Protein Design and Phage Display

Ronald H. Hoess*

Dupont Pharmaceuticals Company, Experimental Station E336/205, Wilmington, Delaware 19880

Received February 2, 2001

Contents

I. Introduction 3205
II. Phage Display 3206
A. Biology of Filamentous Phage 3206
B. Basic Display Formats 3206
C. Library Construction 3209
D. Affinity Selection Methodology 3209
III. Selection Based on Binding 3209
A. Affinity Maturation 3209
B. DNA-Binding Proteins 3210
C. Scaffold Proteins 3211
IV. Phage Display and Folded Proteins 3211
A. Understanding Protein Folding Using Phage Display 3211
B. Applications for Engineering Proteins 3213
V. Phage Display and Catalytic Function 3214
A. Selection Using Transition-State Analogues 3214
B. Selection Using Suicide Inhibitors 3215
C. Selection Linked to Catalysis 3215
VI. Conclusions 3217
VII. Acknowledgments 3217
VIII. References 3217

I. Introduction

The application of phage display for de novo design of proteins appears at first to be an unusual combination of technologies. Recently, however, the power of combining the principles of the two technologies has become increasingly apparent. This review will address recent uses of phage display that can be applied to de novo design problems. Protein design and phage display appear to employ very different principles for elucidating the mechanisms that direct protein folding. Protein design connotes a rational approach to engineering protein structure and function with rules based on first principles of protein folding.[1] The de novo approach has been successful in that a number of rationally designed proteins have been assembled. The design process is relatively slow because it is an iterative process and each design must be characterized biophysically before making the next change until a final design is complete. Increasingly, much of the work is guided by the known structural database and the ability to model designed structures.[2]

* To whom correspondence should be addressed. Phone: (302) 695-7553. Fax: (302) 695-9420. E-mail: ronald.h.hoess@dupontpharma.com.

Ronald H. Hoess was born in Philadelphia, Pennsylvania, in 1948. He received his B.S. degree in Botany from Pennsylvania State University in 1970 and his Ph.D. degree in Genetics from the University of Minnesota in 1975. After postdoctoral work at the Fox Chase Cancer Center and Brown University, he joined the Frederick Cancer Research Center in 1979. There he began working on the Cre-*lox* site-specific recombination system. He continued this work after moving to Dupont Central Research and Development in 1984. In 1991, upon joining what is now Dupont Pharmaceuticals, he began working on phage-display technology. He has published over 60 papers in areas including site-specific recombination, protein folding and phage display. Currently he is a Senior Investigator at Dupont Pharmaceuticals and an Adjunct Associate Professor in the Department of Biochemistry and Biophysics at the University of Pennsylvania Medical School.

By contrast, phage display at first appears to be a totally random approach. Large combinatorial libraries of encoded peptides or proteins are displayed as fusions to the phage capsid. Phage displaying a peptide or protein with a desired property can be selected from the library and decoded by sequencing the phage DNA.[3] Unlike de novo design, selections from a random library are frequently completed in a matter of weeks. An understanding of why certain sequences are selected is only achieved after the experiment is completed and analysis of the resulting structures can be performed.

Ultimately phage display can be applied in a more rational way. Although it is possible to construct large combinatorial libraries, it has become apparent that it is not possible to simply select from such libraries a folded and functional protein by weeks end. While there are other technologies that can surpass phage-display libraries in size and complexity,[4,5] the sequence space even for a relatively small protein is much larger than any library that can be constructed. Thus, a rational approach is required for the application of phage display to protein engineering. Successful application of the technology has been

10.1021/cr000056b CCC: $36.00
Published on Web 08/18/2001

guided by structural information analogously to the way these principles are employed for de novo design.

The initial work on phage-displayed proteins utilized a broad range of proteins found in nature. These proteins provided a framework for studying various aspects of structure and function. After a decade of work, it is reasonable that phage-display technology can be joined with de novo design for the creation of novel proteins. In a logical progression of steps, designed proteins can be refined using phage display to (i) select for properly folded structures, (ii) evolve a binding function, or (iii) evolve a catalytic function. This review will focus on studies that have combined phage display of proteins with protein design to achieve these goals.

