reprinted from:
A. Plückthun, C. Schaffitzel, J. Hanes and L. Jermutus, In vitro selection and evolution of proteins. Adv. Protein Chem. 55, 367-403 (2000).

IN VITRO SELECTION AND EVOLUTION OF PROTEINS

By ANDREAS PLÜCKTHUN, CHRISTIANE SCHAFFITZEL, JOZEF HANES, and LUTZ JERMUTUS

Biochemisches Institut, Universität Zürich, Switzerland

I. Introduction 367
A. *In Vivo* versus *In Vitro* Selections 369
II. The Key to *In Vitro* Protein Evolution: Cell-free Translation 372
A. Increasing *In Vitro* Translation Yields 372
B. Increasing the Fraction of Functional Molecules 373
III. *In Vitro* Selection Strategies 374
A. SELEX 374
B. Ribosome Display 376
C. RNA-peptide Fusion and *In Vitro* Virus 383
D. Selection for Enzymatic Activity *In Vitro* 385
E. Water in Oil Emulsions for Binding Selections (STABLE) 388
IV. Applications of Ribosome Display 388
A. Selection from Libraries with or without Concurrent Sequence Evolution 389
B. Directed Evolution of Binding Affinity and Stability 392
V. Perspectives of Directed *In Vitro* Evolution 398
References 399

I. Introduction

Since the advent of recombinant DNA technology, the engineering of proteins for improved binding specificities, ligand affinities, and stability has become commonplace. Both hypothesis-based "rational" and combinatorial approaches exist to address these tasks. It is the technical advances in the latter and the apparent current limitations in the former which gave rise to this volume.

To rationally (re)design a protein requires detailed structural and, in the case of enzymes, mechanistic information. The technical problem of producing the gene for virtually any protein with any sequence has been solved by gene synthesis and site-directed mutagenesis, but the knowledge of how sequence changes affect protein expression, function and biophysical properties lags far behind. Currently, the predictive accuracy of even the most sophisticated structure-based engineering approaches is often still insufficient to produce the desired design effects without additional experimentation (Dougan *et al.*, 1998; Yelton *et al.*, 1995). Nevertheless, progress has been achieved and is certain to continue (Rubingh, 1997; Hellinga, 1997). The fundamental problem of

ADVANCES IN PROTEIN CHEMISTRY, Vol. 55

0065-3233/00 $35.00

predictions is the multitude of configurations of very similar energy. It is very difficult to decide which way the balance will tip and to predict whether a large number of possible small movements will or will not result in a larger overall conformational change. If no structural and mechanistic information is available, the effects of mutational changes are almost fully unpredictable.

In contrast, for combinatorial approaches the challenge lies in the technology—to actually achieve a Darwinian evolution in a reasonable time. Although the global concepts have been clear for a long time, only recently have the tools become available to exploit this approach in practice. Currently, it appears that combinatorial and evolutionary methods are in the lead for actually improving a particular molecule in practice.

It is useful at this point to define the differences between combinatorial and evolutionary strategies. The underlying principle of the *combinatorial* approach is selection for the desired property from a pool of diverse molecules (a single-pot or "constant" library, which does not change later). In this case, the accessed sequence space equals the functional library size. In the *evolutionary* approach, on the other hand, a given molecule (one starting sequence) or a library (many starting sequences) is continuously diversified, to elicit improved or even entirely novel functions by an iterative process of diversification and selection. Therefore, the accessed sequence space in an evolution experiment is far greater than the initial library size.

Directed evolution mimics the natural process by which protein variants arise and are tested for their fitness in living systems in a series of "generations"—cycles of diversification and selection. A particularly instructive example (which occurs much more rapidly than the phylogenetic evolution of proteins) is the somatic hypermutation of antibodies. During the secondary immune response, antibody V genes undergo point mutations at a frequency of about 10^{-3} per base pair per generation (Berek and Milstein, 1987; Allen *et al.*, 1988). Concomitant with this rapid mutational process, selection of B cells with high-affinity receptors for the immunizing antigen leads to a 10- to 100-fold increase in the average antibody affinity (Berek and Milstein, 1987). This "affinity maturation" can be modeled as an adaptive walk on a rugged sequence landscape (Macken and Perelson, 1989), and it was found that a small number of single mutations is necessary to reach a local optimum in the fitness landscape.

To date, a number of well-established strategies exist to select protein-ligand interactions. These can be exploited to *identify* binding molecules (combinatorial approach) or iteratively *improve* them (evolutionary approach). Additionally, if the binding interaction is restricted to the native

state, the selection can be used to select for the quality of the protein, including expression yield (Jung *et al.*, 1999), folding kinetics and thermodynamic stability (Ruan *et al.*, 1998; Spada *et al.*, 1998), resistance to proteolysis (Sieber *et al.*, 1998; Kristensen and Winter, 1998) or stability in an extreme nonphysiological environment (Kuchner and Arnold, 1997; Jung *et al.*, 1999; Schmidt-Dannert and Arnold, 1999).

Examples of such protein-protein interaction selection systems are phage display (Smith, 1985; Winter *et al.*, 1994), display on other viruses (Kasahara *et al.*, 1994), bacterial surface display (Georgiou *et al.*, 1993; Daugherty *et al.*, 1999), yeast display (Kieke *et al.*, 1997; Boder and Wittrup, 1997), the yeast two hybrid system (Fields and Song, 1989; Chein *et al.*, 1991), and protein-fragment complementation assays (Pelletier *et al.*, 1998). These methods all contain a necessary *in vivo* step, which has a number of disadvantages that will be discussed in the following sections.

Ribosome display (Hanes and Plückthun, 1997) is the first method for screening and selecting functional proteins that is performed entirely *in vitro,* thus circumventing many of the drawbacks of *in vivo* systems. Here, we present the principles underlying ribosome display and some of its applications for generating high affinity and high stability antibodies from given starting molecules of complex libraries and summarize related *in vitro* selection technologies. We also compare *in vitro* selection to *in vivo* methods.

In ribosome display, the physical link between genotype and phenotype is accomplished by mRNA–ribosome–protein complexes, which are directly used for selection. If a library of different mRNA molecules is translated, a protein library results in which each protein is produced from its "own" mRNA and remains connected to it. Since these complexes of the proteins and their encoding mRNAs are stable for several days under the appropriate conditions, very stringent selections can be performed. As all steps of ribosome display are carried out *in vitro,* reaction conditions of the individual steps can be tailored to the requirements of the protein species investigated, as well as the objectives of the selection or evolution experiment. Application of ribosome display has produced scFv fragments of antibodies with affinities in the picomolar range from libraries prepared from immunized mice (Hanes *et al.*, 1998) and more recently from a naive, completely synthetic library (Hanes *et al.*, 2000), and has been used to evolve improved off-rates and stability (Jermutus *et al.*, 2000).

A. In Vivo *versus* In Vitro *Selections*

All *in vivo* methods have in common that the library usually encoded on a plasmid or phage replicon must be transformed into cells, either

bacteria or yeast. These microorganisms then express the protein for an intracellular interaction screen, such as the yeast two hybrid system (Fields and Song, 1989; Chien *et al.*, 1991) or the protein-fragment complementation assay (Pelletier *et al.*, 1998; 1999). Alternatively, bacteria or yeast cells display the protein on their surface (Georgiou *et al.*, 1993; Daugherty *et al.*, 1999; Kieke *et al.*, 1997; Boder and Wittrup, 1997). Finally, the bacteria may be transformed with the library in order to produce phages (e.g., phage display with filamentous or λ phages) that then carry the protein on their surface. Obviously, the library size is determined by the transformation frequency, and typically *Escherichia coli* libraries of 10^{10} to 10^{11} present an upper limit (Dower and Cwirla, 1992). To create libraries even with this size involves significant labor. Importantly, after each *in vitro* randomization step, a new library has to be created and transformed. Libraries screened by yeast display (Boder and Wittrup, 1997) and the yeast two hybrid system (Fields and Song, 1989) are even smaller, due to the generally lower transformation efficiency in yeast.

With ribosome display and other *in vitro* selection systems (see below) no transformation is necessary. Therefore, it is possible to assemble libraries *in vitro* and retain their very large size. Furthermore, it is possible to screen these protein and peptide libraries with 10^{11} or more members, with a new library of point mutants at every generation in an evolution experiment. An increase in library size improves the chance to select for the desired function and in addition increases the diversity of molecules selected. Lancet *et al.* (1993) estimated the relationship between the library size and the best affinity of a member in the library that could be selected. The prediction was that increasing a library from 10^8 to 10^{12} sequences will increase the affinity of the best selected binder up to 300-fold. Therefore, with the possibility of screening very large libraries by *in vitro* selection technology, it becomes more likely that a larger variety of high affinity binders with the desired function are selected. Using ribosome display, it has indeed been possible to select and evolve high affinity antibodies with dissociation constants as low as 80 pM from protein libraries (Hanes *et al.*, 1998; Hanes *et al.*, 2000) in a short time. In this approach, very large libraries are easily accessible because they do not have to be cloned and transformed into cells but can be rapidly assembled *in vitro*.

In vivo, a pre-selection due to the host environment cannot be avoided. Growth disadvantage or even toxicity can lead to a loss of potential candidates. More rapidly growing library members can become over-represented in the culture despite the fact that they are not specifically enriched by the actual selection. Furthermore, folding, transport, aggre-

gation, and proteolytic degradation can often not be controlled effectively by the investigator in an *in vivo* environment, and the final application of the selected molecule may be envisioned for an environment quite different from that of *in vivo* selection. Another important point is that cells are complex genetic entities, and they often find many ways to survive or evade the selection pressure that differ from the ones desired by the investigator. Mutations or recombinations within the plasmid or within the host genome can provide an easy solution for the cell to circumvent the selection pressure.

These undesired selection pressures are substantially reduced *in vitro,* and the translation conditions can be optimized for the protein to be displayed on a case-by-case basis. Ribosome display can also be easily combined with *in vitro* mutagenesis techniques such as mutagenic PCR (Cadwell and Joyce, 1992), DNA shuffling (Stemmer, 1994), the staggered extension process (Zhao *et al.,* 1998) or other recombination-based methods in an evolution experiment (see Section IV, B). Also, if nonproofreading DNA polymerases are used for ribosome display, a diversification of the initial library during the selection cycles will be observed automatically due to mutations introduced during the many PCR steps at the end of each selection cycle. Thereby, the sequence space sampled is much larger than the initial size of the library. In principle, the quality of the pool is iteratively improved, since only proteins that survived the first selection will be used for further diversification. During all subsequent selections the mutated proteins have to compete with their progenitors.

In contrast, if a diversification step needs to be included in an *in vivo* selection strategy in order to evolve the protein under investigation, either a mutator strain needs to be used (Low *et al.,* 1996) or it is necessary to repeatedly switch between the selection procedure *in vivo* (phage, bacteria, yeast) and the mutagenesis step for diversification carried out *in vitro.* The disadvantage of the former case is that mutator strains can also create unwanted mutations in the plasmid and in the host genome, while the latter is a rather laborious procedure, as after each diversification step, the newly created library has to be religated and retransformed. Consequently, only relatively few examples of protein evolution over several cycles of diversification and selection are found in the literature (e.g., Yang *et al.,* 1995; Schier and Marks, 1996; Moore *et al.,* 1997).

Ribosome display has thus two main advantages compared to *in vivo* selection systems: on the one hand *in vitro* technologies allow one to screen very large libraries, since no transformation steps are necessary. On the other hand, subsequent diversification of the library is easy and convenient and every single clone present in the library can conceivably

be evolved. In addition, working *in vitro* allows for tight control of the selection experiment at each step.

II. The Key to *In Vitro* Protein Evolution: Cell-free Translation

A basic understanding of *in vitro* translation is a prerequisite for devising and optimizing cell-free protein selection systems. *In vitro* protein synthesis, independent of its use in selection technology, has received increasing interest in recent years (reviewed by Jermutus *et al.*, 1998). This is due to improvement of protein yields, the increase in detection sensitivity of many analytical methods, and the advent of new technologies such as atomic force microscopy (AFM) (Engel *et al.*, 1999) and fluorescence correlation spectroscopy (FCS) (Eigen and Rigler, 1994; Rigler, 1995) that allow analytical work at low protein concentrations, even down to the single molecule level.

Many applications of cell-free translation rely on the correct folding of the *in vitro* expressed polypeptide into its three-dimensional structure, and this is a prerequisite for all protein selection systems that are based on *in vitro* translation. Because proteins are selected and evolved for functionality, sufficient expression and correct folding in the respective cell-free translation system are a necessity for efficient selection. An attractive advantage of using *in vitro* translations is that, at least in principle, any component of the reaction can be deliberately added or removed. To achieve any improvement in yield, however, separate consideration of both the actual translation and the folding is necessary. Even in optimized systems, however, translation yields are not similar for all globular proteins.

A. *Increasing* In Vitro *Translation Yields*

There are many hypotheses about the underlying mechanisms of differences in translation yields, and to date no cell-free translation system has been engineered that allows a high expression and quantitative folding of any given protein sequence. At least three problems need to be solved regarding total protein production. First, mRNA secondary structures can inhibit translation initiation or stall elongating ribosomes (Kozak, 1989; Yu *et al.*, 1994). This becomes especially important if RNA hairpin loops are further stabilized in nonphysiological conditions such as the relatively high Mg^{2+} concentrations typically used in a standard S30 *E. coli* translation system. However, using the degeneracy of the genetic code, this limitation can be removed by silent mutagenesis of the primary sequence. Second, RNase and protease sensitivity can reduce expression yield by removing the template RNA or the synthesized protein (see

Section III, B, 2 for discussion of *E. coli* RNases). Because the recognition sequences of these enzymes are different in each organism and are in many cases unknown, this problem might be solved by removing these enzymes from the translation system with immunoprecipitation.

A third and more general bottleneck is tRNA availability. Any cell-free expression system contains endogenous aminoacyl-tRNA-synthetases, and usually a tRNA pool from the same organism is added for the translation reaction. Because the genetic code is degenerate, the pool contains tRNAs with different anti-codons for the same amino acid. The concentrations for these different tRNAs vary, resulting in rare codons on the mRNA level. Although some reports have suggested that these codons might be important for cotranslational folding (Komer *et al.*, 1999; Thanaraj and Argos, 1996), this point remains controversial and inclusion of rare codons generally decreases protein yield. Ribosome stalling at rare codons can either trigger 10Sa-RNA-mediated proteolytic degradation (Roche and Sauer, 1999) or premature translation termination (Komar *et al.*, 1999). Simply increasing the total concentration of the whole tRNA pool would not change the molar ratios of the different tRNAs that are competing for aminoacylation at the synthetase. As a consequence, codon rarity will persist. The only way to resolve the problem lies in adding a tRNA pool with different molar rations of the tRNAs accepting the same amino acid (De Pasquale and Kanduc, 1998). *In vivo,* protein translation was found to be mediated by changes of the tRNA pool composition (Kanduc, 1997; Hentze, 1995). *In vitro* transcribed tRNAs, which can be obtained with T7 RNA polymerase, are substrates for aminoacyl-tRNA-synthetases *in vitro.* This opens the door for the rational design of the tRNA pool to be added in cell-free translation.

