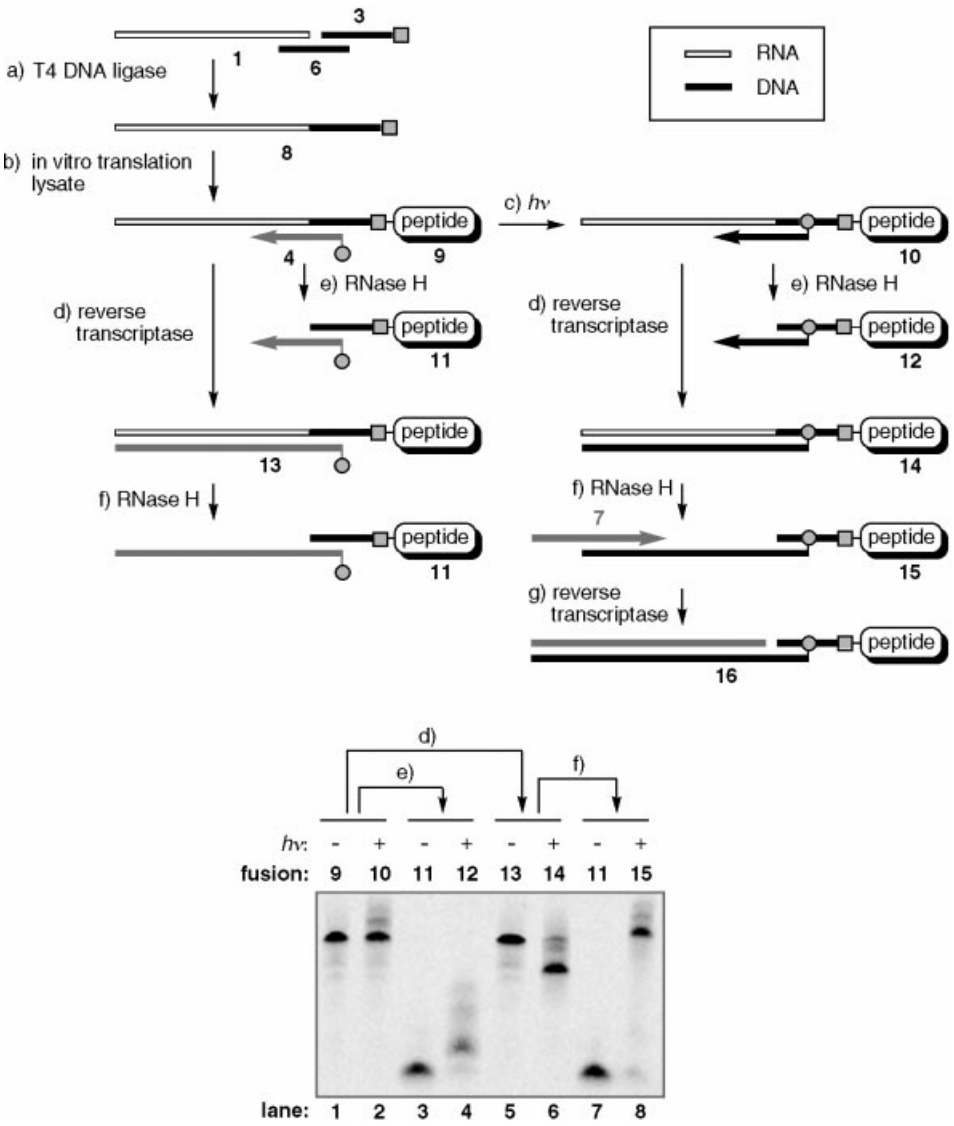

Figure 2. *Synthetic scheme for the preparation of cDNA – peptide fusion of Type I. Steps d) – f) below outline the control reactions after omission of the photo-crosslink step c). Fusion products containing ^{35}S-labeled peptide were detected by autoradiography after denaturing polyacrylamide gel electrophoresis (PAGE). Noncrosslinked and photocrosslinked fusion products were applied to the gel in pairs.*

