FEBS Letters 539 (2003) 24–28 FEBS 27064

Selection based on the folding properties of proteins with ribosome display

Tomoaki Matsuura[1], Andreas Plückthun*

Biochemisches Institut, Universität Zürich, Winterthurerstr. 190, CH-8057 Zürich, Switzerland

Received 7 January 2003; revised 7 February 2003

First published online 24 February 2003

Edited by Giovanni Cesareni

Abstract Ribosome display is a powerful tool for selecting and evolving protein functions through ligand-binding. Here, this in vitro system was used to perform selection based on the folding properties of proteins, independent of specific ligand-binding. The selection is based on two properties of misfolded proteins: (1) increased sensitivity to proteolysis and (2) greater exposure of hydrophobic area. By targeting these properties, we show that compactly folded and soluble proteins can be enriched over insoluble and random coil proteins. This approach may be especially useful for selection and evolution of folded proteins from random sequence libraries.

Key words: Ribosome display; In vitro; Protein folding

1. Introduction

The primordial evolution of proteins, ultimately resulting in highly sophisticated structures with finely tuned functions, has remained largely mysterious. It is clear that the evolutionary fine-tuning in cells must have gone through cycles of mutation and selection, but the early emergence of folded structures remains enigmatic. Proteins must obviously fulfill their functional requirements, which frequently involve specific binding events, but must also be highly soluble, stable, efficiently folded and resistant to cellular proteases to result in reasonable half-lives. It is also noteworthy that the number of different folds which can be predicted to be encoded in all genomes is rather limited [1]. It thus appears that a major part of the diversity of natural proteins has exploited the diversification of existing folds, but it is not clear whether many more folds may be physically possible.

There is great interest in generating novel proteins [2,3], and the challenge is to establish a system capable of obtaining such molecules. Our desire is to develop a system that can mimic the course of natural protein evolution, can perform rounds of selection and diversification rapidly, and is capable of applying selection pressures for function, solubility and protease resistance. Our goal was to carry out these events completely in vitro, in order to have access to much larger libraries than would be possible by transforming libraries into cells. We have previously developed a method for selecting proteins for their binding properties by using ribosome display [4] (Fig. 1A). Here we report the development of a selection system for solubility and protease resistance of proteins, independent of particular ligand-binding properties, by using the ribosome display technology.

2. Materials and methods

2.1. Preparation of plasmids and mRNA

Plasmid pRD-n1n2_2 is a derivative of pQE30 (Qiagen) and part of the construct is schematically shown in Fig. 1B. Genes encoding the proteins to be displayed were polymerase chain reaction (PCR) amplified to introduce *Nco*I and *Bam*HI sites at each end. The PCR products were then digested and ligated into the vector pRD-n1n2_2 using the *Nco*I/*Bam*HI restriction sites. The ligation products were directly PCR amplified with oligonucleotides SDA-pqe (AGACCACAACGGTTTCCCTCTAGAAATAATTTTGTTTAACTTTAAGAAAGAGGAGAAATTAACTATGAGA) and pqe1-(GATCTATCAACAGGAGTCCAAGCTCA). The products were subsequently PCR amplified again with oligonucleotides T7B (ATACGAAATTAATACGACTCACTATAGGGAGACCACAACGG) and T3te_pD-(CGGCCCACCCGTGAAGGTGAGCC) to introduce the 5′- and 3′-stem loops and to make the final construct (Fig. 1B) for in vitro transcription. All PCRs were performed with Taq polymerase (Gibco). PCR products were directly used for in vitro transcription using T7 RNA polymerase (New England Biolabs) and purified as described previously [5].

2.2. In vitro translation for ribosome display and gel filtration for the purification of ternary complexes

In vitro translation for ribosome display was carried out essentially as previously described [5]. Following a 7 min translation at 37°C, the reaction was stopped by five-fold dilution with ice-cold wash buffer (WBK; 50 mM Tris–acetate, pH 7.5; 150 mM NaCl; 50 mM magnesium acetate; 0.5 M KCl) with 2.5 mg/ml heparin. After centrifugation at 20 000 × *g* for 5 min, the supernatant, which contains ternary complexes of mRNA, encoded protein and ribosomes, was used for subsequent experiments. When necessary, 300 μl of translation solution was applied to a gel filtration column to separate the intact ternary complexes from other proteins (proteases) and to exchange the buffer [6].