Moreover, it was recently shown that the deliberate removal of phosphatases from the *in vitro* translation mixture by immunoprecipitation results in increased protein yields (Shen *et al.*, 1998). As proposed earlier (Jermutus *et al.*, 1998), the elimination of one of the many causes of fast ATP and GTP depletion extends the time of synthesis, and, as a consequence, the total amount of produced protein. It should be possible to similarly remove proteases and nucleases (see above). Together with new ATP regeneration systems to keep the biochemical energy level at a high steady state (Kim and Swartz, 1999), the optimization measures mentioned should increase the level of synthesis of most proteins, and this should directly improve *in vitro* selection.

B. Increasing the Fraction of Functional Molecules

Cotranslational misfolding or aggregation reduces the active fraction of synthesized proteins and might also decrease the expression yield by

inhibiting the translating ribosome. However, correct folding can at present only be improved on a case-to-case basis, even though the use of standardized cocktails of beneficial factors is conceivable. The optimization of cell-free production of single-chain Fv antibody fragments (scFv) has been investigated in detail in an *E. coli* translation system (Ryabova *et al.*, 1997). The factors contributing to efficient folding and proper disulfide bond formation were identified. It was suggested that chaperones, mostly DnaK and DnaJ, increase the amount of soluble scFv in an *E. coli* S30 cell-free system, but do not affect the amount of functional proteins, indicating that there are soluble, misfolded species. Eukaryotic PDI (protein disulfide isomerase), a eukaryotic protein catalyzing disulfide bond formation and rearrangement, added cotranslationally in the cell-free system, increased the amount of functional antibodies. Evidence was provided that the isomerization reaction, and not the net disulfide bond formation, is the rate-limiting step for the *in vitro* folding of scFv fragments. The *in vitro* production of disulfide-containing, native molecules is thus possible, and the use of appropriate redox conditions and addition of folding catalysts can have a significant effect on the preparative production of biologically active proteins *in vitro*. However, different proteins may need other chaperone cocktails to maximize functionality.

III. *In Vitro* Selection Strategies

In the following sections, an overview of the different approaches for *in vitro* selection is provided. Although this chapter focuses on proteins, we want to briefly explain a nucleic acid selection method that formed the basis of the *in vitro* evolution of functional proteins.

A. SELEX

In 1990, Tuerk and Gold introduced a technology termed *S*ystematic *E*volution of *L*igands by *EX*ponential Enrichment (SELEX). By using SELEX, it is possible to exponentially enrich and evolve RNA ligands in multiple rounds *in vitro* from a random oligonucleotide pool. Many protein motifs and functions can be mimicked by folded RNA structures (Roberts and Ja, 1999). Currently, this method is widely employed to screen for nucleic acid ligands (aptamers) binding to numerous targets with potential applications in diagnostics and biotechnology (Osborne *et al.*, 1997). Furthermore, SELEX has been used to isolate novel nucleic acid based catalysts for a variety of reactions (Gold *et al.*, 1995).

In SELEX, multiple rounds of *in vitro* transcription of random nucleic acid pools, affinity selection, and RT-PCR are performed, thus giving rise to exponential amplification of the selected molecules. The principle underlying SELEX is schematically depicted in Figure 1. After several selection cycles, the binders can subsequently be cloned and sequenced and then characterized. In SELEX, genotype and phenotype are simultaneously represented by the same RNA molecule, since it exerts its function through its three-dimensional structure, which is in turn determined by its nucleotide sequence. The chemical and functional diversity of RNA can be further increased by addition of cofactors such as histidine (Roth and Breaker, 1998) and divalent cations (Tarasow *et al.*, 1997) to the selection.

Nevertheless, RNA molecules have some severe disadvantages as ligands, and they have been almost completely replaced by peptides and proteins in evolution. RNA is a polyanion and thus frequently selects positive charges as the target (Hermann and Patel, 2000), thereby restricting the epitopes on a protein that can be blocked. Furthermore, RNA is extremely prone to degradation by ubiquitous RNases. To actually use an aptamer in any application, the RNA has first to be stabilized by introducing stable nucleotide analogs that are obtained either by the

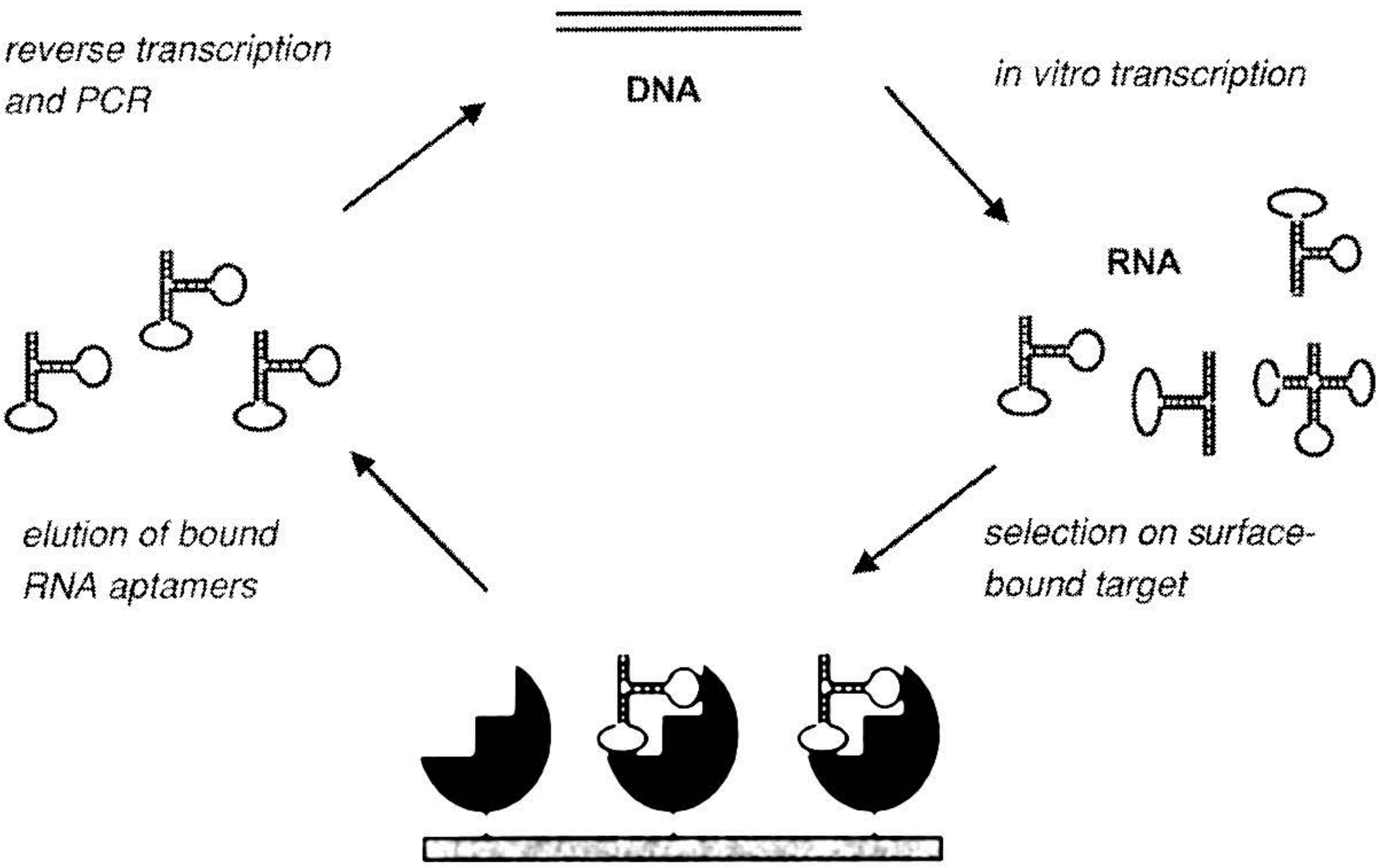

FIG. 1. SELEX. A DNA oligonucleotide pool is transcribed *in vitro*. The resulting RNA is directly used in affinity selection against an immobilized target. RNA molecules that bind (termed "aptamers") are subsequently eluted. By RT-PCR, an oligonucleotide pool enriched for binders can be regenerated and used for a new round of SELEX.

addition of phosphorothioates or the substitution of 2′-OH groups by 2′-NH_2 and 2′-F (Eaton *et al.*, 1997; Ruckman *et al.*, 1998). Such molecules cannot be synthesized by enzymes in preparative amounts, but must be prepared by large-scale synthesis. The initial (and remaining) appeal of SELEX was the very rapid generation of high affinity binders from very large initial libraries (10^{15} to 10^{16} sequences), as the RNA transcripts directly constitute the ligands.

B. Ribosome Display

In their original publication about SELEX, Tuerk and Gold (1990) already speculated that a similar approach could be adapted to protein selection. They referred to experiments describing the isolation of particular mRNAs from a pool of variants by immunoprecipitation of the nascent polypeptides present in the mRNA-ribosome-polypeptide complexes (Korman *et al.*, 1982; Kraus and Rosenberg, 1982). In fact, soon after the publication of SELEX (Tuerk and Gold, 1990) a patent application was filed (Kawasaki, 1991), proposing a similar approach to enrich peptides from libraries.

The first experimental demonstration of the ribosome display technology was the selection of short peptides from a library using an *E. coli* S30 *in vitro* translation system (Mattheakis *et al.*, 1994; Mattheakis *et al.*, 1996). The concept pursued by Mattheakis and coworkers (1994) for peptides was then used for the development of ribosome display of functional proteins by use of the *E. coli* S30 *in vitro* translation system (Hanes and Plückthun, 1997). However, for this purpose it was necessary to significantly modify and optimize the experimental conditions of ribosome display to make this technology efficient enough for the display of correctly folded, functional proteins.

1. Principle of Ribosome Display

The principle of ribosome display is depicted in Figure 2. A DNA library, encoding a polypeptide in a special ribosome display cassette (discussed in Section III, B, 2) is either directly used for coupled *in vitro* transcription-translation, or first transcribed *in vitro* to mRNA, purified, and subsequently used for *in vitro* translation (discussed in Section III, B, 3). During *in vitro* translation ribosomal complexes (mRNA-ribosome-polypeptide) form that contain a functionally folded protein emerging from the ribosomal tunnel and most probably still connected to the tRNA at its C-terminal end. For the protein to fold, it must have an unstructured region occupying the ribosomal tunnel, which must be encoded downstream of the gene region that encodes the folded protein (see Section III, B, 2). The composition and reaction conditions of the

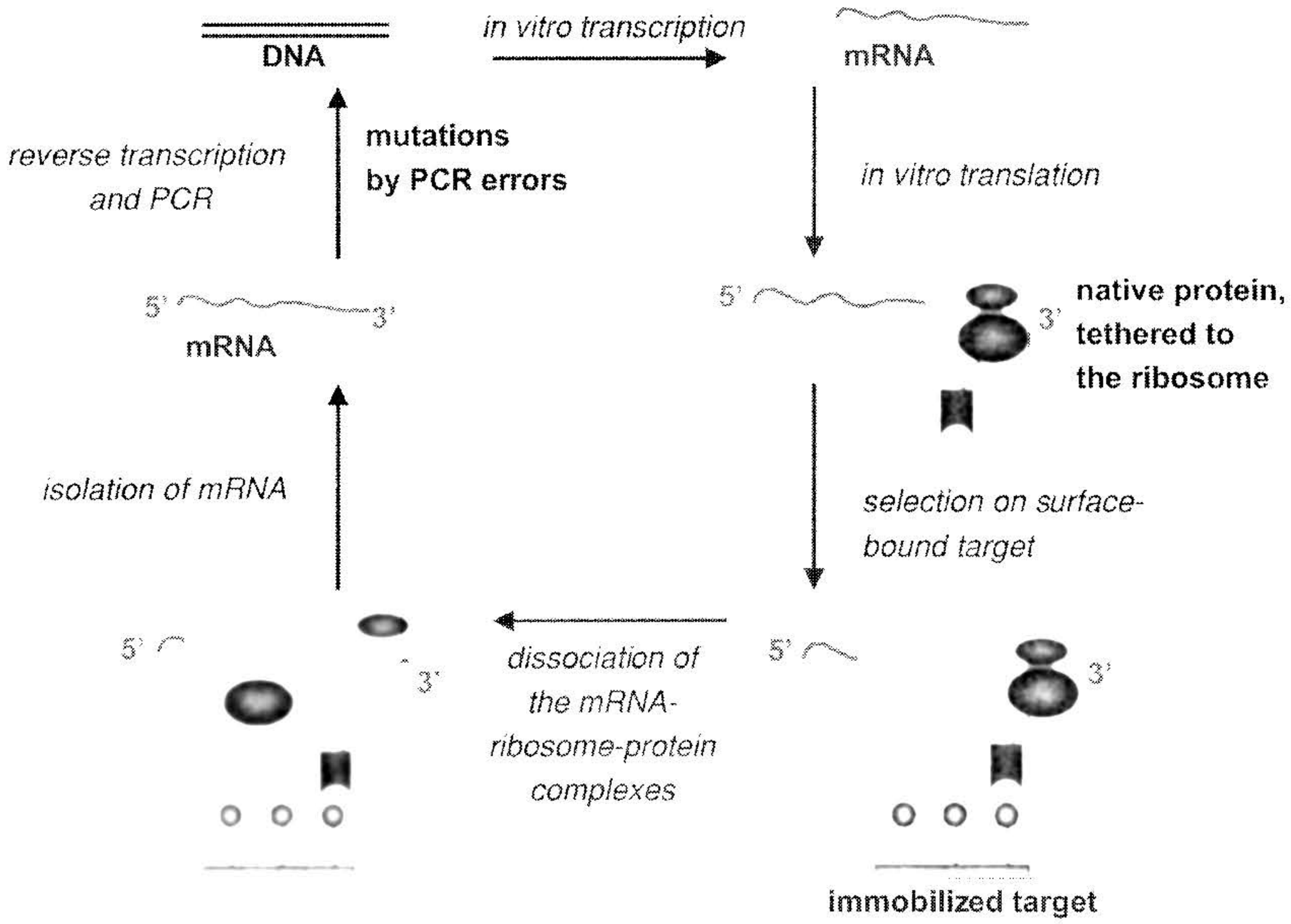

FIG. 2. Ribosome display. A library of proteins (e.g., scFv fragments of antibodies) is transcribed and translated *in vitro*. The resulting mRNA lacks a stop codon, giving rise to linked mRNA-ribosome-protein complexes. These are directly used for selection on the immobilized target. The mRNA incorporated in 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.

in vitro translation must be commensurate with folding of the particular protein in question (see Section III, B, 3).