An alternative and more direct route to cDNA – protein fusions Type II was developed by employing the branched puromycin linker **5** (Figure 1). Linker **5** was prepared by DNA solid-phase synthesis on puromycin – CPG by using a branched phosphoramidite. Synthesis of the reverse transcription primer portion was achieved with 5′-phosphoramidites. After selective deprotection of the branch, the synthesis was concluded by attachment of the 5′-terminal decanucleotide carrying the crosslinker psoralen. The mRNA construct **2** coded for the same ORF as mRNA **1** and carried in addition an UA psoralen photo-crosslinking site, followed by an UAA stop codon and a poly-A stretch as a handle for oligo-dT purification.[16] The psoralen-modified linker **5** could be photo-crosslinked to mRNA **2** in high yield (Figure 3). The crosslinked conjugate **17** was used for in vitro translation without further purification. The reaction was carried out under similar conditions as above (Type I) and furnished mRNA – protein fusion **18**, which was then purified on oligo-dT cellulose. Reverse transcription of **18** followed by RNase H digestion yielded the cDNA fusion **20** with a

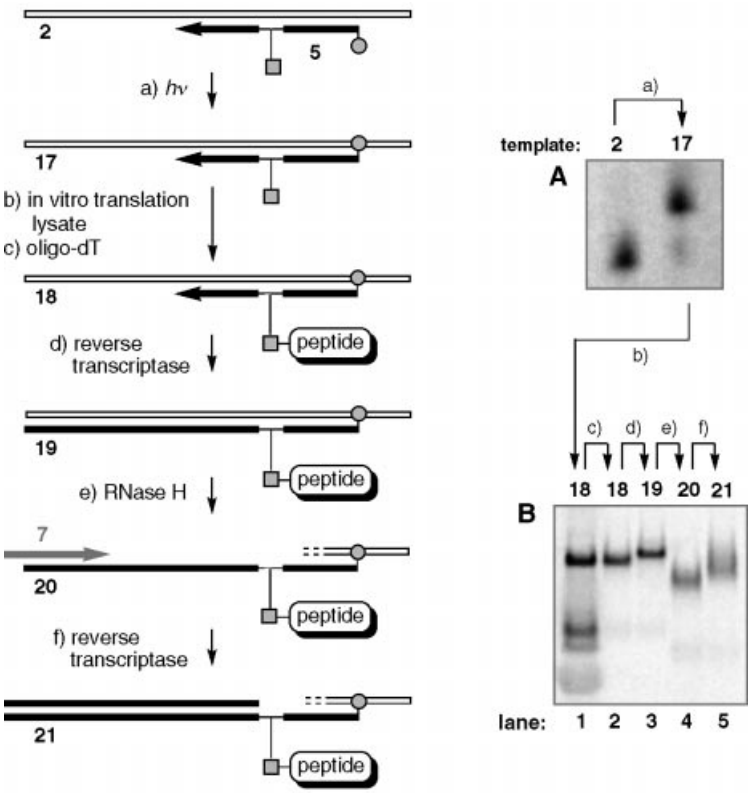

Figure 3. *Synthetic scheme for the preparation of cDNA – peptide fusion of Type II. Electrophoretic denaturing gel analysis followed by UV-shadowing was used to determine the photo-crosslinking yield of step a (gel A). The reaction products of steps b – f were analyzed by electrophoresis on a non-denaturing NuPage gel (gel B). Fusion products carry the ^{35}S-radiolabel in the peptide portion and were detected by autoradiography.*

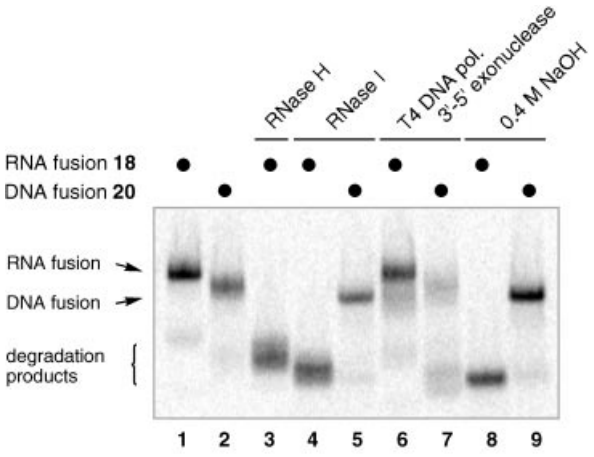

Figure 4. *Stability of mRNA– and cDNA – peptide fusion towards nucleases and sodium hydroxide. Incubations were performed for 30 min at 37 °C, followed by product analysis on a NuPage gel.*

gel shift similar to the mRNA fusion **18**, thus demonstrating a stable crosslink between the cDNA genotype and the encoded peptide carrying the ^{35}S methionine radiolabel (Figure 3). Its subsequent conversion to the double-stranded cDNA – protein fusion **21** followed the same reaction steps as described for the synthesis of the Type I fusions **15** and **16**.