2.3. Ribosome display selection with hydrophobic interaction chromatography (HIC)

For the HIC selection, the translation reaction was stopped by five-fold dilution with ice-cold WBK. The buffer was then exchanged to WBK with desired KCl concentrations by gel filtration. The eluate (300 μl) was then added to a 100 μl bed volume of HIC beads (butyl-, octyl-, phenyl-agarose; Pharmacia). After shaking for 1 h at 4°C,

*Corresponding author. Fax: (41)-1-635 5712.
E-mail address: pluecktthun@bioc.unizh.ch (A. Plückthun).

[1] Present address: Symbiotic Engineering Laboratory, Department of Bioinformatic Engineering, Graduate School of Information Science and Technology, Osaka University, 2-1 Yamadaoka, Suita, Osaka 565-0871, Japan.

Abbreviations: HIC, hydrophobic interaction chromatography; PCR, polymerase chain reaction; RT, reverse transcription

0014-5793/03/$22.00
doi:10.1016/S0014-5793(03)00178-9

beads were removed and the supernatant was applied to a gel filtration column to exchange the buffer to WBKT (WBK with 0.25 M KCl and 0.1% Tween 20). The eluate was mixed with BSA (final concentration of 5 mg/ml), and 0.1 μg of anti-tetra-his-tag antibody (Qiagen) was added. After shaking for 1 h at 4°C, 20 μl of protein-L agarose (Santa Cruz Biotechnology), which had been washed with WBKT, was added and the solution was shaken for another 1 h. The beads were then washed five times with 1 ml of WBKT, with a shaking period of 5 min in between the washing steps. mRNA still bound to the beads was eluted with elution buffer (50 mM Tris, pH 7.6, 150 mM NaCl, 20 mM EDTA, 50 μg/ml *Saccharomyces cerevisiae* RNA). mRNA was purified, reverse transcribed and PCR amplified as previously described [5]. Reverse transcription (RT) was carried out with the oligonucleotide dN2_rd2 (CGGTTGCGGTACCCATAG) and PCR with oligonucleotides SDA-pqe and dN2_rd2.

2.4. Ribosome display selection by proteolysis

For proteolytic selection, translation was carried out as described in Section 2.3. After complexes were stabilized by dilution with WBK, 10 μl of trypsin (Sigma) was added to 300 μl of the reaction mixture to give a series of appropriate concentrations (see Fig. 5) and incubated for 1 h at 4°C. The samples were then centrifuged at 20 000 × *g* for 5 min and applied to a gel filtration column to exchange the buffer to WBKT and to separate proteases from the ternary complexes. The complexes were selected and the recovered mRNA was RT-PCR amplified as described in Section 2.3.

When both HIC and proteolytic selections were combined, first the ternary complexes were treated with proteases as described above, and the buffer was exchanged to WBK with 3 M KCl by gel filtration, which simultaneously removes the proteases. The HIC selection was then performed on the eluate.

3. Results

3.1. Design of the constructs used for ribosome display

In the construct used for ribosome display, the protein-coding sequences consist of two parts: the N-terminal part, which encodes the protein to be selected, and the C-terminal part, which serves as a spacer sequence (Fig. 1B). The spacer sequence provides sufficient distance between the displayed protein and the ribosome, allowing the protein to fold into its correct conformation [5]. Up to now, either a portion of the sequence from gene III from the phage M13mp19 [5] or a portion of the helical region of *tol*A [6] or the extended region of *ton*B [7] from *Escherichia coli* have been used as a C-terminal tether. These spacers are very likely to be unstructured in the context of ternary complexes. However, as we want to perform selections based on the folding properties of the displayed protein, the spacer sequence should constitute a well-folded protein, and thus aforementioned sequences cannot be used for this study, as the spacer has to be more stable than the displayed proteins. We therefore chose amino acids T20–V109 of protein D (pD), a structured part of the capsid pro-