In the *E. coli* system, it is important to stop the *in vitro* translation reaction by rapid cooling on ice. The reaction is usually diluted several-fold in prechilled buffer containing the components for stabilization of the ribosomal complexes. In the *E. coli* system, the ribosomal complexes can be very efficiently stabilized by low temperature and by high Mg^{2+} concentrations (50 mM), and then used for affinity selection. It is believed that high Mg^{2+} "condenses" the ribosome by binding to the rRNA, making it difficult for the peptidyl-tRNA to dissociate or be hydrolyzed. The low temperature probably slows down the hydrolysis of the peptidyl-tRNA ester bond, and perhaps also the thermal motions, which would facilitate dissociation of the peptidyl-tRNA. Such complexes are stable for up to several days.

a. Binding Selection. In principle, the affinity selection of the ribosomal complexes can be performed in two different ways. Either ligands

can be immobilized on a surface (panning tubes or microtiter wells, for instance), or biotinylated ligands can be used, which bind to the proteins displayed on the ribosomal complexes and are subsequently captured by streptavidin-coated magnetic beads. Usually, the panning is performed for one hour at temperatures equal to or below 4°C for routine enrichments. However, shorter as well as longer incubation times (as long as twenty days) are possible and can be advantageous. The former can be useful if all binders are to be captured, regardless of their affinity, while the latter is appropriate if very high affinity binders are to be evolved (see Section IV, B, 2). The advantage of performing the panning in solution is that the ligand concentration is well defined and the ligand is mostly in its native conformation. Therefore, the number of unspecifically bound ribosomal complexes is usually lower. Furthermore, nonspecific binding during affinity selection of ribosomal complexes can be decreased using diluted, autoclaved milk and heparin (Hanes *et al.*, 1998). Heparin may additionally act also as an RNase inhibitor.

b. Elution. After affinity selection, nonspecifically bound ribosomal complexes are removed by intensive washing with magnesium-containing buffer. Removal of the stabilizing Mg^{2+} ions with an excess of EDTA causes dissociation of all bound complexes, allowing the mRNA of bound ribosomal complexes to be directly isolated and obviating the need to elute the binder from the target. Alternatively, by competitive elution of bound ribosomal complexes with free ligand, followed by mRNA isolation of eluted complexes, only the mRNA present in complexes containing a functional and specific binding protein is isolated, possibly leading to higher enrichment factors. However, this approach might be difficult to apply for binders with a very high affinity for antigen.

c. Amplification. The isolated mRNA is then used for RT-PCR, and the amplified DNA can be used for the next cycle of ribosome display. When magnetic beads are used for selection, RT-PCR may also be performed directly with the washed beads (He and Taussig, 1997). After each round of ribosome display, a portion of the DNA can be analyzed by cloning and sequencing and by ELISA or RIA.

2. *The Ribosome Display Construct*

The features of the ribosome display construct are summarized in Figure 3. On the DNA level, the construct requires a T7 promoter for efficient *in vitro* transcription to mRNA. On the mRNA level, the construct contains, as a regulatory sequence for translation, either a prokaryotic ribosome binding site (Shine and Dalgarno, 1975) if the *E. coli*

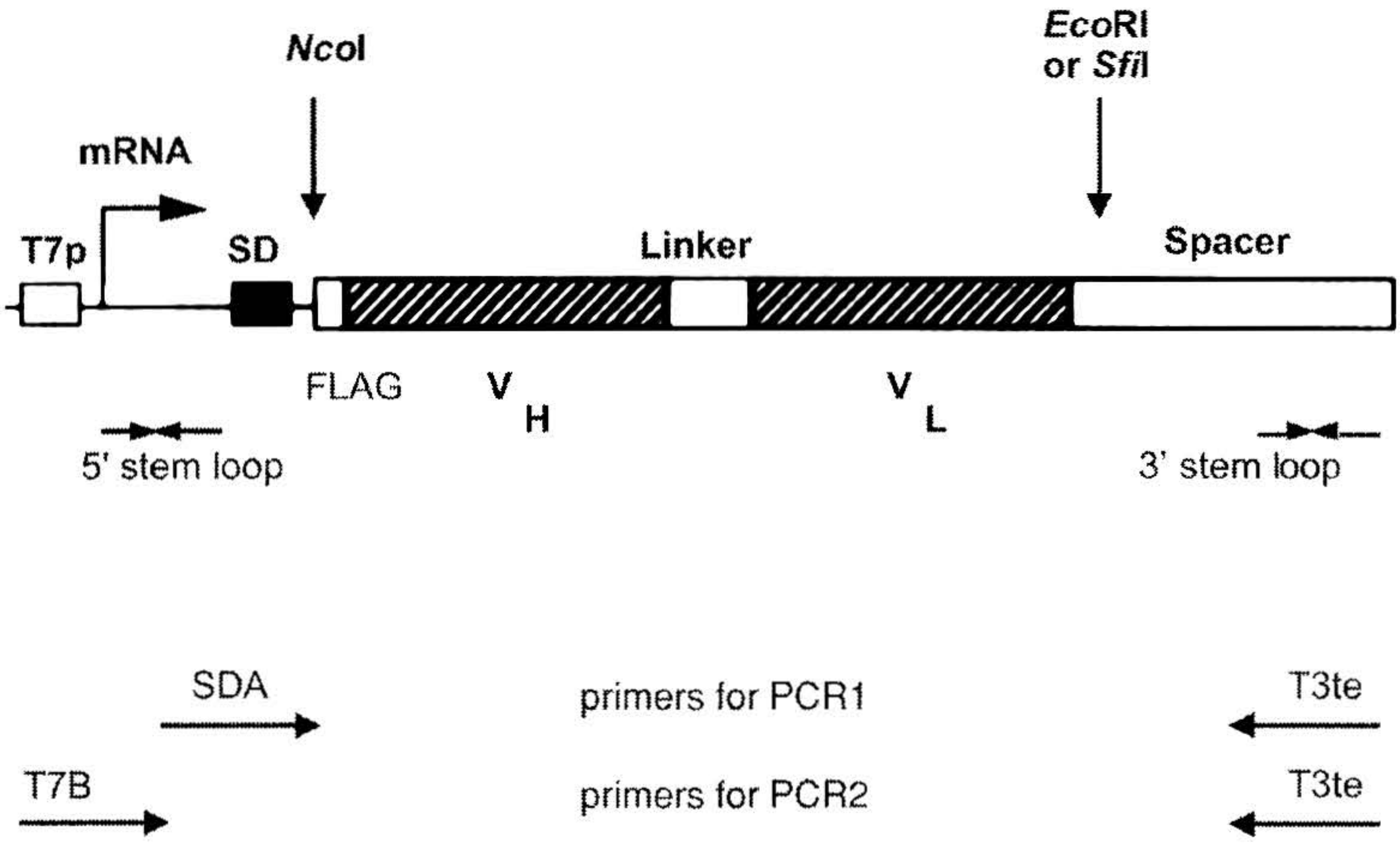

FIG. 3. DNA construct used for *E. coli* ribosome display, as illustrated for a single-chain antibody. The promoter (T7) is followed by a Shine-Dalgarno sequence (SD) and the protein of interest (here a scFv construct) containing an N-terminal FLAG tag for detection. The variable domains V_H and V_L are joined by a glycine/serine-rich linker. A spacer (tether) is cloned in frame behind the sequence of the antibody scFv fragment without a stop codon. Sequences encoding RNA stem-loop structures are present both at the 5′ and 3′ ends. In ribosome display, after the reverse transcription step (with primer T3te), two subsequent PCR steps are used to reintroduce the Shine-Dalgarno sequence (PCR1; primers SDA, T3te) and the T7 promoter (PCR2; primers T7B, T3te) to regenerate the complete scFv construct.

system is used, or a Kozak consensus and enhancer sequence (Kozak, 1984) if the eukaryotic ribosome display system is used. This sequence is followed by the open reading frame encoding the protein to be displayed, followed by a spacer sequence fused in frame to the protein.

The coding region ends with the protein sequence—that is, there is no stop codon present. In the prokaryotic system the presence of a stop codon would result in the binding of the release factors (Grentzmann *et al.*, 1995; Tuite and Stansfield, 1994) and the ribosome recycling factor (Janosi *et al.*, 1994) to the mRNA-ribosome-protein complexes. This would then lead to the release of the protein by hydrolysis of the peptidyl-tRNA (Tate and Brown, 1992), thereby dissociating the ribosomal complexes (Fig. 4A). A similar mechanism exists in eukaryotic systems (Frolova *et al.*, 1994; Zhouravleva *et al.*, 1995).

Another important prerequisite for efficient ribosome display in the *E. coli* system is the elimination of the 10Sa-RNA (Ray and Apirion, 1979). 10Sa-RNA is a stable bacterial RNA with a tRNA-like structure, but having an extended Ω-loop (Komine *et al.*, 1994). If a truncated

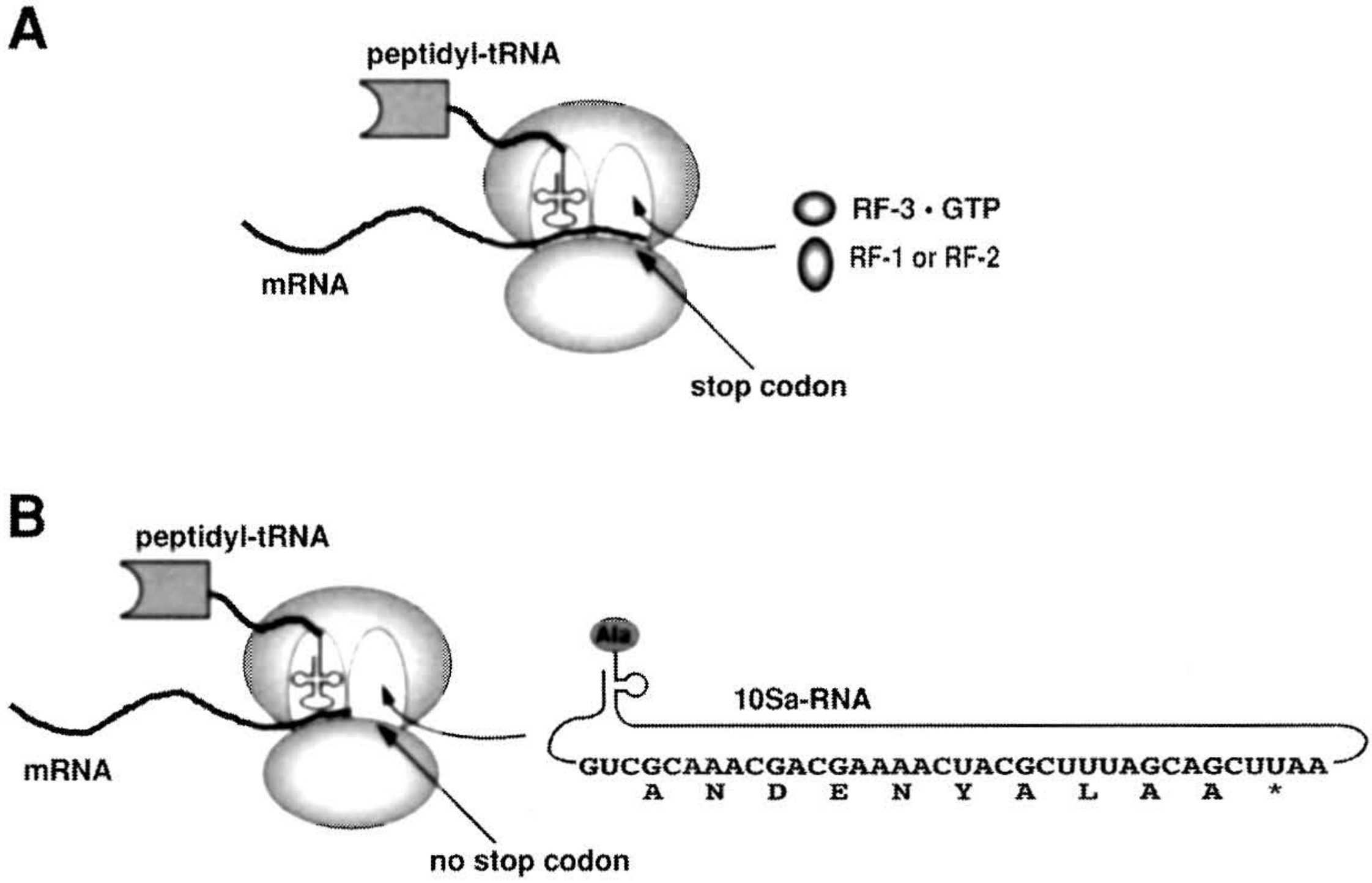

FIG. 4. Role of the stop codon and 10Sa-RNA in *E. coli* translation. (A) When a stop codon is encountered, a complex of two release factors, RF-1 and RF-3 or RF-2 and RF-3, binds instead of the tRNA. The release factor RF-1 recognizes the stop codons UAA and UAG, while RF-2 recognizes UAA and UGA. The binding of the release factor complex results in hydrolysis of the peptidyl-tRNA and release of the peptide. (B) The role of 10Sa-RNA. If truncated mRNA without a stop codon is translated in *E. coli*, the ribosome stops at the end of the mRNA. 10Sa-RNA can then bind to the ribosomal A site and 10Sa-RNA can act as tRNA by transferring an alanine to the truncated protein. Subsequently, 10Sa-RNA acts as mRNA and a peptide tag with the indicated sequence is added to the truncated protein. 10Sa-RNA encodes a stop codon and therefore the protein is released and then degraded by proteases specifically recognizing this C-terminal tag.

mRNA, lacking a stop codon, is translated *in vivo* in *E. coli*, the 10Sa-RNA binds to the ribosomal tRNA acceptor site. This results in a carboxy-terminal modification of the truncated polypeptide by addition of a peptide tag encoded by the 10Sa-RNA and subsequent release from the ribosome (Fig. 4B). The released protein, tagged with this sequence, is finally degraded by a tail-specific protease (Keiler *et al.*, 1996; Roche and Sauer, 1999).