In addition to the gel analysis, a set of experiments was carried out with cDNA – protein fusions of Type II to confirm their structure. First, the nature of the peptide portion was investigated by specific binding to affinity reagents. Our model mRNA **2** encoded for the Flag (DYKDDDDK) and Strep tag II (WSHPQFEK) epitope sequences. After incubation with Anti-Flag M2 Affinity Gel followed by washing steps, 57% of the cDNA fusion **20** could be retained on the matrix thus verifying the presence of the correctly expressed Flag peptide sequence. Binding to StrepTactin sepharose occurred to a somewhat lesser extent, leading to 18% binding of fusion **20**.

To examine the nature of the nucleic acid portion, we subjected fusion constructs **18** and **20** to different nuclease and alkaline treatments (Figure 4). Upon addition of RNase H to mRNA – protein fusion **18**, the mRNA strand was cleaved in the double-stranded mRNA/cDNA hybrid region, leaving only the single-stranded mRNA 3′-overhang remaining bound to peptide and linker **5** (Figure 4, lane 3). Incubation with RNase I led to complete digestion of the mRNA leading to the faster moving peptide – DNA – linker conjugate (Figure 4, lane 4). As expected, single-stranded cDNA – protein fusion **20** proved to be stable to RNase treatment (Figure 4, lane 5). However, treatment of fusion **20** with T4 DNA polymerase that exhibits 3′-5′ exonuclease activity led to degradation of the single-stranded cDNA (Figure 4, lane 7), whereas RNA fusion **18** appeared to remain largely unchanged under these conditions (Figure 4, lane 6). However, treatment of the mRNA fusion **18** with 0.4 M sodium hydroxide led to the degradation of the mRNA (Figure 4, lane 8) but left the cDNA fusion **20** untouched (Figure 4, lane 9). Interestingly, the sharpness of the cDNA fusion band greatly improved upon treatment with base and RNase I (Figure 4, compare lane 2 with lanes 5 and 9). We attribute this to a somewhat incomplete mRNA digestion with RNase H during preparation of **20** (see also step e in Figure 3). As a result, the product was likely to display some distribution in nucleic acid length as seen by the diffuse appearance of bands during gel electrophoresis. RNase I or sodium hydroxide removed mRNA more efficiently, thus leading to a sharper band for cDNA fusion **20**.

To further demonstrate the increased stability of cDNA – protein fusions over mRNA – protein fusions, single- and double-stranded fusions **9**, **13**, **15**, and **16** were incubated with chinese hamster ovarian cell membranes at room temperature (Table 1). Aliquots were withdrawn over time and analyzed by gel electrophoresis to determine fusion half-lives. As expected, the mRNA fusions **9** and **13** were most rapidly degraded. The addition of VRC ribonuclease inhibitor reduced the rate of mRNA decomposition by about fourfold but did not completely abolish nuclease activity. Single-stranded cDNA fusion **15** was also subject to nuclease degradation, but showed a more than twofold increased half-life over mRNA-fusion **9**. Double-stranded cDNA fusion **16** was completely stable for two hours. No more radioactivity could be detected after gel-analysis of samples which were exposed to cell membranes for 24 h. This suggested

Table 1. Half-life values of peptide fusions in the presence of chinese hamster ovarian cell membranes. VRC = vanadyl ribonucleside complex.

Fusion construct		VRC	$t_{1/2}$	
ssRNA		**9**	– +	35 min 120 min
cDNA/RNA		**13**	– +	12 min 50 min
ssDNA		**15**	–	80 min
dsDNA		**16**	–	> 2 h, < 24 h

that residual protease activity lead to degradation of the protein portion of the fusion which carries the radioactive label (^{35}S).