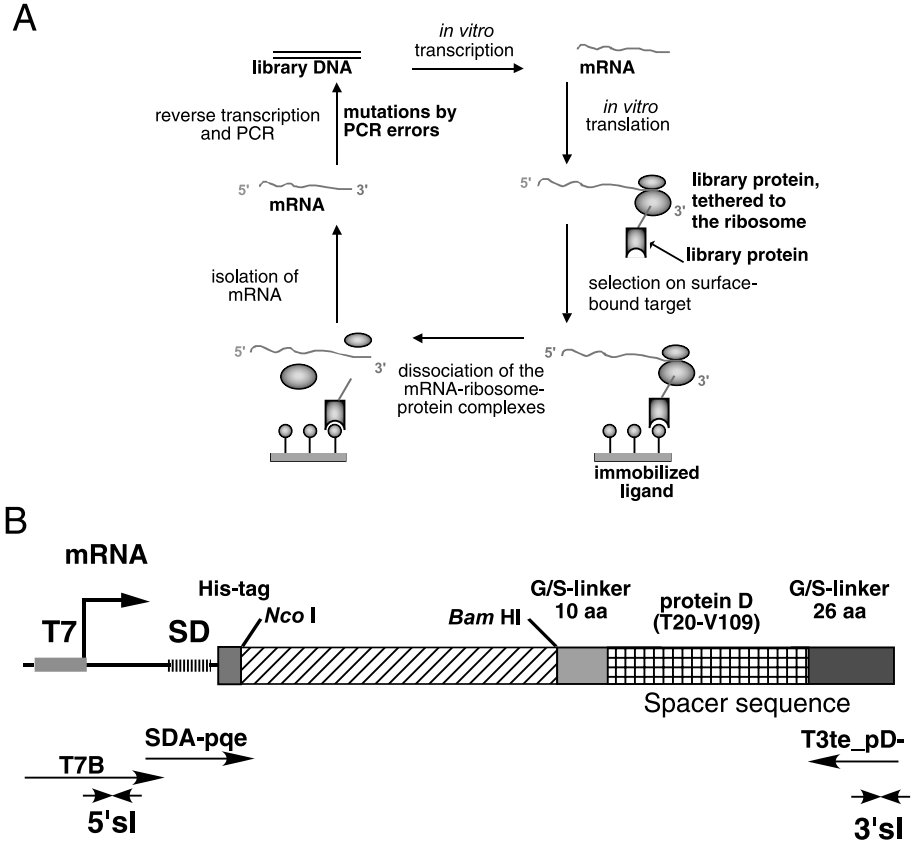

Fig. 1. A: Principle of ribosome display selection. A library of DNA molecules encoding the proteins of interest is transcribed and translated in vitro. The resulting mRNA lacks a stop codon, giving rise to ternary complexes of mRNA-ribosome-encoded protein. These are directly used for an affinity selection. The mRNA from the bound complexes is eluted and purified. RT-PCR can introduce mutations and yields a DNA pool enriched for binders that can be used for the next iteration. B: DNA construct used for *E. coli* ribosome display described here. The T7 promoter is followed by a Shine–Dalgarno sequence (SD), an N-terminal his-tag and the protein of interest (here RC1/RC2, FB, or pD). The protein of interest, which is cloned with *Nco*I/*Bam*HI restriction sites, is followed by a 10 amino acid (aa) Gly/Ser linker, T20-V109 of pD and a 26 aa Gly/Ser linker. Sequences encoding stem–loop structures are present at both ends (5′-sl and 3′-sl).