At both ends of the mRNA, the ribosome display construct should include stemloops, 5′- and 3′-stemloops are known to stabilize mRNA against RNases *in vivo* as well as *in vitro*. The presence of stemloops is important, especially in the *E. coli* ribosome display system, because the extract used for *in vitro* translation contains high RNase activities. To date, five of twenty *E. coli* RNases have been shown to contribute to mRNA degradation (Hajnsdorf *et al.*, 1996), and they are probably all present in the S30 extract. The efficiency of ribosome display was in-

creased approximately 15-fold (Hanes and Plückthun, 1997), when a 5′-stemloop derived from the T7 gene 10 upstream region and a 3′-stemloop derived from the terminator of the *E. coli* lipoprotein were introduced into the ribosome display construct. A similar improvement in efficiency was observed when using the same 5′-stemloop and the 3′-stemloop derived from the early terminator of phage T3 (Hanes and Plückthun, 1997). The stemloop structures may protect the mRNA particularly from degradation by the exonucleases PNPase and RNaseII, which act from the 3′-end of the mRNA (Hajnsdorf *et al.*, 1996), and against RNaseE, which recognizes the 5′-end (Bouvet and Belasco, 1992).

A protein tail, which is the same in all library members, is fused to the C-terminus of the ribosome display construct and serves as a spacer. This spacer has two main functions. First, it tethers the synthesized protein to the ribosome. Second, it keeps the structured part of the protein outside the ribosome and allows its folding and interaction with ligands, without clashing with the ribosomal tunnel. The ribosomal tunnel covers between 20 and 30 C-terminal amino acids of the nascent polypeptide chain during protein synthesis and can therefore prevent the folding of the protein (Malkin and Rich, 1967; Smith *et al.*, 1978).

A ribosome display construct (the library in the ribosome display format) can be prepared completely *in vitro* either by ligation of the DNA library to the spacer region or by assembly PCR of the DNA library and the spacer. All the above-mentioned features, which are important for ribosome display (T7 promoter, ribosome binding site, and stem-loop structure), are then introduced by PCR (Fig. 3).

3. *The* In Vitro *Translation Step of Ribosome Display*

a. Coupled versus Uncoupled System. The ribosomal complexes can be generated either by a coupled *in vitro* transcription-translation system (Mattheakis *et al.*, 1994; He and Taussig, 1997) using a DNA library as a template, or the mRNA can be first prepared by *in vitro* transcription, purified and subsequently used for *in vitro* translation (Hanes and Plückthun, 1997; Gersuk *et al.*, 1997). The coupled system is much simpler than the uncoupled one, but especially in the case of the *E. coli* S30 system, it yields fewer functional ribosomal complexes (Hanes *et al.*, unpublished experiments). Furthermore, T7 RNA polymerase, which is used for *in vitro* transcription, requires reducing agents such as 2-mercaptoethanol for its stability. If disulfide-containing proteins are displayed in the coupled system, the presence of reducing agents during *in vitro* translation markedly decreases or may even abolish their folding efficiency and thereby the activity of the displayed protein. A separate transcription step under reducing conditions followed by the translation

step under oxidizing conditions can be used to solve this dilemma. This problem may in principle also be overcome by preparing T7 RNA polymerase without reducing agent, but enzyme activity has then to be carefully monitored.

b. Time and Temperature of In Vitro *Translation.* Every *in vitro* translation system differs in optimal translation time and temperature, both of which can influence the yield. In the *E. coli in vitro* translation system the translation reaction is usually performed at 37°C (Mattheakis *et al.*, 1994; Hanes and Plückthun, 1997), where the folding efficiency was found to be favorable. Although *in vitro* folding usually gives higher yields at low temperature, the combined temperature effects of the translation reaction, chaperone-assisted folding, escape from nucleases and proteases, and other unknown factors seem to be most favorable at 37°C for the yield of functional proteins.

The time of translation is also very important, especially for uncoupled systems. During *in vitro* translation, protein synthesis follows a saturation curve reaching a plateau after 30 minutes (Ryabova *et al.*, 1997). At the same time the mRNA is continuously degraded with a half-life of approximately 5 to 10 minutes. Thus, an optimal time exists at which the concentration of intact mRNA-ribosome-protein complexes that can be used for selection is maximal. The optimal time for the *E. coli* system is around 7 minutes.

In eukaryotic systems the optimal translation times are usually longer. For instance, an *in vitro* translated truncated lysozyme in a wheat germ system was still present in the ribosomal fraction after 30 minutes of translation (Haeuptle *et al.*, 1986). The translation time in the coupled system is not such a critical parameter, compared to the uncoupled system, since the mRNA is continuously being produced. The reaction time can therefore be extended to 30–60 minutes when using the *E. coli* (Mattheakis *et al.*, 1996) or the rabbit reticulocyte systems (He and Taussig, 1997). A longer translation time is not recommended because some crucial components necessary for translation or transcription become limiting and low molecular weight compounds generated during the translation accumulate and eventually inhibit the *in vitro* translation (Jermutus *et al.*, 1998).

c. Additives to the In Vitro *Translation.* The addition of several components, stabilizing either the mRNA or the ribosomal complexes or improving the protein folding during the *in vitro* translation reaction, can increase the overall efficiency of ribosome display. RNasin, an inhibitor of certain mammalian RNases, was first used in the wheat germ system

(Gersuk *et al.*, 1997), but it was not reported whether it had any effect, and it has no effect in the *E. coli* system.

To stop the translation reaction and further stabilize the ribosomal complexes, cycloheximide can be added in the eukaryotic system (Gersuk *et al.*, 1997). For the same purpose chloramphenicol, an antibiotic that inhibits bacterial protein synthesis by binding to the 23S ribosomal RNA in the peptidyl transferase center, can be used in the *E. coli* system (Mattheakis *et al.*, 1994). However, chloramphenicol was found to have no influence on the efficiency of *E. coli* ribosome display (Hanes and Plückthun, 1997).

Protein disulfide isomerase (PDI) was found to be important in catalyzing disulfide bond formation of antibody fragments, and it improved the efficiency of *E. coli* ribosome display of antibodies threefold when used during the *in vitro* translation reaction (Hanes and Plückthun, 1997) (see also Section II, B). A fourfold improvement of ribosome display was observed when 10Sa-RNA, which is involved in degradation of truncated proteins (see Section III, B, 2), was inhibited by using an antisense DNA oligonucleotide directed against the 10Sa-RNA (Hanes and Plückthun, 1997).

4. Applications of Ribosome Display

In contrast to the other *in vitro* selection technologies discussed below, where to date mostly model enrichments have been reported, more examples on directed evolution are available for ribosome display, and these are reported in a separate section (see Section IV). Briefly, experiments on peptide and protein libraries are summarized, and the directed evolution of distinct biophysical properties is discussed.

C. RNA-Peptide Fusion and In Vitro *Virus*

1. Principle of RNA-Peptide Fusion

A somewhat different approach to couple phenotype and genotype was designed by Roberts and Szostak (1997) and independently by Nemoto *et al.* (1997), who linked a peptide covalently to its encoding mRNA. In this technology, called "RNA-peptide fusion" (Roberts and Szostak, 1997; Roberts, 1999; Roberts and Ja, 1999) or "*in vitro* virus" (Nemoto *et al.*, 1997), mRNA is transcribed *in vitro*, purified, and subsequently ligated at its 3′-terminus to a puromycin-tagged DNA-linker (Fig. 5). This RNA-DNA construct with a puromycin at its 3′-end is again purified and then translated *in vitro*. The ribosome stalls upon reaching the RNA-DNA junction, allowing the puromycin to enter the peptidyltransferase

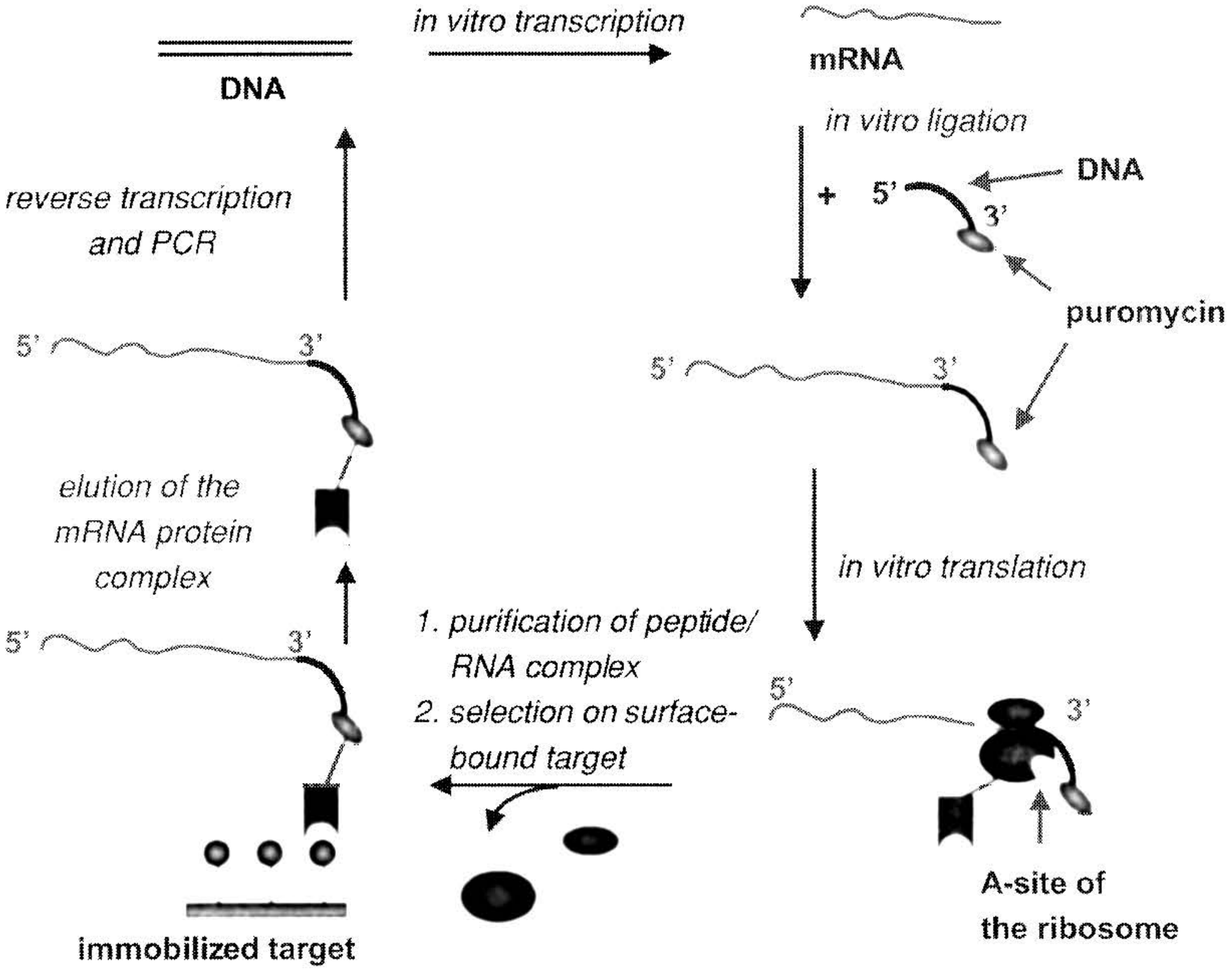

FIG. 5. Protein-RNA fusion. Covalent RNA-protein complexes can be generated by ligation of a DNA-puromycin linker to the *in vitro* transcribed mRNA. During *in vitro* translation, the ribosome stalls at the RNA-DNA junction. Puromycin can then bind to the ribosomal A-site. The nascent polypeptide is thereby transferred to puromycin. The resulting covalently linked complex of mRNA, puromycin, and peptide can be used for selection experiments. After affinity selection, the bound complexes are eluted and subsequently the mRNA is amplified by RT-PCR.

site and covalently couple to the nascent peptide. Optimization of the DNA-linker length minimized cross-contamination, which would result in polypeptides fused to a nonrelated mRNA (Roberts and Szostak, 1997). The resulting covalently linked complex of mRNA-DNA hybrid, puromycin, and the encoded polypeptide can then be dissociated from the ribosome and used for affinity selection. After selection, the bound RNA-peptide fusions can be eluted and amplified by reverse transcription and PCR. One critical and time-consuming step during each cycle is the ligation of the DNA-puromycin linker to the mRNA, which requires careful handling of the mRNA. The advantage of RNA-peptide fusion is that the covalently coupled complexes of mRNA and peptide are robust and therefore should allow more stringent selection conditions than in noncovalently coupled systems, but it is unclear how much the functional library size is decreased by the additional manipulations necessary in this procedure.

2. *Applications of mRNA-peptide Fusion*

In a model enrichment, the DNA encoding a peptide with the myc-epitope was diluted in different ratios in a DNA pool encoding random peptide sequences. The myc-epitope peptide was enriched by immunoprecipitation with an anti-c-myc antibody by a factor of twenty to forty per cycle of mRNA-peptide fusion (Roberts and Szostak, 1997). More recently, the mRNA-peptide fusion approach has been used to select peptides for binding to this anti-c-myc antibody from a library of 2×10^{13} molecules (Roberts and Ja, 1999). In these experiments, the enrichment factor was reported to be 200 per round and is therefore similar to the enrichment factors observed with ribosome display.

Protein-ligand interactions can not only be secreened or selected *in vitro,* but also can be directly characterized for particular interaction partners. Nemoto *et al.* (1999) applied the mRNA-peptide fusion technology to fluorescently label the displayed proteins in order to study protein-protein interactions by fluorescence polarization measurements.

D. Selection for Enzymatic Activity In Vitro

A long-standing desire of biochemists has been to generate catalysts, either by design or by selection (Arnold, 1998; Forrer *et al.,* 1999). One strategy, used with some success in the identification of catalytic antibodies, has been to select by phage display molecules that bind to a transition-state analog (Schultz and Lerner, 1995; Arkin and Wells, 1998). Another possibility is to covalently trap phages expressing active catalyst with suicide inhibitors (Baca *et al.,* 1997; Janda *et al.,* 1997).

However, a major limitation of this approach is that appropriate suicide inhibitors or transition state analogs are not available for most enzymatic reactions. Furthermore, binding to the transition-state analog does not necessarily correlate well with catalysis, nor does improved binding automatically generate an improvement in catalysis (Baca *et al.,* 1997).

The fundamental problem with selection for catalysis is that the product of an enzyme leaves the catalytic protein. Thus, even when genetic information of the catalyst is physically connected to the catalytic protein, the phenotype (i.e., the efficiency of the reaction) does not remain connected. In other words, in a mixture of catalysts of different efficiency, the information—which of the molecules has actually produced most of the product—becomes lost. As a consequence, product and enzyme have to remain physically connected for an efficient evolutionary process.

There are two principal ways of achieving this. Either a direct physical link between catalyst and substrate (and thus product) must be created or, alternatively, a compartmentalization of catalyst and replication ma-

chinery must be designed. The problem was solved by nature with the emergence of cells as compartments.