Discussion

In the synthesis of cDNA – protein fusion Type I, a psoralen photo-crosslinking reaction was used to covalently attach the reverse transcription primer to the mRNA – peptide fusion molecule **9**. To confer specificity to the crosslinking reaction, the DNA – linker region in **9** was designed to code for a reverse transcription primer hybridization sequence and for a d(TA) psoralen intercalation site (Figure 1).[20] Although the sequence of linker **3** differed slightly in sequence from the polynucleotide linkers described previously, in vitro translation of **8** produced the fusion product **9** in a yield similar to the yields reported for templates carrying poly-dA derived puromycin linkers (Figure 2).[15]

In the synthesis of the Type II cDNA – protein fusion, the use of the branched puromycin linker **5** that carried a covalently bound reverse transcription primer rendered the post-translational photo-crosslinking step unnecessary (Figure 3). The novel linker design was based on a recently developed and improved strategy for the preparation of mRNA – puromycin templates, which used a psoralen interstrand crosslink as a pausing site for the ribosome.[16] According to these results, the highest fusion yield was achieved when the 3′-terminal puromycin was attached through highly flexible triethylenglycol phosphate (TEG) units to the hybridization sequence carrying the 5′-psoralen. Based on these findings, dCdC-puromycin was attached through six TEG units to a branched phosphate building block in the hybridization sequence of linker **5**. Sequence specific hybridization of **5** to mRNA **2** followed by psoralen interstrand crosslinking formed the mRNA – linker conjugate **17** in good purity for direct fusion formation in rabbit reticulocyte lysate. The fusion yield of **17** was equal to the yields which were typically obtained with templates carrying linear (nonbranched) photolinkers.[16]

Binding of cDNA – protein fusion **20** through the N-terminal Flag-epitope to the anti-Flag antibody was more efficient (57%) than binding of the C-terminal Strep-tag to the StrepTactin affinity reagent (18%). This could be in part due to steric hindrance or electrostatic repulsion caused by the proximity of the C-terminal tag to the comparatively large and highly charged nucleic acid portion during binding to the anti-Flag antibody. Alternatively, the 3′-terminal mRNA – primer duplex may cause an antisense effect and block the ribosome from completely translating the StrepTag II coding region. The duplex formed between linker **5** and mRNA **2** reaches into the ORF and might lead, at least partially, to ribosome stalling and premature fusion formation. The incorporation of a linker sequence following a structurally important coding sequence could alleviate this effect and ensure complete translation of the ORF.

Degradation experiments confirmed our expectations that the cDNA fusion is resistant to ribonuclease digestion and is stable at alkaline pH, whereas the mRNA fusion is completely degraded under these conditions (Figure 4). A comparison of half-lives after incubation with chinese hamster ovarian cell membranes showed that cDNA – protein fusions **15** and **16** are significantly more robust than the mRNA fusion **9** and **13** under these conditions (Table 1). Addition of the nuclease inhibitor vanadyl ribonucleoside complex (VRC) increased the half-lives of mRNA – protein fusions by about fourfold to reach times which were comparable to the half-lives of cDNA – protein fusions in the absence of inhibitor. The lability of the mRNA/cDNA – peptide construct **13** over the single-stranded mRNA – peptide fusion **9** suggested RNase H activity in the membrane preparation. The double-stranded cDNA – peptide fusion **16** was most inert towards degradation and showed a half-life which was at least fourfold larger than the half-life found for the mRNA – protein fusions in the absence of ribonuclease inhibitor.

The present cDNA fusion technology should be most useful for protein selection in vitro. Libraries of cDNA – protein fusions can be prepared and polypeptide ligands to a desired protein target can be isolated by iterative rounds of affinity selection and amplification (Figure 5). Double-stranded fusions are preferentially employed in the affinity selection step to prevent single-stranded nucleic acid sequences from forming functional structures and to ensure that the protein moiety alone is responsible for the function in the fusion molecule. The cDNA/mRNA hybrid form can be easily obtained by primer elongation of the cDNA fusion template of Type II and does not involve a heating step for sequence-specific annealing of an endogenous primer as described previously.[15] Thus, cDNA fusion technology should be most beneficial in in vitro selection protocols employing antibody, antibody(mimic) or cellular protein libraries, and where a heating step would lead to denaturation of the fused protein. This could cause misfolded structures to be carried into a selection.