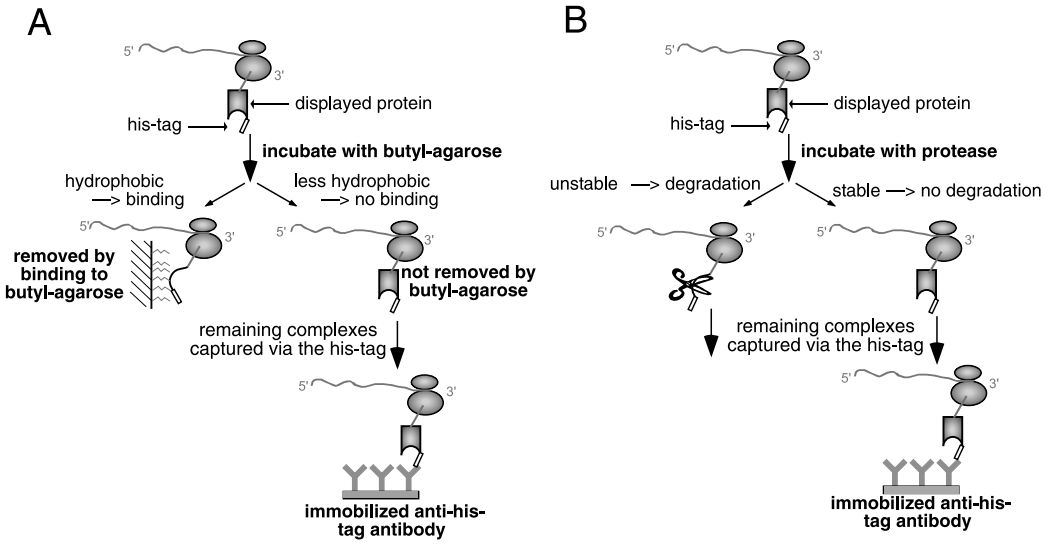

Fig. 2. A: Scheme of the selection based on protein hydrophobicity. Ternary complexes are generated from mRNAs encoding the proteins of interest. Complexes are then incubated with beads possessing hydrophobic groups in the presence of high concentration of salt. Those with greater hydrophobicity will bind the beads and those which remain unbound are recovered using the specific interactions between the N-terminal his-tag, anti-his-tag antibody and immobilized protein-L. B: Scheme of the selection based on protease resistance. After the complexes are generated, they are incubated with protease. Those that are sensitive to proteases are degraded, whereas those having a compact structure are not. Ternary complexes displaying undigested proteins are recovered via the his-tag.

tein from phage Lambda [8], as a spacer sequence. pD has been shown to be stable, soluble and very well expressed in *E. coli* and monomeric in solution [9]. T20–V109 of pD has also been shown to be more resistant to proteolytic digestion than the full-length molecule, while retaining the biophysical properties of the full-length construct (Forrer, P. and Plückthun, A., unpublished data). The pD sequence is followed by a 26 amino acid residue Gly/Ser linker, which is presumed to remain in the ribosomal tunnel, when ternary complexes are formed.

The efficacy of ribosome display has been successfully demonstrated in a variety of selection schemes [6,7] to select and further evolve the functional properties of proteins. The stability of a protein also can be improved using ribosome display [10] by exploiting specific ligand-binding properties for selection. We wanted to further adapt the ribosome display technology to select proteins based on their folding properties, but independent of their ligand-binding properties. Two concepts are schematically shown in Fig. 2A,B.

In this study we used three types of model proteins that have very different folding properties. We tested two proteins that are soluble but exhibit random coil formation (RC1 and RC2) as shown by circular dichroism spectroscopy, one protein that forms amyloid-like fibrils (FB), and the structured portion S16–V109 of a compact native model protein, pD, a major capsid protein from bacteriophage Lambda [8,9]. RC1, RC2 and FB were found in a protein library that was constructed previously [11] by combining de novo designed secondary structure modules, and the amino acid sequences are shown in Table 1. The properties of these proteins are described in detail elsewhere [11].

3.2. *Effect of KCl concentration on the stability of the ternary complexes*

As the binding of proteins to hydrophobic ligands is dependent on salt concentration, the ribosomal complexes must be stable at high concentrations of salts to remain selectable (Fig. 2A). The most commonly used salt for HIC is $(NH_4)_2SO_4$; however, ternary complexes precipitated at concentrations above 2 M. We thus tested the stability of ternary complexes in different KCl concentrations. For this purpose, we incorporated a histidine-tag (his-tag) at the N-terminus of each construct to capture stable complexes (Fig. 1B), by using an anti-his-tag antibody, which is itself captured by protein-L agarose beads. We found that the amount of recovered mRNA was very low in the absence of KCl, but between 0.5 M and 3 M KCl, the complexes were stable (Fig. 3). Interestingly, even the mRNA of the protein that forms amyloid-like FB could be recovered, suggesting that when the ternary complexes are formed, the protein molecules are sufficiently segregated by ribosomes, which consist mostly of RNA, to prevent aggregation.