In practice, however, it turns out to be difficult to focus selection pressure on just one enzyme in a cellular environment. The challenge of these approaches lies in correctly directing selection pressure (Zhao and Arnold, 1997). Nature is surprisingly inventive in finding other solutions to escape the selection pressure than through the designed library. Many of the cautionary notes about microorganisms evading the selection pressure particularly apply to catalysis.

The selection for catalytic turnover is therefore more complicated than for binding to a transition state analog or a suicide inhibitor. It requires the possibility to physically separate positive library members from the rest (Barbas *et al.*, 1997; Janda *et al.*, 1997; Pedersen *et al.*, 1998). For instance, antibodies against the product can be used to capture product-containing phages (Tawfik *et al.*, 1993). This requires that the substrate and the enzyme are both displayed on the same phage (Pedersen *et al.*, 1998). The substrate can also be linked noncovalently to the phages displaying the protein (Demartis *et al.*, 1999). In principle, the described strategies for selection of catalysts can also be used with *in vitro* techniques such as ribosome display, but this has not yet been reported.

1. Principle of "Cell-like Compartments"

Tawfik and Griffith (1998) reported an *in vitro* selection strategy for catalytic activity using compartmentalization. Here, each member of the DNA library is encapsulated in an aqueous compartment in a water in oil emulsion. The compartments are generated from an *in vitro* transcription-translation system, and contain the components for protein synthesis. The dilution is chosen such that, on average, the water droplets contain less than one DNA molecule. The DNA is transcribed and translated *in vitro* in the presence of substrate, which is covalently attached to the DNA. Only translated proteins with catalytic activity convert the substrate to the product. Subsequently, all DNA molecules are recovered from the water droplets and the DNA linked to the product is separated from the unmodified DNA linked to the educt, which requires a method to discriminate between both. The modified DNA can then be amplified by PCR and used for a second selection cycle. The principle of this approach is depicted in Figure 6.

2. Applications of "Cell-like Compartments"

As a model system, a DNA methylase was chosen (Tawfik and Griffiths, 1998). DNA encoding the methyltransferase activity was methylated by

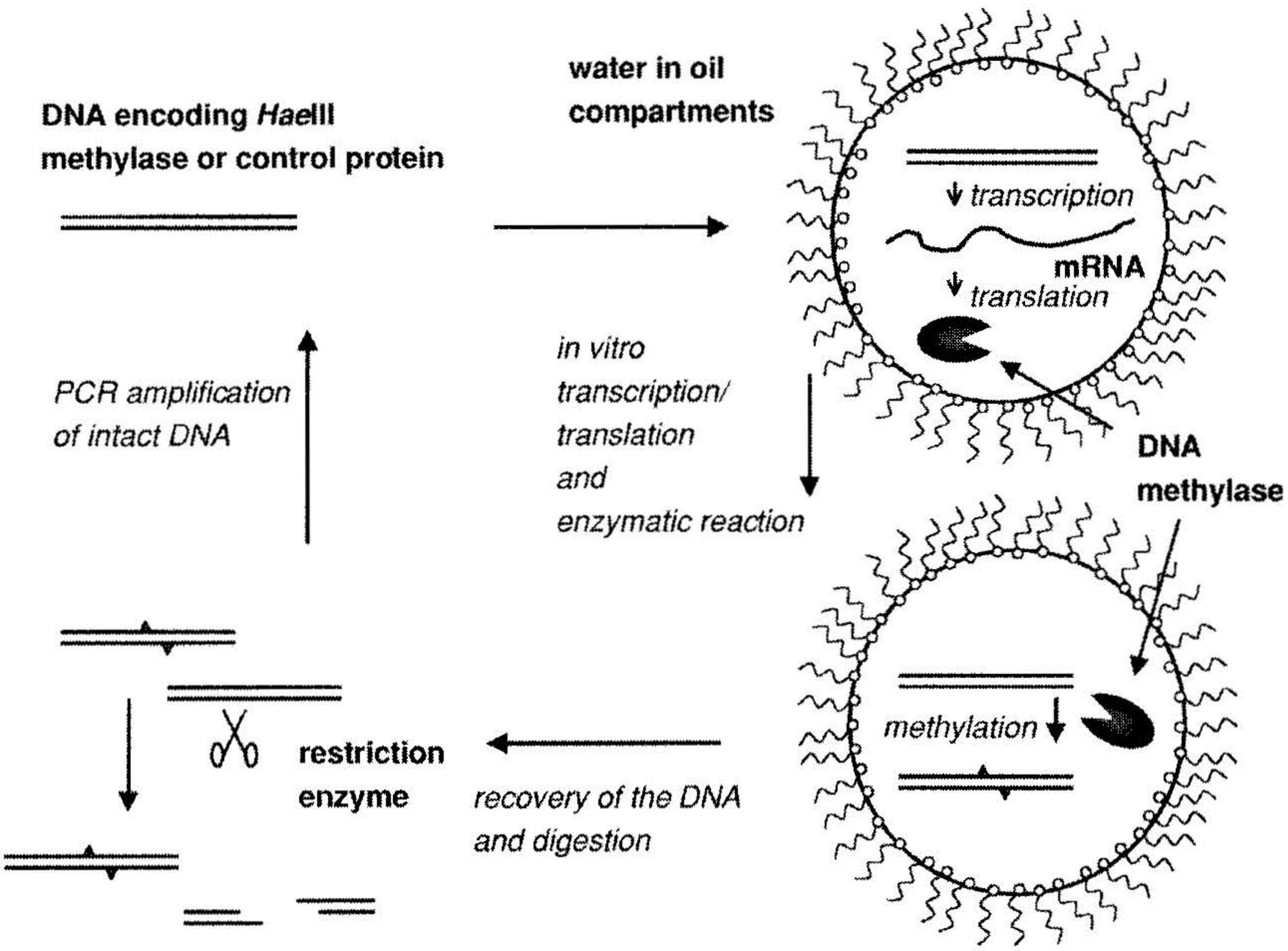

FIG. 6. Selection for enzymatic activity (DNA methylation) by compartmentalization. DNA encoding *Hae*III methylase is diluted with unrelated DNA (encoding dihydrofolate-reductase). This mixture is dispersed together with a reaction mixture for *in vitro* transcription and translation to form water in oil compartments. The dilution is chosen such that each compartment contains one DNA molecule on average. In the aqueous compartments, the genes are transcribed and translated. In compartments in which an active methylase is translated, the DNA can be methylated and is subsequently recovered from the emulsion and digested by a restriction enzyme. The methylated DNA (encoding *Hae*III methylase) is protected against the digestion, remains intact and is subsequently amplified by PCR.

the enzyme and thereby protected against digestion by a restriction endonuclease. The unmodified DNA was degraded and the intact methylated DNA was amplified by PCR. In a model enrichment, DNA encoding *Hae*III methylase was diluted in different ratios with DNA encoding an irrelevant protein and could be enriched in a single round by a factor of 5000.

The challenge in this technology is to generalize this compartment approach to reactions other than DNA modification. The discrimination of DNA-bound product and DNA linked to educt will be an important step in this endeavor.

E. Water in Oil Emulsions for Binding Selections (STABLE)

1. Principle of Water in Oil Emulsions for Binding Selections (STABLE)

In an approach similar to the "cell-like compartments," Doi and Yanagawa (1999) used biotinylated DNA to display peptides fused to streptavidin in compartments of water in oil emulsions. The method was named streptavidin-biotin linkage in emulsions, STABLE (Doi and Yanagawa, 1999). Upon *in vitro* translation each translated peptide is displayed as a fusion to streptavidin that binds to its encoding biotinylated DNA in its compartment. The resulting protein-DNA fusions can then be recovered and used for affinity selection. To avoid cross-contamination, biotin has to be added before recovery because much more streptavidin will be produced in each compartment than biotinylated DNA is present. The selected DNA-protein complexes can then be amplified by PCR. The principle of this selection strategy is shown in Figure 7.

One advantage of protein-DNA complexes is that DNA is much more stable as genotype than mRNA. Also, DNA encoding a stop codon can be displayed, facilitating the design of libraries from natural sources (cDNA libraries). However, this advantage may be offset by additional handling steps in this procedure.

It should be noted, however, that streptavidin forms tetramers; therefore four copies of the peptide will be displayed per DNA and a multivalency effect during the selection step is hard to avoid, which may make selection for high affinity more difficult.

2. Applications of STABLE

As a model enrichment, Doi and Yanagawa (1999) selected for Ni-NTA binders from a library of decamer peptides. The reported enrichment factor for the his-tag fused to streptavidin was, however, only ten. This is probably due to the low efficiency of the protein DNA fusion formation, which was estimated to be only 1%.

IV. Applications of Ribosome Display

In vitro selection technology can be used in principle for three tasks. We will discuss these in the following sections for ribosome display, as currently the examples are mostly available from this approach. The first task is to identify a protein or a peptide from a pool of variants. Here, ribosome display is used exclusively as a method for *selection,* and changes are not actively introduced in the original pool. In this case, proofreading polymerases are used during the PCR amplification steps (see below). A second application is to *evolve* a given protein or peptide

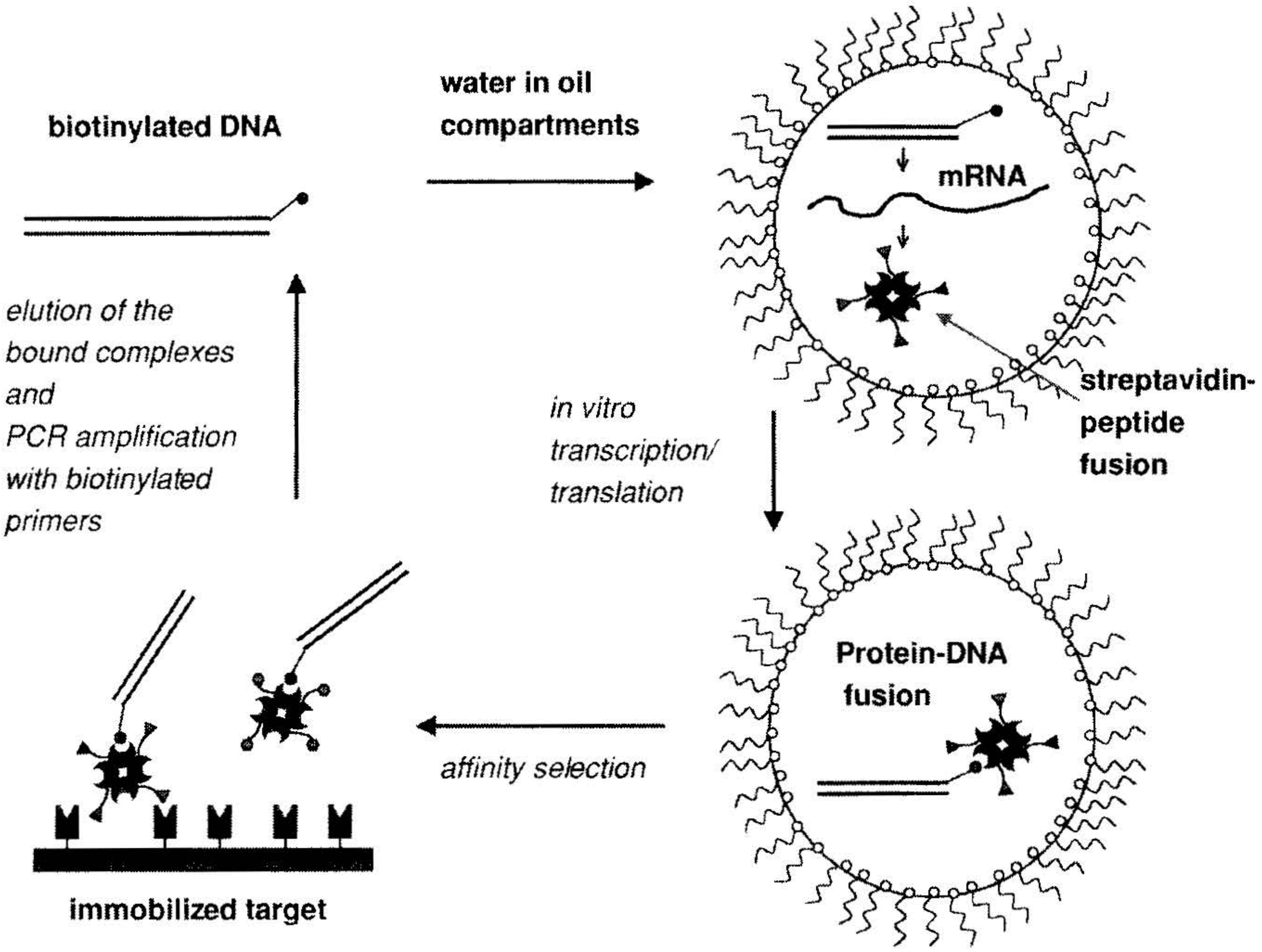

FIG. 7. STABLE. A biotinylated DNA library encoding streptavidin-random peptide fusions is dispersed together with the reaction mixture for *in vitro* transcription and translation in water in oil compartments. The dilution is chosen such that each compartment contains a single DNA molecule on average, which is transcribed and translated. The encoded streptavidin-peptide fusion is synthesized in the compartment and can bind to its encoding biotinylated DNA. The DNA-protein complexes are subsequently recovered from the emulsion and used for affinity selection. The DNA of the bound complexes is then eluted and amplified by PCR.

to obtain new or improved functions or properties. Third, it is possible to combine *both* tasks: select from a pool of variants, when an evolution of the selected clones occurs simultaneously. This requires a diversification of the pool after each selection cycle. In the simplest case, this is accomplished by the inherent error rate of a low-fidelity DNA polymerase used for the PCR amplification steps of ribosome display.

A. Selection from Libraries with or without Concurrent Sequence Evolution

1. Peptide Display

The first successful application of ribosome display system was demonstrated for the display of a library of random peptide decamers

(Mattheakis *et al.*, 1994). A library of 10^{12} DNA molecules was used with *E. coli* ribosome display utilizing a coupled *in vitro* transcription-translation system. This library was selected for binding to the monoclonal antibody D32.39, which originally bound dynorphin B, a 13-residue opioid peptide, with 0.29 nM affinity. Five cycles of ribosome display resulted in several different peptides with affinities to the antibody ranging from 7.2 to 140 nM affinity. Yet, a peptide with a sequence similar to dynorphin B was not isolated.

A similar approach was followed for displaying a random library of 20-mers using a wheat germ uncoupled transcription and translation system (Gersuk *et al.*, 1997), and several peptides were isolated that bound to prostate-specific antigen, but not to bovine serum albumin. No affinities were reported.