A further advantage of using cDNA – protein fusion technology for in vitro selection protocols includes the resistance of the cDNA genotype towards degradation by RNases. This robustness should prove beneficial for in vitro selections when an RNase-free preparation of the target cannot be obtained. Applications include the selection of affinity ligands to receptor proteins, which are difficult to obtain in active form from recombinant sources, and which need to be presented on membranes or on whole cells.

The robustness of cDNA – protein fusions in alkali could allow for post-translational modification of the protein portion. For

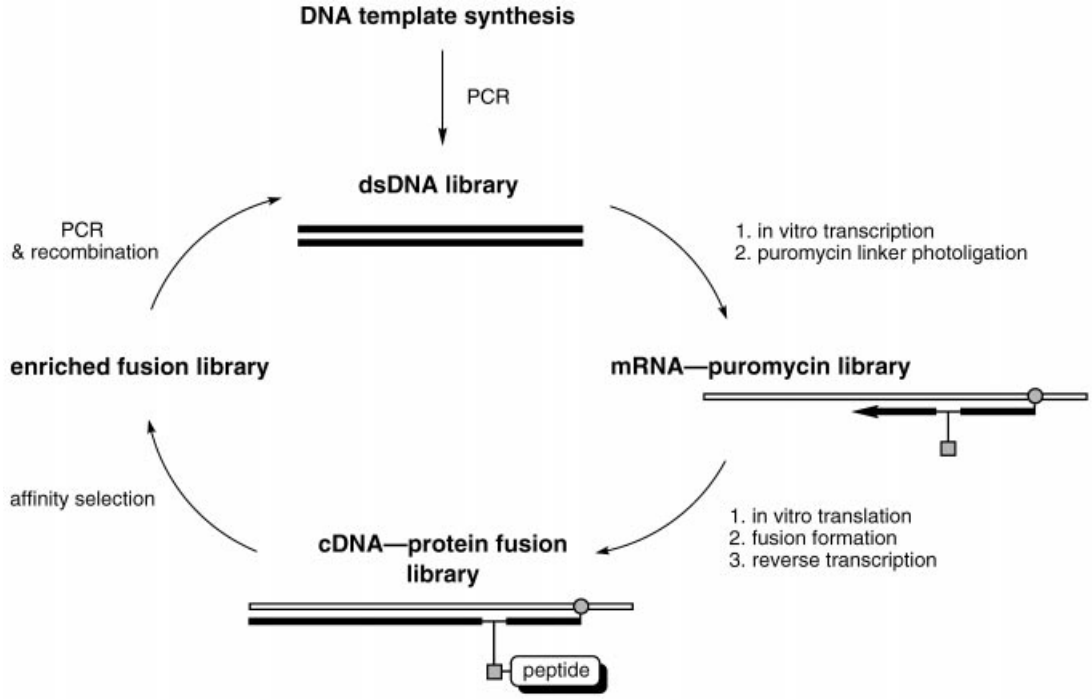

Figure 5. *In vitro selection scheme. A protein library is prepared according to the present protocol for the preparation of Type II cDNA fusions. Target-binding fusion molecules can be isolated in the affinity selection step and can subsequently be amplified to afford an enriched pool of binding sequences.*[2, 15]

example, base-catalyzed β-elimination of phosphate in phosphoserine or phosphothreonine followed by nucleophilic addition of nonnatural residues should be possible without destroying the cDNA genotype.[21] An in vitro selection protocol without intervening amplification, similar to the protocol described for the isolation of DNA aptamers with modified nucleotides,[22] could be applied for the efficient identification of nonnatural polypeptide ligands from chemically modified cDNA – protein fusion libraries.

In summary, the robustness of cDNA – protein fusions should provide flexibility in the choice of selection conditions and should prove most useful in in vitro selection applications where mRNA – protein fusions would be degraded. We believe that this robustness and the ease of their preparation should make cDNA fusions a versatile tool for in vitro peptide and protein display applications.