Table 1
Amino acid sequences of the model proteins in one letter code

RC1	MRRSEMMKRAGSPIARAAGSERKFFEKFMKKGSRTVSIVGSEKRIAQRMFKKGSSSSGSPRKIAEFMQEGSRVVVITGSRTLTVSGSERQAIEQLIRK-GSPAAARRGSEQALFRRAARAGSPKKALQALAARGS
RC2	MRRSEEKAIRRMAAKGSRVVVFSGSRYVTVYGSPQKAAKKFIQAGSEAFLKAGSEMQEFRGSPLERIAGSRVYLVFGSERRAMEKALREGSPAFRKSG-SPLFKALGSPELMRKGSPIMAAAGSRSIYVSGS
FB	MRRSTSSGSRGIVIVGSRTISVVGSRSLVYGSRVLTISGSRLVLVLGSRSFYLYGSDDS

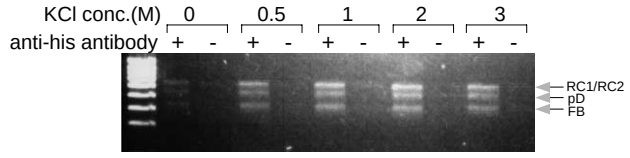

Fig. 3. Effect of KCl concentrations on the stability of ternary complexes. After in vitro translation, ternary complexes were placed in WBK with different KCl concentrations. The buffer was then exchanged to WBKT. Affinity selection of his-tagged protein was performed in the presence (+) or absence (−) of anti-his-tag antibody. The mRNA from complexes still bound to the solid support was isolated and RT-PCR amplified. Products were analyzed by agarose gel electrophoresis. The PCR fragments encoding the three types of proteins are indicated. It can be seen that all complexes are stable and can be enriched, except in the absence of KCl.

3.3. Selection based on protein hydrophobicity

We then tested whether proteins with unfavorable folding properties could be preferentially removed using HIC (Fig. 2A). Ternary complexes were generated from the mRNAs of three different model proteins (RC1/RC2, FB and pD), which were mixed in equal proportion. Complexes were then incubated with agarose beads displaying different hydrophobic ligands (butyl, octyl and phenyl) in the presence of 3 M KCl. Those complexes that remained unbound were captured via their N-terminal his-tag, and mRNA pools were amplified by RT-PCR. We found that butyl-agarose very efficiently removed the fibril forming protein and partially removed those exhibiting random coils, whereas pD was retained, and therefore enriched (Fig. 4A). When using octyl or phenyl as the hydrophobic ligand, however, there were very few mRNA molecules recovered of any protein, regardless of the KCl concentration, suggesting that most of the ternary complexes were bound to these more hydrophobic ligands.

We also tested the effect of KCl concentration on the enrichment of pD using butyl-agarose (Fig. 4B). The enrichment

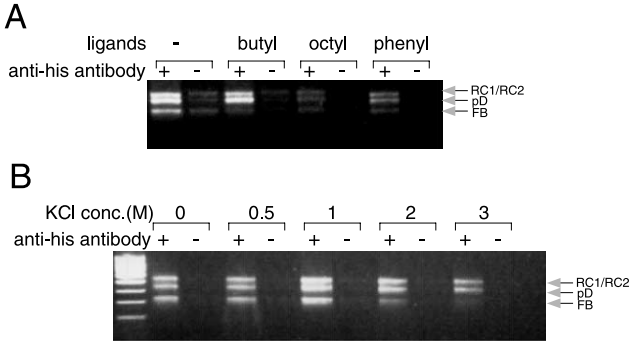

Fig. 4. Selection based on the hydrophobicity of the displayed protein. A: Ternary complexes of the model proteins were incubated with (+) or without (−) butyl-, octyl- or phenyl-agarose beads. After the incubation, beads were removed and the affinity selection of his-tagged protein was performed in the presence (+) or absence (−) of anti-his-tag antibody. RT-PCR products were analyzed by agarose gel electrophoresis. It can be seen that a large fraction of all proteins are removed by octyl- and phenyl-agarose, while FB is removed selectively with butyl-agarose. B: Influence of KCl concentrations on the removal of proteins with unfavorable behavior. Ternary complexes were incubated with butyl-agarose beads at different KCl concentrations. Affinity selection of his-tagged protein was performed in the presence (+) or absence (−) of anti-his-tag antibody. RT-PCR products were analyzed by agarose gel electrophoresis. The selectivity is best with 3 M KCl.