Because the peptide sequences are very short in such systems, few (if any) errors are introduced by the polymerase and even error-prone PCR techniques are not ideal for introducing sequence changes. Therefore, unless a cassette mutagenesis strategy is followed, the benefit of ribosome display for an evolutionary approach is not exploited with displayed peptides. What is exploited, however, is the possibility to use larger libraries than those possible with *in vivo* selection systems.

2. *Protein Display*

The *E. coli* ribosome display system had to be considerably optimized for efficient display of folded proteins. These improvements (explained in more detail in Section III, B) involved the use of RNase inhibitors, the design of hairpins at either end of the RNA and a separate transcription and translation step allowing control over the individual redox requirements, and lead to higher yields, greater stability, and reduced nonspecific binding of the complexes (Hanes and Plückthun, 1997; Hanes *et al.*, 1999).

In a model system of two scFv fragments of an antibody, a 10^9-fold enrichment was achieved by five cycles of ribosome display with an average enrichment of about 100-fold per cycle (Hanes and Plückthun, 1997). All selected scFvs had mutated during five cycles of ribosome display, possessing between zero and four amino acid changes, compared to the original sequence.

Subsequently, it was demonstrated that it is possible to select and evolve scFv antibody fragments from immune libraries using the *E. coli* system. Only three rounds of ribosome display were necessary to isolate a family of scFv fragments binding to a peptide variant of the GCN4 leucine zipper, which exists as a random coil in solution (Hanes *et al.*, 1998). Most of the isolated scFvs had again acquired mutations, with

zero to five amino acid changes compared to their consensus sequence, the most likely progenitor scFv, which was present in the library before selection. The best scFv had a dissociation constant of $(4 \pm 1) \times 10^{-11}$ M, measured in solution. The likely common progenitor of the related scFvs bound the antigen with a 65-fold lower affinity than the best binder. This result demonstrated for the first time that from the PCR-based mutants, proteins can be evolved to higher affinity with ribosome display.

The display of proteins using the rabbit reticulocyte ribosome display system was also reported. Here a small library, derived from a scFv derivative of an antibody, V_H-linker-V_L-C_L, binding to progesterone, was used for selection (He and Taussig, 1997). The authors used a coupled *in vitro* transcription-translation system in the presence of 2 mM DTT. However, this concentration of reducing agents can in some cases prevent the folding of disulfide containing proteins such as antibodies.

The eukaryotic ribosome display was subsequently used for selection of human antibody scFv fragments binding progesterone from a library prepared from transgenic mice (He *et al.*, 1999). In this case, a proofreading polymerase was used for the PCR amplification steps included in the ribosome display protocol. It was thereby demonstrated that it is possible to use ribosome display as a method exclusively for selection by virtually maintaining the original library repertoire.

In a direct comparison of the rabbit reticulocyte system to the *E. coli* ribosome display system in a model study, the *E. coli* system turned out to be more efficient for the display of the model scFv constructs tested (Hanes *et al.*, 1999).

A more direct proof that protein variants of higher affinity are truly created during the *E. coli* ribosome display procedure was obtained by using HuCAL (Knappik *et al.*, 2000), a very large synthetic antibody library of 2×10^9 independent members (Hanes *et al.*, 2000). This naive library was applied for six rounds of ribosome display selection using insulin as the antigen. In three independent experiments different scFv families with different framework combinations were isolated. Since the library used was completely synthetic, consisting of forty-nine framework combinations, with the CDR3 regions of both variable domains randomized (Knappik *et al.*, 2000), the starting scFv sequences were known. Thus, any mutations could be directly identified as being generated during the ribosome display procedure. By sequence comparison to the original members of the library it was apparent that all the antibodies selected were not part of the initial library. By using nonproofreading polymerases, mutations were introduced into the enriched pool during the PCR amplification step that is part of each ribosome display cycle.

Thereby, new diversity was generated and each single member of the library began to diversify. This procedure closely mimics the process of somatic hypermutation of antibodies during secondary immunization. The final products of selection are different families of closely related sequences stemming from a common progenitor that started to evolve during ribosome display. A biophysical comparison of the isolated scFvs to their progenitors revealed that all selected scFvs had mutated and significantly improved their affinities to the antigen up to 40-fold by these mutations. The best scFvs had affinities in the low picomolar range (Hanes *et al.*, 2000).

B. Directed Evolution of Binding Affinity and Stability

Directed evolution consists of cycles of diversification and selection. Because ribosome display takes place entirely *in vitro*, it can ideally be combined with *in vitro* methods of generating sequence diversity. Since true evolution requires diversification in each cycle, this facile alternation between *in vitro* mutagenesis and *in vitro* selection is one of the attractions of the ribosome display method.

1. Introducing Diversity

Depending on the particular protein under consideration, either a focused library or a randomization of the whole gene encoding the protein may be more appropriate. Clearly, this depends on the prior knowledge of the system and its history. If binding affinity is the target function, and the binding site is known, it can be advantageous to first target the residues presumed to be in direct contact with the ligand and then target the whole gene to affect long-range interactions and second sphere residues, which may have subtle effects on the positioning of the direct contact residues. If the binding residues are not exactly known or if they may already be the outcome of an *in vivo* evolution (e.g., in the case of a natural antibody), or in a target function that invariably involves the whole protein—such as stability or folding efficiency—whole gene randomization is probably more effective.

Directed mutagenesis with oligonucleotides encoding mixtures of amino acids (Hermes *et al.*, 1990), error-prone PCR in the presence of nonphysiological metal ions such as Mn^{2+} (Lin-Goerke *et al.*, 1997) or dNTP analogues (Zaccolo *et al.*, 1996) can all be used to randomize the gene either at a particular sequence position or over the whole gene sequence. If a family of homologous proteins with similar function is available, recombination methods such as DNA shuffling (Stemmer, 1994), family shuffling (Crameri *et al.*, 1998), StEP (*St*aggered *e*xtension

*p*rocess, Zhao *et al.*, 1998) and RPR (*r*andom-*p*riming *in vitro r*ecombination, Shao *et al.*, 1998) can be used to generate a library. Furthermore, recombination is especially important in between the selection cycles to generate new diversity in the selected pool.

In many of the approaches, the number of mutations per gene can be adjusted, and it is likely that optimal mutation rates exist. On one hand, it is important to have a reasonably large collection of mutants in order to screen sufficient sequence space. On the other hand, it is absolutely necessary to preserve the function of the protein. If too many mutations are introduced, the harmful or destructive mutations can neutralize the beneficial effect of other mutations in one gene. It is generally assumed that *evolution* occurs by steps of increasing fitness and that the sequence of a functional protein must form a continuous path that can be traversed by single mutational steps without passing nonfunctional or less adapted forms (Spetner, 1970; Macken and Pereson, 1989; Smith, 1970; Gillespie, 1984). Favorable double mutations, with one mutation leading to an unfavorable intermediate, have been considered to be rather rare and therefore not important in evolution (Smith, 1970; Gillespie, 1984), even though this is difficult to exclude in general. Therefore, several investigators examined the question of an optimal mutation rate (Kepler and Perelson, 1995 Leigh, 1973; Sasaki and Isawa, 1989) in different contexts. However, as the tolerance of proteins to mutation and the frequency of beneficial mutations is poorly understood at best, further work will be required before this fundamental question can be answered.

2. *Evolving the Affinity of Ligand-binding Proteins*

There are, in principle, two ways to select for improved binding constants. The first strategy relies on equilibrium selection. Here, a displayed library of proteins is incubated with low amounts of ligand such that the concentration of the cognate partner is below the dissociation constant of the protein-ligand complex. On reaching equilibrium, interactions with higher affinity should be favored and, as a consequence, those binders should become enriched during selection. However, the equilibration time τ increases with decreasing ligand concentration:

$$\tau \sim \frac{1}{k_{on}[L] + k_{off}} \tag{1}$$

where k_{on} and k_{off} are the kinetic constants and $[L]$ denotes the ligand concentration. If k_{off} is assumed to be very small (e.g., 10^{-6} s^{-1}), which should be the case for tight binders, the first term with ligand concentra-

tions in the picomolar range becomes negligible at a typical k_{on} of 10^5 M^{-1} s^{-1} and the system needs more than a week to equilibrate! Furthermore, weak interactions are not actively excluded from the selection process. These theoretical considerations are in line with experimental observations: In equilibrium selections with the aim of obtaining tighter binders no evolution of the entire sequence pool was observed (Jermutus *et al.*, unpublished observations). Therefore, only very few improved binders could be detected.

Alternatively, the kinetic constants k_{on} and k_{off} can be targeted directly. While k_{on} is primarily controlled by translational and rotational diffusion as well as orientation factors and ranges usually from 10^5 to 10^7 M^{-1} s^{-1} (Northrup and Erickson, 1992), k_{off} of typical receptor-ligand interactions can vary over several order of magnitudes (10^{-1} to 10^{-6} s^{-1}). Off-rate selection has the potential to significantly improve binding affinity (Hawkins *et al.*, 1992; Yang *et al.*, 1995; Boder and Wittrup, 1997, Chen *et al.*, 1999). Furthermore, the selection time can be controlled easily, such that a selection for predefined kinetic constants is feasible. Provided that the ligand can be obtained in sufficient amounts and can be tagged, the protocol for off-rate evolution is straightforward: The displayed protein library is equilibrated with a low concentration of tagged antigen, usually in the range of the starting K_D. In the next step, a high molar excess of free, untagged antigen is added and the incubation is continued for increasing time periods in each evolution round. By adding the free antigen in excess, any dissociation of displayed protein with its tagged ligand becomes irreversible. In this strategy, weakly interacting molecules are titrated from the selection process. Mutants with faster off-rates are actively trapped by free antigen and washed away, even if they are present at very high concentrations. As a consequence, the background of low affinity binding proteins surviving the selection pressure is reduced.

In a directed *in vitro* evolution of antibody affinity (Jermutus *et al.*, 2000), ribosomal complexes coding for the protein library were first equilibrated with nanomolar concentrations of biotinylated antigen, and then an excess of free antigen was added. After time periods that increased in each round, complexes still binding the biotinylated antigen were rescued by the addition of streptavidin magnetic beads. The mRNA coding for these proteins was purified and served as the template for the next evolution round. An initial library of the fluorescein-binding antibody c12 with a starting K_D of 1.2 nM was created by error-prone PCR with the dNTP analogues dPTP {6-(2-deoxy-β-D-ribofuranosyl)-3,4-dihydro-8*H*-pyrimido[4,5-c][1,2]-oxazin-7-one-5′-triphosphate} and 8-oxo-dGTP (Zaccolo *et al.*, 1996). Despite the fact that RNA is generally regarded a labile molecule, the selection could be carried out for more

than 10 days, provided that a high magnesium concentration and low temperature were maintained. Three selected mutants were cloned into *E. coli*, expressed and purified.

Measurements of k_{on}, k_{off}, and K_D showed that k_{off} and as a consequence K_D had indeed been improved by more than one order of magnitude (Jermutus *et al.*, 2000). The evolved scFvs all contained multiple mutations, compared to the parent molecule. Between four and eleven amino acids (mean value of 7.2) were mutated per scFv. The majority of these mutations are located in permissive positions on the surface of the molecule, in areas unlikely to be directly involved in antigen recognition. Only two positions are mutated in the majority of all sequences, indicating a strong selection: The mutation of His L94 to tyrosine in CDR L3 affects a residue that points straight into the antigen binding site (Fig. 8A). The mutation of Asp H101 to glycine, alanine, or serine in CDR H3 affects a residue on the outer side of CDR H3, pointing away from the antigen binding pocket. This substitution will have the effect of breaking the salt bridge and increasing the flexibility of CDR H3 such that it can adopt a more "open" conformation (see legend to Fig. 8A for details). The multiple occurrence of Asp H101 replacements shows that this solution for improving the off-rate of the scFv fragment has been found several times independently during the *in vitro* evolution process.

3. *Increasing Protein Stability*

The energy difference between the native and the denatured state of a protein is very small and typically in the range of 5 to 15 kcal/mol. As a consequence, seemingly minor changes in the structure can lead to reduced stability and/or open up new pathways leading to misfolding. This observation provides the rationale for evolving proteins to higher stabilities *in vitro*: As it is difficult to introduce conditions unfavorable for folding without affecting other components of the *in vitro* system, the best strategy is to reduce the stability of the wild-type protein such that it no longer functions. The selection is then for additional mutations that lead to a regain of function to compensate for the loss incurred by the original destabilization. Mutants with increased folding rates or higher intrinsic stabilities should then be selected. The selection design must be carefully considered, as the level of destabilization will affect the selection background. This background is caused by proteins that survive the selection process without improved properties. The higher the stability of the initial pool, the more difficult it is to select for improved mutants. After a sufficient number of evolution cycles, single mutants that have adapted to the selection stress can be identified. By then removing the stress, such as by reversing a destabilizing mutation or