Materials and Methods

Oligonucleotides: The model mRNA substrates **1**: GGG ACA AUU ACU AUU UAC AAU UAC AAU GGA CUA CAA GGA CGA UGA CGA UAA GGG CGG CUG GUC CCA CCC CCA GUU CGA GAA GGC AUC CGC U and **2**: GGG ACA AUU ACU AUU UAC AAU UAC AAU GGA CUA CAA GGA CGA UGA CGA UAA GGG CGG CUG GUC CCA CCC CCA GUU CGA GAA GGC AUC CGC UAU UUA AAA AAA AAA AAA AAA AAA A were prepared by T7 transcription (Megashortscript transcription kit, Ambion, TX, USA) from dsDNA templates. After transcription, the mRNAs were purified by denaturing polyacrylamide gel electrophoresis. The modified oligonucleuotides **3**: 5′-pd(AAA AAA AAA ACG GCT ATA TAA AAA AAA CC)-Pu, **4**: 5′ psoralen C2-TAG CCG TTT TTT TTT TAG CGG ATG C and **5**: 5′ ggt caa gct ctt-branch [5′ psoralen C6 -TAG CGG ATG C 3′] spacer$_6$ CC-Pu [uppercase = standard DNA-3′-phosphoramidites; lowercase = DNA-5′-phosphoramidites; spacer = spacer-9 phosphoramidite; Pu = puromycin-CPG (all Glen Research, VA, USA); branch = asymmetric branching amidite (Clontech, CA, USA)] were synthesized on an Expedite Synthesizer Model 8909 (PerSeptive Biosystems, MA, USA) according to recommended protocols for the corresponding phosphoramidites. For the branched construct **5**, the main chain was synthesized first and concluded with a final capping step. Next, the levulinyl protecting group was removed from the branching unit through treatment with 0.5 M hydrazine monohydrate in pyridine – acetic acid for 15 min at room temperature. Automated synthesis was then resumed and the side chain sequences (indicated in square brackets) were attached. The oligonucleotides were fully deprotected in concentrated ammonium hydroxide for 8 h at 55 °C and purified by denaturing polyacrylamide gel electrophoresis. The DNA sequences **6**: d(TTT TTT TTT TAG CGG ATG C) and **7**: d(TAA TAC GAC TCA CTA TAG GGA CAA TTA CTA TTT ACA ATT) were purchased from Oligos etc. (OR, USA) and used without further purification.

Synthesis of cDNA – protein fusions of Type I (Figure 2): mRNA **1** and linker **3** were hybridized to template DNA **6** and enzymatically ligated by T4 DNA ligase as described previously.[15] After purification by electrophoresis on a denaturing polyacrylamide gel, the resulting mRNA – linker construct was used as template for in vitro translation using the Retic Lysate IVT kit from Ambion. Reactions contained ligated mRNA **8** (50 pmol), creatine phosphate (10 mM), potassium acetate (150 mM), magnesium chloride (0.5 mM), 0.1 mM of each amino acid exept methionine, [^{35}S] methionine (150 μCi; Amersham) and 67 % v/v of lysate in a total volume of 300 μL and were carried out for 30 min at 30 °C. To promote the subsequent formation of mRNA – peptide fusions, KCl and $MgCl_2$ were added to 590 mM and 50 mM final concentrations, respectively, in a volume of 500 μL. Incubation was continued for 60 min at 20 °C. Products were isolated by diluting the lysate into binding buffer (10 mL; 100 mM Tris·HCl