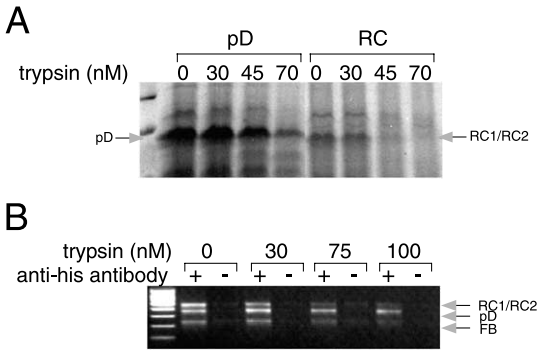

Fig. 5. Selection based on the protease resistance of the displayed protein. A: Ternary complexes of pD and RC1/RC2 were generated using [^{35}S]methionine. Complexes were incubated with different concentrations of trypsin and analyzed by SDS–PAGE. pD is more stable than RC1/RC2. B: Ternary complexes of the model proteins were incubated with different concentrations of trypsin. After the removal of trypsin, affinity selection of his-tagged protein was performed in the presence (+) or absence (−) of anti-his-tag antibody. RT-PCR products were analyzed by agarose gel electrophoresis. Enrichment of pD is observed.

of pD was more pronounced at higher concentrations of KCl, as the hydrophobicity increases under these conditions. Therefore, it appears that the selection pressure for the removal of hydrophobic molecules can be adjusted by changing the KCl concentration.

3.4. Selection based on protease resistance

Misfolded or unstructured proteins are not sufficiently compact to be protected from proteolytic digestion. It has been shown that this type of selection pressure can be applied to select for compactly folded proteins using phage display [12–14]. We aimed to apply the same concept in a fully in vitro system by using ribosome display (Fig. 2B).

We first wanted to test the protease resistance of three types of model proteins (RC1/RC2, FB and pD). After complexes were generated, they were treated with proteases and were analyzed by sodium dodecyl sulfate–polyacrylamide gel electrophoresis (SDS–PAGE) (Fig. 5A). FB, however, was not able to migrate into the gel, possibly due to its fibril formation [11], and was thus omitted from this analysis. We could see that the RC1/RC2 mixture was mostly degraded by trypsin at concentrations above 45 nM, whereas pD still remains fully intact (Fig. 5A). Gel filtration experiments showed that undigested pD remained connected to the ribosome (data not shown), indicating that the ribosome itself is sufficiently stable against trypsin. However, when using chymotrypsin, obvious differences in protease resistance between pD and the RC1/RC2 mixture were not observed (data not shown), and therefore further experiments were only continued with trypsin. We then investigated whether the complexes treated with trypsin could be recovered according to the protease resistance of the displayed protein. Starting from a mixture of mRNAs encoding the three types of model proteins, we saw a clear enrichment of pD with increasing concentrations of trypsin (Fig. 5B). These results indicate that a selection based on the protease resistance of the displayed protein can be achieved with ribosome display.

3.5. Combining HIC and proteolytic selections

We then investigated whether the two selection methods,

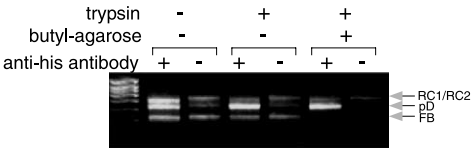

Fig. 6. Effect of combining the selections based on surface hydrophobicity and protease resistance of the displayed protein. Ternary complexes of the model proteins were incubated either with (+) or without (−) 70 nM trypsin, followed by an incubation with (+) or without (−) butyl-agarose beads. Affinity selection of his-tagged protein was performed in the presence (+) or absence (−) of anti-his-tag antibody. RT-PCR products were analyzed by agarose gel electrophoresis. Combining the two selection criteria enhances the enrichment of pD.