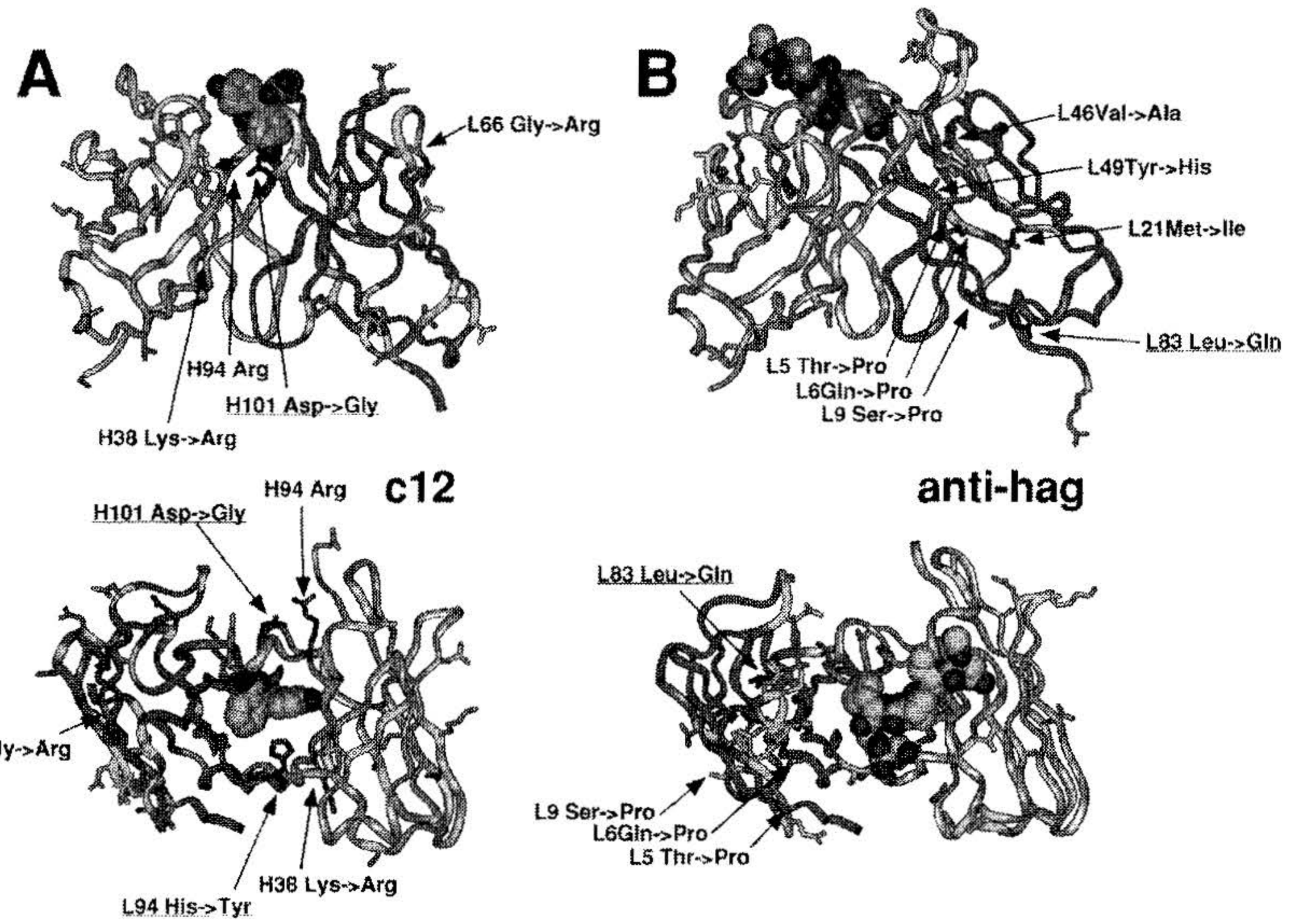

FIG. 8. Localization of the mutated residues after ribosome display. (A) Homology model of the antibody c12 with docked fluorescein. The strongly selected mutations L94 His to Tyr and H101 Asp to Gly, Ser or Ala are labeled and underlined. The mutation in position H101 destroys the salt bridge to Arg H94, which will most likely lead to a more open conformation of CDR3. Other mutations, which have been selected in several clones and are believed to indirectly contribute to binding, are labeled. The substitution of Lys H38 to Arg is very likely to have a stabilizing influence, as it participates in a highly conserved charge cluster with Glu H46 and Asp H86. The mutation of Gly L66 to Arg is expected to have a profound effect on the conformation of the outer loop and thus indirectly affect the geometry of the antigen binding site. (B) Experimental structure of the anti-influenza hemagglutinin antibody Fab 17/9 (PDB entry lifh). The strongly selected mutations L83 Leu to Gln is labeled and underlined. This residue is located at the bottom of the V_L domain and in intact antibodies contributes to the interface between variable and constant domain. However, it is exposed in scFv fragments, and thus a hydrophilic residue here may be beneficial. Mutations in buried and semi-buried positions are concentrated in the V_L domain (Met L21 Ile, Val L46 Ala, Tyr L49 His), while the V_H domain accumulated predominantly surface and interface mutations. In one clone, a stretch of consecutive proline residues (Thr L5 Pro, Gln L6 Pro, Ser L9 Pro, with L8 being a native Pro) was selected. Probably these residues lead to a stabilization by limiting the conformational degrees of freedom of the unfolded state even though they will not be able to form all of the main chain H-bonds of the original sequence and will therefore lead to some conformational adjustments. The positions of the remaining mutated residues are indicated in gray without labeling, illustrating how ribosome display leads to a targeting of the whole sequence.

removing denaturants, more stable variants of the wild-type proteins can be created. To illustrate this method a few examples are given below.

If certain amino acids (such as disulfide bonds or crucial amino acids in the hydrophobic core) are indispensable for protein stability, these positions can be changed by site-directed mutagenesis (Proba *et al.*, 1998). To avoid back-mutations during the evolution process or the selection of a residual wild-type contamination, the pool is amplified after each round of ribosome display with a primer that reintroduces the destabilizing mutation. If the mutation is not close to one of the termini, the coding sequence has to be amplified in two parts, which are then reassembled by PCR. Thus, to evolve improved stabilities this strategy first removes known crucial stabilizing factors to select for compensatory mutations at different positions.

Another approach may be used that focuses on the nascent protein chain. Since ribosome display depends on *in vitro* translation and folding to the functional state, the folding of the nascent chain can, in the case of disulfide-containing proteins, be inhibited by adding DTT (Jermutus *et al.*, 2000). Similar approaches should in principle be possible by adding proteases and suitable amounts of detergents and denaturants, even though this will have to be tested for each particular case. Likewise, important chaperones might be removed from the translation mix by immunoprecipitation. Another, more speculative strategy could make use of published *in vitro* translation systems from extremo- or thermophiles such as *Thermus thermophilus* (Watanabe *et al.*, 1980, Ueda *et al.*, 1991). Here folding, and eventually even selection, could occur at high temperatures.

Single-chain Fv antibody fragments contain two conserved disulfide bridges. These are important stability elements, and the removal of the disulfide bond usually results in a significant loss of activity. Based on this observation a strategy for evolving improved stability was defined (Jermutus *et al.*, 2000). An anti-hemagglutinin scFv with a stability of about 24 kJ/mol and midpoint of denaturation of 4.5 M urea was used as a test case. Because it was shown previously that oxidizing conditions during *in vitro* translation were necessary for maximal yields of functional protein, more stable mutants were selected by choosing a reducing redox potential during the translation step in ribosome display. Over five rounds the selection pressure was gradually increased by increasing the DTT concentration from 0.5 mM to 10 mM, corresponding to a final redox potential of at least −300 mV (assuming less than 0.1% oxidized species at equilibrium). Mutants could only survive the selection process if they folded to a stable structure in the presence of DTT and retained their antigen binding activity. For this purpose, ribosomal

complexes were incubated under reducing conditions on immobilized antigen (hag-peptide) and washed only briefly to avoid any selection for tighter binders. Since the selection is designed to enrich mutants that regain sufficient functionality, higher affinities are a possible selection shortcut (Jung *et al.*, 1999). After five selection rounds single mutants were cloned into *E. coli*, expressed, and purified. The most stable protein had increased its stability by about 30 kJ/mol, shifted its denaturation midpoint by 0.9 M urea, and displayed an m-value from the denaturation curve very close to the theoretical value for a two-state transition. From urea renaturation experiments under reducing conditions it was concluded that the evolved mutants could quantitatively refold in the presence of DTT. In contrast, the transition of the wild-type protein under reducing conditions indicated a population of nonnative species remaining after refolding and, thus, incomplete reversibility. Moreover, unlike the wild-type protein, all the mutants could be functionally expressed in the cytoplasm. Sequencing revealed that the mutants had acquired three to seven mutations in the coding sequence (mean value of 4.8). From modeling (Fig. 8B) and biophysical analysis of these mutants it could be concluded that they had all used different mutation strategies to adapt to the selection pressure. Both experiments, the maturation of off-rate and stability, indicated that the mutants had used different lineages during the evolution process, which is most probably due to the large library size in each ribosome display selection.

V. Perspectives of Directed *In Vitro* Evolution

The *in vitro* evolution of proteins is now a reality. To date, most of the evolutionary experiments have been carried out with ribosome display, but applications of the other described technologies will surely follow. The increased library size during selection and the experimental ease of including complex diversification techniques make *in vitro* selection techniques the methods of choice for the deliberate alteration of protein characteristics. However, *in vitro* selections will only be successful if protocols can be designed that direct the evolution process to the intended phenotype and minimize the risk of selection shortcuts. While selection for binding, improved affinity, and increased stability have now been described in the literature, more challenging goals such as enzymatic activity, expressability, or, in the case of scFvs, shifts in monomer-dimer equilibrium will need considerable effort for designing generally applicable selection strategies.

It is likely that future developments will address the automation of *in vitro* evolution technologies, as chemical processes are in general easier

to automate than biological ones. Cellular processes are variable and dependent on more parameters, some of which are difficult or impossible to influence.

In vitro protein evolution will complement but not replace hypothesis-driven or structure-aided engineering, as it would be very uneconomical to not make use of available knowledge regarding crucial residues, interactions, or known structural transitions. Directed evolution is suited to making the small adjustments that are beyond today's predictive methods. Further, instead of repeatedly discovering the same features in every directed evolution experiment, it can be advantageous to "dope" a library with possible mutations in the suspected positions. Thus, a combination of both methods—structure-based rough sketching and evolutionary fine-tuning—is likely to become a standard approach for solving practical problems in protein engineering and design. A very important corollary of this perspective is that there is great merit in detailed biophysical study of the effect of point mutations, as this knowledge will greatly facilitate the design of "smart libraries."

In conclusion, there are four key advantages in carrying out selections and evolutionary refinements *in vitro.* First, it is rapid, as no cellular cloning is involved. Second, the size of libraries is only limited by the amount of DNA (or RNA) that can be handled. Third, it is, in general, easier to focus the selection pressure on the quantity in question *in vitro* than in the highly variable context of a living cell. Fourth, the interfacing of selection from complex libraries and their simultaneous evolution is more convenient, as both can be carried out *in vitro. In vitro* protein evolution thus has a bright future.

ACKNOWLEDGMENTS

C. S. and L. J. are supported by a predoctoral Kékule-fellowship from the Fonds der Chemischen Industrie (Germany). The authors would like to thank Stephen Marino, Julia Henckel, and Imre Berger for critical review of the manuscript.

REFERENCES

Allen, D., Simon, T., Sablitzky, F., Rajewsky, K., and Cumano, A. (1988). *EMBO J.* **7,** 1995–2001.

Andersen, P. S., Stryhn, A., Hansen, B. E., Fugger, L., Engberg, J., and Buus, S. (1996). *Proc. Natl. Acad. Sci. USA* **93,** 1820–1824.

Arkin, M. R., and Wells, J. A. (1998). *J. Mol. Biol.* **284,** 1083–1094.

Arnold, F. H. (1998). *Nat. Biotechnol.* **16,** 617–618.

Baca, M., Scanlan, T. S., Stephenson, R. C., and Wells, J. A. (1997). *Proc. Natl. Acad. Sci. USA* **94,** 10063–10068.

Barbas, C. F., Heine, A., Zhong, G., Hoffmann, T., Gramatikova, S., Bjornestedt, R., List, B., Anderson, J., Stura, E. A., Wilson, I. A., and Lerner, R. A. (1997). *Science* **278,** 2085–2092.

Barbas, C. F., Bain, J. D., Hoekstra, D. M., and Lerner, R. A. (1992). *Proc. Natl. Acad. Sci. USA* **89,** 4457–4461.

Berek, C., and Milstein, C. (1987). *Immunol. Rev.* **96,** 23–41.

Berger, S. L., and Birkenmeier, C. S. (1979). *Biochemistry* **18,** 5143–5149.

Boder, E. T., and Wittrup, K. D. (1997). *Nat. Biotechnol.* **15,** 553–557.

Bouvet, P., and Belasco, J. G. (1992). *Nature* **360,** 488–491.

Cadwell, R. C., and Joyce, G. F. (1992). *PCR Methods Appl.* **2,** 28–33.

Chen, Y., Wiesmann, C., Fuh, G., Li, B., Christinger, H. W., McKay, P., de Vos, A. M., and Lowman, H. B. (1999). *J. Mol. Biol.* **293**, 865–881.

Chien, C. T., Bartel, P. L., Sternglanz, R., and Fields, S. (1991). *Proc. Natl. Acad. Sci. USA* **88,** 9578–9582.

Chothia, C., and Lesk, A. M. (1987). *J. Mol. Biol.* **196,** 901–917.

Crameri, A., Raillard, S. A., Bermudez, E., and Stemmer, W. P. (1998). *Nature* **391,** 288–291.

Daugherty, P. S., Olsen, M. J., Iverson, B. L., and Georgiou, G. (1999) *Protein Eng.* **12,** 613–621.

Demartis, S., Huber, A., Viti, F., Lozzi, L., Giovannoni, L., Neri, P., Winter, G., and Neri, D. (1999). *J. Mol. Biol.* **286,** 617–633.

De Pasquale, C., and Kanduc, D. (1998). *Biochem. Mol. Biol. Int.* **45,** 1005–1009.

Doi, N., and Yanagawa, H. (1999). *FEBS Lett.* **457,** 227–230.

Dougan, D. A., Malby, R. L., Gruen, L. C., Kortt, A. A., and Hudson, P. J. (1998). *Protein Eng.* **11,** 65–74.

Dower, W. J., and Cwirla, S. E. (1992). In: Chang, D. C., Chassy, B. M., Saunders, J. A., and Sowers, A. E. (eds.), *Guide to electroporation and electrofusion.* Academic Press, San Diego, 291.

Eaton, B. E., Gold, L., Hicke, B. J., Janjic, N., Jucker, F. M., Sebesta, D. P., Tarasow, T. M., Willis, M. C., and Zichi, D. A. (1997). *Bioorg. Med. Chem.* **5,** 1087–1096.

Eigen, M., and Rigler, R. (1994). *Proc. Natl. Acad. Sci. USA* **91,** 5740–5747.

Engel, A., Lyubchenko, Y., and Müller, D. (1999). *Trends Cell Biol.* **9,** 77–80.

Fields, S., and Song, O. (1989). *Nature* **340,** 245–246.

Forrer, P., Jung, S., and Plückthun, A. (1999) *Curr. Opin. Struct. Biol.* **9,** 514–520.

Frolova, L., Le Goff, X., Rasmussen, H. H., Cheperegin, S., Drugeon, G., Kress, M., Arman, I., Haenni, A. L., Celis, J. E., Philippe, M., *et al.* (1994). *Nature* **372,** 701–703.

Georgiou, G., Poetschke, H. L., Stathopoulos, C., and Francisco, J. A. (1993). *Trends Biotechnol.* **11,** 6–10.

Gersuk, G. M., Corey, M. J., Corey, E., Stray, J. E., Kawasaki, G. H., and Vessella, R. L. (1997). *Biochem. Biophys. Res. Commun.* **232,** 578–582.

Gillespie, J. H. (1984). *Evolution* **38,** 1116–1129.

Gold, L., Polisky, B., Uhlenbeck, O., and Yarus, M. (1995). *Ann. Rev. Biochem.* **64,** 763–797.

Gram, H., Marconi, L. A., Barbas, C. F., Collet, T. A., Lerner, R. A., and Kang, A. S. (1992). *Proc. Natl. Acad. Sci. USA* **89,** 3576–3580.