pH 8.0 at 25 °C, 10 mM EDTA, 1 M NaCl, 0.25 % v/v Triton X-100) and adding 10 mg oligo-dT cellulose type 7 (Pharmacia, NJ, USA). Samples were rotated for 60 min at 4 °C. The solid support was then washed with ice-cold binding buffer (5 mL) that was devoid of EDTA, followed by elution with aliquots of H_2O (100 μL). Product was found in fractions 2 and 3, which were then combined. The total yield of mRNA fusion **9** was determined by scintillation counting of the incorporated [^{35}S] methionine to be 1.6 pmol (3.2 % of input mRNA **8**). The oligo-dT-purified material **9** (20 μL) was mixed with primer **4** (0.5 μL, 50 μM) and first strand buffer (6 μL, Superscript II kit from GibcoBRL; 250 mM Tris·HCl pH 8.3 at 25 °C, 375 KCl, 15 mM $MgCl_2$) and briefly heated to 80 °C for 2 min, followed by slowly cooling to 0 °C. Psoralen photo-crosslink formation was induced by irradiating the sample for 15 min at 0 °C with $\lambda > 310$ nm [450 W medium pressure immersion lamp (ACE Glass), equipped with a Pyrex absorption sleeve in a Quartz immersion well]. Next, a dNTP mix (0.6 μL; 25 mM each), 0.1 M DTT (3 μL) and Superscript II reverse transcriptase (0.4 μL; 80 units) were added and cDNA synthesis was carried out for 60 min at 42 °C. The RNA portion was then removed from **14** by continuing incubation for 60 min at 37 °C after addition of RNase H (0.5 μL; 1 unit; Promega) to yield single-stranded cDNA – protein fusion **15**. Finally, double-stranded DNA **16** was generated by adding primer **7** (50 pmol) incubating for another 60 min at 42 °C. Product analysis was carried out by electrophoresis on denaturing 6 % TBE – urea polyacrylamide gels (Novex), followed by visualizing of the ^{35}S-labeled product bands through exposing on phosphorimager screen. Control reactions were performed following the same reaction sequence but ommitting the psoralen photo-crosslink formation (Figure 2, steps d – f)

Synthesis of cDNA – protein fusions of Type II (Figure 3): The branched linker construct **5** (5 μM) was annealed to mRNA **2** (2.5 μM) in 25 mM Tris buffer (pH 7.0 at 25°C), containing NaCl (100 mM) and crosslinked by irradiation for 15 min at room temperature in a borosilicate glass vial (Kimble/Kontes, NJ, USA) using a handheld multiwavelength UV lamp model UVGL-25 (UVP, CA, USA) set to long wave. Product analysis was performed by electrophoresis on a 6 % TBE – urea polyacrylamide gel followed by visualizing through UV shadowing and indicated nearly quantitative conversion of the starting material. The photo-ligated product **17** was used for in vitro translation without further separation from remaining unligated mRNA and excess linker. In vitro translation and fusion formation reactions were performed as described for Type I with 100 pmol input mRNA in a 300 μL total volume. After purification on oligo-dT cellulose we obtained mRNA fusion **18** (5.5 pmol). Its conversion into single-stranded and double-stranded cDNA – protein fusions **20** and **21**, respectively, was done by reverse transcription (Superscript II kit from GibcoBRL) and RNase H (Promega) treatment as described for Type I fusion.

Flag immunoprecipitation and StrepTactin binding: ^{35}S-labeled cDNA – protein fusion **20** (10 μL; specific acticity 4000 cpm μL^{-1}) was added to Anti-Flag M2 Affinty Gel (20 μL; Sigma, MO, USA) in 300 μL buffer containing 50 mM Tris·HCl pH 7.4 at 25°C, 1 % NP 40, NaCl (150 mM), EDTA (1 mM), Na_3VO_4 (1 mM), and NaF (1 mM). A second sample was prepared with StrepTactin sepharose (20 μL; Genosys, TX, USA) in 300 μL buffer containing 100 mM Tris·HCl pH 7.1 at 25°C, EDTA (1 mM) and yeast tRNA (0.5 mg mL^{-1}). The mixtures were rotated for 1 h at 4 °C and then transferred to a Ultrafree-MC filter unit (0.45 μm; Millipore, MA, USA). The buffers were removed by centrifugation, and the residues washed with ice-cold buffer (5 × 300 μL). The cDNA – protein fusion-affinity reagent interaction was disrupted by heating the residues in 400 μL NuPage gel-loading buffer to 70 °C for 5 min. An aliquot of 20 μL was removed and quantitated by scintillation counting. The amount of fusion bound to the anti-Flag gel was found to be 57 % as being determined by the relative amount of radioactivity recovered; binding to StrepTactin occurred with 18 % efficiency. For comparison, binding of the same construct as mRNA – peptide fusion took place with 57 and 63 %, respectively.[16]

Nuclease and base treatment: To cDNA fusion **20** (10 μL) or mRNA fusion **18** (10 μL) in reverse transcription buffer we added either RNase H (0.2 μL; 0.4 units), RNase I (0.2 μL; 2 units), T4 DNA polymerase (0.2 μL; 0.6 units; 3′-5′ exonuclease activity) or 2.0 M NaOH (2.5 μL). Samples were incubated for 30 min at 37 °C and then analyzed on a 4 – 12 % NuPage polyacrylamide gel (Novex) followed by autoradiography (Figure 3).