HIC and proteolysis, could be combined. After in vitro translation using the mixed mRNA of the model proteins, ternary complexes were first treated with protease, followed by an incubation with butyl-agarose beads. Those that survived both of the selection steps were recovered and RT-PCR amplified. We could see almost a complete enrichment of pD when both selections were combined (Fig. 6), indicating that these two can be applied sequentially.

4. Discussion

In this study, we have shown that selection pressures based on the folding properties of displayed proteins, but independent of their ligand-binding properties, can be applied by using ribosome display; the HIC selection efficiently removes aggregation-prone molecules, while protease digestion removes proteins forming largely random coil. Ribosome display is capable of selecting functional proteins in vitro, and by combining the strategies described here, we can now mimic a part of natural protein evolution by applying three different selection pressures for function, solubility and protease resistance.

We also found that the ternary complexes can remain soluble, regardless of the intrinsic solubility of the displayed protein. This property may be used to alter the folding behavior of proteins that are not produced in functional form in *E. coli* due to their aggregation. Waldo et al. have reported a system which is capable of screening soluble proteins in *E. coli* [15], based on the idea that solubly expressed proteins should fluoresce when fused with green fluorescent protein (GFP), while aggregating proteins will prevent GFP from folding, and thus block fluorescence. However, this screening is carried out in vivo and thus not as easily applicable to very large libraries. We can now use this combined selection strategy to enrich well-folded molecules in vitro from artificial libraries, e.g. those built from secondary structure modules [11]. Furthermore, the present strategy can be implemented before or after a selection for ligand-binding to filter out aggregation-prone molecules from the pool of binders.

Acknowledgements: We thank Andreas Ernst for helpful discussions throughout the course of this work, and Drs. Casim Sarkar, David Zechel, and Patrik Forrer for critical reading of the manuscript. The authors also thank members of the Plückthun group for detailed discussions on the ribosome display technology. This work was supported by the National Center of Competence in Research Structural Biology and the Forschungskredit der Universität Zürich.

References

[1] Chothia, C. (1992) Nature 357, 543–544.
[2] Keefe, A.D. and Szostak, J.W. (2001) Nature 410, 715–718.
[3] Minard, P., Scalley-Kim, M., Watters, A. and Baker, D. (2001) Protein Sci. 10, 129–134.
[4] Hanes, J. and Plückthun, A. (1997) Proc. Natl. Acad. Sci. USA 94, 4937–4942.
[5] Hanes, J., Jermutus, L. and Plückthun, A. (2000) Methods Enzymol. 328, 404–430.
[6] Amstutz, P., Pelletier, J.N., Guggisberg, A., Jermutus, L., Cesaro-Tadic, S., Zahnd, C. and Plückthun, A. (2002) J. Am. Chem. Soc. 124, 9396–9403.
[7] Hanes, J., Schaffitzel, C., Knappik, A. and Plückthun, A. (2000) Nat. Biotechnol. 18, 1287–1292.
[8] Yang, F. et al. (2000) Nat. Struct. Biol. 7, 230–237.
[9] Forrer, P. and Jaussi, R. (1998) Gene 224, 45–52.
[10] Jermutus, L., Honegger, A., Schwesinger, F., Hanes, J. and Plückthun, A. (2001) Proc. Natl. Acad. Sci. USA 98, 75–80.
[11] Matsuura, T., Ernst, A. and Plückthun, A. (2002) Protein Sci. 11, 2631–2643.
[12] Sieber, V., Plückthun, A. and Schmid, F.X. (1998) Nat. Biotechnol. 16, 955–960.
[13] Riechmann, L. and Winter, G. (2000) Proc. Natl. Acad. Sci. USA 97, 10068–10073.
[14] Finucane, M.D., Tuna, M., Lees, J.H. and Woolfson, D.N. (1999) Biochemistry 38, 11604–11612.
[15] Waldo, G.S., Standish, B.M., Berendzen, J. and Terwilliger, T.C. (1999) Nat. Biotechnol. 17, 691–695.