Grentzmann, G., Brechemier-Baey, D., Heurgue-Hamard, V., and Buckingham, R. H. (1995). *J. Biol. Chem.* **270,** 10595–10600.

Griffiths, A. D., Williams, S. C., Hartley, O., Tomlinson, I. M., Waterhouse, P., Crosby, W. L., Kontermann, R. E., Jones, P. T., Low, N. M., Allison, T. J., *et al.* (1994). *EMBO J.* **13,** 3245–3260.

Haeuptle, M. T., Frank, R., and Dobberstein, B. (1986). *Nucleic Acids Res.* **14,** 1427–1448.

Hajnsdorf, E., Braun, F., Haugel-Nielsen, J., Le Derout, J., and Regnier, P. (1996). *Biochimie* **78,** 416–424.
Hanes, J., Schaffitzel, C., Knappik, A., and Plückthun, A. (2000). Submitted.
Hanes, J., Jermutus, L., Schaffitzel, C., and Plückthun, A. (1999). *FEBS Lett.* **450,** 105–110.
Hanes, J., Jermutus, L., Weber-Bornhauser, S., Bosshard, H. R., and Plückthun, A. (1998). *Proc. Natl. Acad. Sci. USA* **95,** 14130–14135.
Hanes, J., and Plückthun, A. (1997). *Proc. Natl. Acad. Sci. USA* **94,** 4937–4942.
Hawkins, R. E., Russell, S. J., and Winter, G. (1992). *J. Mol. Biol.* **226,** 889–896.
He, M., Menges, M., Groves, M. A., Corps, E., Liu, H., Brüggemann, M., and Taussig, M. J. (1999). *J. Immunol. Methods* **231,** 105–117.
He, M., and Taussig, M. J. (1997). *Nucleic Acids Res.* **25,** 5132–5134.
Hellinga, H. W. (1997). *Proc. Natl. Acad. Sci. USA* **94,** 10015–10017.
Hentze, M. W. (1995). *Curr. Opin. Cell Biol.* **7,** 393–398.
Hermann, T., and Patel, D. J. (2000). *Science* **287,** 820–825.
Hermes, J. D., Blacklow, S. C., and Knowles, J. R. (1990). *Proc. Natl. Acad. Sci. USA* **87,** 696–700.
Hoogenboom, H. R. (1997). *Trends Biotechnol.* **15,** 62–70.
Hoogenboom, H. R., and Winter, G. (1992). *J. Mol. Biol.* **227,** 381–388.
Hurle, M. R., and Gross, M. (1994). *Curr. Opin. Biotechnol.* **5,** 428–433.
Jakobovits, A., Green, L. L., Hardy, M. C., Maynard-Currie, C. E., Tsuda, H., Louie, D. M., Mendez, M. J., Abderrahim, H., Noguchi, M., Smith, D. H., *et al.* (1995). *Ann. N. Y. Acad. Sci.* **764,** 525–535.
Janda, K. D., Lo, L. C., Lo, C. H. L., Sim, M. M., Wang, R., Wong, C. H., and Lerner, R. A. (1997). *Science* **275,** 945–948.
Janosi, L., Shimizu, I., and Kaji, A. (1994). *Proc. Natl. Acad. Sci. USA* **91,** 4249–4253.
Jermutus, L., Honegger, A., Schwesinger, F., Hanes, J., and Plückthun, A. (2000). Submitted.
Jermutus, L., Ryabova, L. A., and Plückthun, A. (1998). *Curr. Opin. Biotechnol.* **9,** 534–548.
Jung, S., Honegger, A., and Plückthun, A. (1999). *J. Mol. Biol.* **294,** 163–180.
Kanduc, D. (1997). *Arch. Biochem. Biophys.* **342,** 1–5.
Kasahara, N., Dozy, A. M., and Kan, Y. W. (1994). *Science* **266,** 1373–1376.
Kawasaki, G. (1991). *PCT Int. Appl.* WO 91/05058.
Keiler, K. C., Waller, P. R., and Sauer, R. T. (1996). *Science* **271,** 990–993.
Kepler, T. B., and Perelson, A. S. (1995). *Proc. Natl. Acad. Sci. USA* **92,** 8219–8223.
Kieke, M. C., Cho, B. K., Boder, E. T., Kranz, D. M., and Wittrup, K. D. (1997). *Protein Eng.* **10,** 1303–1310.
Kim, D. M., and Schwartz, J. R. (1999). *Biotechnol. Bioeng.* **66,** 180–188.
Knappik, A., Ge, L., Honegger, A., Pack, P., Fischer, M., Wellnhofer, G., Hoess, A., Wolle, J., Plückthun, A., and Virnekäs, B. (2000). *J. Mol. Biol.* **296,** 57–86.
Köhler, G., and Milstein, C. (1975). *Nature* **256,** 495–497.
Komar, A. A., Lesnik, T., and Reiss, C. (1999). *FEBS Lett* **462,** 387–391.
Komine, Y., Kitabatake, M., Yokogawa, T., Nishikawa, K., and Inokuchi, H. (1994). *Proc. Natl. Acad. Sci. USA* **91,** 9223–9227.
Korman, A. J., Knudsen, P. J., Kaufman, J. F., and Strominger, J. L. (1982). *Proc. Natl. Acad. Sci. USA* **79,** 1844–1848.
Kozak, M. (1989). *Mol. Cell. Biol.* **9,** 5134–5142.
Kozak, M. (1984). *Nature* **308,** 241–246.
Kraus, J. P., and Rosenberg, L. E. (1982). *Proc. Natl. Acad. Sci. USA* **79,** 4015–4019.
Kristensen, P., and Winter, G. (1998). *Fold. Des.* **3,** 321–328.
Kuchner, O., and Arnold, F. H. (1997). *Tibtech* **15,** 523–530.

Lancet, D., Sadovsky, E., and Seidemann, E. (1993). *Proc. Natl. Acad. Sci. USA* **90,** 3715–3719.
Leigh, E. G., Jr. (1973). *Genetics* **73,** Suppl 73:1–18.
Lin-Goerke, J. L., Robbins, D. J., and Burczak, J. D. (1997). *Biotechniques* **23,** 409–412.
Lonberg, N., Taylor, L. D., Harding, F. A., Trounstine, M., Higgins, K. M., Schramm, S. R., Kuo, C. C., Mashayekh, R., Wymore, K., McCabe, J. G., *et al.* (1994). *Nature* **368,** 856–859.
Low, N. M., Holliger, P. H., and Winter, G. (1996). *J. Mol. Biol.* **260,** 359–568.
Macken, C. A., and Perelson, A. S. (1989). *Proc. Natl. Acad. Sci. USA* **86,** 6191–6195.
Makeyev, E. V., Kolb, V. A., and Spirin, A. S. (1999). *FEBS Lett.* **444,** 177–180.
Malkin, L. I., and Rich, A. (1967). *J. Mol. Biol.* **26,** 329–346.
Marks, J. D., Hoogenboom, H. R., Bonnert, T. P., McCafferty, J., Griffiths, A. D., and Winter, G. (1991). *J. Mol. Biol.* **222,** 581–597.
Mattheakis, L. C., Bhatt, R. R., and Dower, W. J. (1994). *Proc. Natl. Acad. Sci. USA* **91,** 9022–9026.
Mattheakis, L. C., Dias, J. M., and Dower, W. J. (1996). *Methods Enzymol.* **267,** 195–207.
Matthews, D. J., and Wells, J. A. (1993). *Science* **260,** 1113–1117.
Moore, J. C., Jin, H. M., Kuchner, O., and Arnold, F. H. (1997). *J. Mol. Biol.* **272,** 336–347.
Nemoto, N., Miyamoto-Sato, E., Husimi, Y., and Yanagawa, H. (1997). *FEBS Lett.* **414,** 405–408.
Nemoto, N., Miyamoto-Sato, E., and Yanagawa, H. (1999). *FEBS Lett.* **462,** 43–46.
Nissim, A., Hoogenboom, H. R., Tomlinson, I. M., Flynn, G., Midgley, C., Lane, D., and Winter, G. (1994). *EMBO J.* **13,** 692–698.
Northrup, S. H., and Erickson, H. P. (1992). *Proc. Natl. Acad. Sci. USA* **89,** 3338–3342.
Osborne, S. E., Matsumura, I., and Ellington, A. D. (1997). *Curr. Opin. Chem. Biol.* **1,** 5–9.
Pedersen, H., Holder, S., Sutherlin, D. P., Schwitter, U., King, D. S., and Schultz, P. G. (1998). *Proc. Natl. Acad. Sci. USA* **95,** 10523–10528.
Pelletier, J. N., Arndt, K. M., Plückthun, A., and Michnick, S. W. (1999). *Nat. Biotechnol.* **17,** 683–690.
Pelletier, J. N., Campbell-Valois, F. X., and Michnick, S. W. (1998). *Proc. Natl. Acad. Sci. USA* **95,** 12141–12146.
Proba, K., Wörn, A., Honegger, A., and Plückthun, A. (1998). *J. Mol. Biol.* **275,** 245–253.
Ray, B. K., and Apirion, D. (1979). *Mol. Gen. Genet.* **174,** 25–32.
Rigler, R. (1995). *J. Biotechnol.* **41,** 177–186.
Roberts, R. W. (1999). *Curr. Opin. Chem. Biol.* **3,** 268–273.
Roberts, R. W., and Ja, W. W. (1999). *Curr. Opin. Struct. Biol.* **9,** 521–529.
Roberts, R. W., and Szostak, J. W. (1997). *Proc. Natl. Acad. Sci. USA* **94,** 12297–12302.
Roche, E. D., and Sauer, R. T. (1999). *EMBO J.* **18,** 4579–4589.
Roth, A., and Breaker, R. R. (1998). *Proc. Natl. Acad. Sci. USA* **95,** 6027–6031.
Ruan, B., Hoskins, J., Wang, L., and Bryan, P. N. (1998). *Protein Sci.* **7,** 2345–2353.
Rubingh, D. N. (1997). *Curr. Opin. Biotechnol.* **8,** 417–422.
Ruckman, J., Green, L. S., Beeson, J., Waugh, S., Gillette, W. L., Henninger, D. D., Claesson-Welsh, L., and Janjic, N. (1998). *J. Biol. Chem.* **273,** 20556–20567.
Ryabova, L. A., Desplancq, D., Spirin, A. S., and Plückthun, A. (1997). *Nat. Biotechnol.* **15,** 79–84.
Sasaki, A., and Isawa, Y. (1989). *Theor. Popul. Biol.* **39,** 201–239.
Schier, R., and Marks, J. D. (1996). *Hum. Antibodies Hybridomas* **7,** 97–105.
Schmidt-Dannert, C., and Arnold, F. H. (1999). *Trends Biotechnol.* **17,** 135–136.
Schultz, P. G., and Lerner, R. A. (1995). *Science* **269,** 1835–1842.
Shao, Z., Zhao, H., Giver, L., and Arnold, F. H. (1998). *Nucleic Acids Res.* **26,** 681–683.

Shen, X.-C., Yao, S.-L., Terada, S., Nagamune, T., and Suzuki, E. (1998). *Biochem. Eng. J.* **2,** 23–28.
Shine, J., and Dalgarno, L. (1975). *Nature* **254,** 34–38.
Sieber, V., Plückthun, A., and Schmid, F. X. (1998). *Nat. Biotechnol.* **16,** 955–960.
Smith, G. P. (1985). *Science* **228,** 1315–1317.
Smith, W. P., Tai, P. C., and Davis, B. D. (1978). *Proc. Natl. Acad. Sci. USA* **75,** 5922–5925.
Spada, S., Honegger, A., and Plückthun, A. (1998). *J. Mol. Biol.* **283,** 395–407.
Spetner, L. M. (1970). *Nature* **226,** 948–949.
Stemmer, W. P. (1994). *Nature* **370,** 389–391.
Tarasow, T. M., Tarasow, S. L., and Eaton, B. E. (1997). *Nature* **389,** 54–57.
Tate, W. P., and Brown, C. M. (1992). *Biochemistry* **31,** 2443–2450.
Tawfik, D. S., Green, B. S., Chap, R., Sela, M., and Eshhar, Z. (1993). *Proc. Natl. Acad. Sci. USA* **90,** 373–377.
Tawfik, D. S., and Griffiths, A. D. (1998). *Nat. Biotechnol.* **16,** 652–656.
Thanaraj, T. A., and Argos, P. (1996). *Protein Sci.* **5,** 1594–1612.
Tuerk, C., and Gold, L. (1990). *Science* **249,** 505–510.
Tuite, M. F., and Stansfield, I. (1994). *Mol. Biol. Rep.* **19,** 171–181.
Ueda, T., Tohda, H., Chikazumi, N., Eckstein, F., and Watanabe, K. (1991). *Nucleic Acids Res.* **19,** 547–552.
Vaughan, T. J., Williams, A. J., Pritchard, K., Osbourn, J. K., Pope, A. R., Earnshaw, J. C., McCafferty, J., Hodits, R. A., Wilton, J., and Johnson, K. S. (1996). *Nat. Biotechnol.* **14,** 309–314.
Watanabe, K., Oshima, T., Iijima, K., Yamaizumi, Z., and Nishimura, S. (1980). *J. Biochem.* (*Tokyo*) **87**(1), 1–13.
Winter, G., Griffiths, A. D., Hawkins, R. E., and Hoogenboom, H. R. (1994). *Ann. Rev. Immunol.* **12,** 433–455.
Yang, W. P., Green, K., Pinz-Sweeney, S., Briones, A. T., Burton, D. R., and Barbas, C. F., 3rd (1995). *J. Mol. Biol.* **254,** 392–403.
Yelton, D. E., Rosok, M. J., Cruz, G., Cosand, W. L., Bajorath, J., Hellstrom, I., Hellstrom, K. E., Huse, W. D., and Glaser, S. M. (1995). *J. Immunol.* **155,** 1994–2004.
Yu, A., Barreiro, V., and Haggard-Ljungquist, E. (1994). *J. Virol.* **68,** 4220–4226.
Zaccolo, M., Williams, D. M., Brown, D. M., and Gherardi, E. (1996). *J. Mol. Biol.* **255,** 589–603.
Zhao, H., and Arnold, F. H. (1997). *Curr. Opin. Struct. Biol.* **7,** 480–485.
Zhao, H., Giver, L., Shao, Z., Affholter, J. A., and Arnold, F. H. (1998). *Nat. Biotechnol.* **16,** 258–261.
Zhouravleva, G., Frolova, L., Le Goff, X., Le Guellec, R., Inge-Vechtomov, S., Kisselev, L., and Philippe, M. (1995). *EMBO J.* **14,** 4065–4072.