Stability in biological media: 5 nM of either mRNA fusions **9** or **13**, or cDNA fusions **15** or **16** were incubated with CHO-K1 cell membranes (3 μg μL^{-1}; Receptor Biology, MD, USA) in 50 mM Tris·HCl pH 8.3 at 25 °C, Kcl (75 mM), $MgCl_2$ (3 mM) and DTT (10 mM) at room temperature. Additional sample solutions of mRNA fusions **9** and **13** were prepared which contained vanadyl ribonucleside complex (VRC; 20 mM) to inhibit ribonuclease activity. Aliquots were taken after 0, 5, 15, 30, 60, 120 min and 24 h, and analyzed by electrophoresis on 4 – 12 % NuPage polyacrylamide gels (Novex) followed by exposure on phosphorimager screen. The relative amounts of remaining fusion were plotted against incubation time and half-lives were graphically extracted from the resulting curves (Table 1).

We thank L. Bloom for technical assistance and M. Matteucci, M. Wright, D. Lipovsek for suggestions on the manuscript.

[1] G. P. Smith, *Science* **1985**, *228*, 1315.
[2] R. W. Roberts, *Curr. Opin. Chem. Biol.* **1999**, *3*, 268.
[3] J. Hanes, A. Plückthun, *Curr. Top. Microbiol. Immunol.* **1999**, *243*, 107.
[4] L. C. Mattheakis, R. R. Bhatt, W. J. Dower, *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 9022.
[5] M. He, M. J. Taussig, *Nucleic Acids Res.* **1997**, *25*, 5132.
[6] J. Hanes, A. Plückthun, *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 4937.
[7] J. Hanes, L. Jermutus, S. Weber-Bornhauser, H. R. Bosshard, A. Plückthun, *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 14130.
[8] M. G. Cull, J. F. Miller, P. J. Schatz, *Proc. Natl. Acad. Sci. USA* **1992**, *89*, 1865.
[9] N. Doi, H. Yanagawa, *FEBS Lett.* **1999**, 227.
[10] K. FitzGerald, *Drug Discov. Today* **2000**, *5*, 253.
[11] Y. Liu, S. Saha, E. Haggard-Ljungquist, *J. Mol. Biol.* **1993**, *231*, 361.
[12] G. Lindahl, *Virology* **1970**, *42*, 522.
[13] R. W. Roberts, J. W. Szostak, *Proc. Natl. Acad. Sci. USA* **1997**, *94*, 12297.
[14] N. Nemoto, E. Miyamoto-Sato, Y. Husimi, H. Yanagawa, *FEBS Lett.* **1997**, *414*, 405.
[15] R. Liu, J. E. Barrick, J. W. Szostak, R. W. Roberts, *Methods Enzymol.* **2000**, *318*, 268.
[16] M. Kurz, K. Gu, P. A. Lohse, *Nucleic Acids Res.* **2000**, *28*, e83.
[17] G. Cho, A. D. Keefe, R. Liu, D. S. Wilson, J. W. Szostak, *J. Mol. Biol.* **2000**, *297*, 309.
[18] T. P. Hopp, K. S. Prikett, C. Price, R. T. Libby, C. J. March, P. Cerretti, D. L. Urdal, P. J. Conlon, *Biotechnology* **1988**, *6*, 1205.
[19] T. G. Schmidt, J. Koepke, R. Frank, A. Skerra, *J. Mol. Biol.* **1996**, *255*, 753.
[20] U. Pieles, U. Englisch, *Nucleic Acids Res.* **1989**, *17*, 285.
[21] T. Mega, N. Nakamura, T. Ikenaka, *J. Biochem. (Tokyo)* **1990**, *107*, 68.
[22] J. Smith, E. V. Anslyn, *Angew. Chem.* **1997**, *109*, 1956; *Angew. Chem. Int. Ed. Engl.* **1997**, *36*, 1879.

Received: December 14, 2000 [F 169]