II. Phage Display

A. Biology of Filamentous Phage

Nearly all of the work on protein engineering using phage display has been done using the filamentous phage M13 or the closely related phage fd. This is in part due to the ease with which the phage can be manipulated and also to our detailed understanding of the viral life cycle and phage structure. Upon infection of its host, the single-stranded M13 genome is first converted to a double-stranded replicative form which serves as a template for the production of viral proteins and single-stranded DNA progeny.[6] The single-stranded DNA is eventually extruded from the host cell through the inner membrane. During this process the viral coat proteins, which are anchored in the membrane, encapsulate the DNA as it traverses the membrane. The released virus is a long flexible rod about 1 μm in length. Most of the viral coat consists of the major coat protein gp8 of which there are 2700 copies per phage. At one end of the virus are five copies of the minor coat proteins gp3 and gp6, and at the opposite end are two other minor coat proteins gp7 and gp9 (Figure 1). All of the coat proteins are synthesized with N-terminal signal sequences that direct them via the secretory machinery to the membrane.

NMR and X-ray structures have recently been reported for the minor capsid protein gp3.[7−9] Since gp3 has been the primary scaffold of choice for the display of proteins, these structures have been useful for understanding function and the relationship to displayed proteins. Structurally gp3 is comprised of three separate domains connected by glycine-rich linker regions. The amino terminal domain (N1) and the middle domain (N2) have important roles in viral infection. N2 is responsible for binding to the tip of the F pilus of the *E. coli* host.[10] Once the phage is bound, the pilus retracts bringing the viral particle close to the cell surface.[11] At the cell surface N1 binds the TolA receptor which then leads to penetration of the host by the viral DNA.[12−14] If the N1 and N2 domains are deleted, the resulting viral particles are noninfectious.[10,15] The carboxy-terminal domain (CT) is essential for viral morphogenesis and cannot be deleted.[16,17] Interestingly, carboxy-terminal fusions to gp3 have also been shown to be tolerated despite the fact that the CT domain appears buried in the viral capsid.[18]

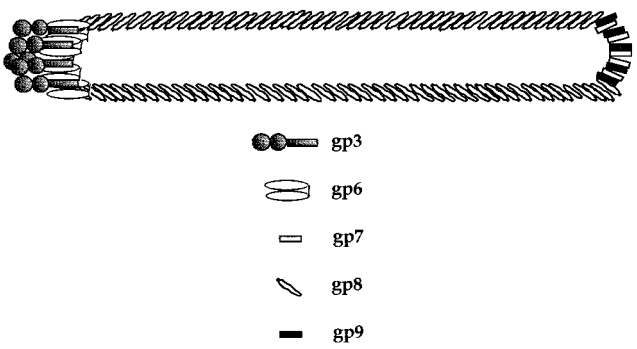

Figure 1. Schematic diagram of the filamentous phage capsid structure. The capsid is comprised primarily of 2700 copies of gp8. At one end, shown here to the right, is the site for initiation of capsid formation where there are five copies each of the small hydrophobic proteins gp7 and gp 9. At the opposite end of the capsid are two additional minor coat proteins, gp3 and gp6, each in five copies. All of the capsid proteins have been utilized for the display of proteins.

The major coat protein gp8 has also been characterized both by X-ray crystallography and by NMR.[19,20] The 50 amino acid protein forms a slightly bent α helix when assembled on the viral capsid. Displayed peptides have been fused to the amino terminus, although total replacement of all 2700 copies of wild-type gp8 with fusion gp8 can only occur with peptides up to six amino acids in length.[21−24] This is an obvious limitation for the display of proteins fused to gp8, with the number being displayed per phage considerably less than 2700. One explanation for this limitation is that the pore structure which the virus must pass through in order to be extruded from the cell is not large enough to accommodate a virion where all gp8 molecules have an additional protein domain fused to them.

The remaining coat proteins have not been characterized structurally, but this has not prevented their use as vehicles for protein display. Because of its orientation on the viral particle, gp6 has been used to display protein fusions at its C-terminus rather than the typical N-terminal fusions of the other coat proteins. This makes it amenable for the display of cDNA-encoded libraries.[25] Recently gp7 and gp9 have been shown to tolerate fusions at their amino termini. Since the two proteins appear to interact with one another, they are ideal for display of dimeric proteins such as antibodies.[26]

B. Basic Display Formats

For the display of proteins on the surface of phage, DNA encoding the protein is inserted between the signal sequence of the coat protein and the amino terminus of the mature coat protein (the signal sequence is proteolytically removed upon secretion). The first peptide display libraries were cloned directly into the viral genome so that all five copies of gp3 display the peptide fusion. There are a number of concerns associated with displaying proteins using this strategy. Fusion of an entire protein domain to gp3 may compromise the infectivity function of gp3. In addi-

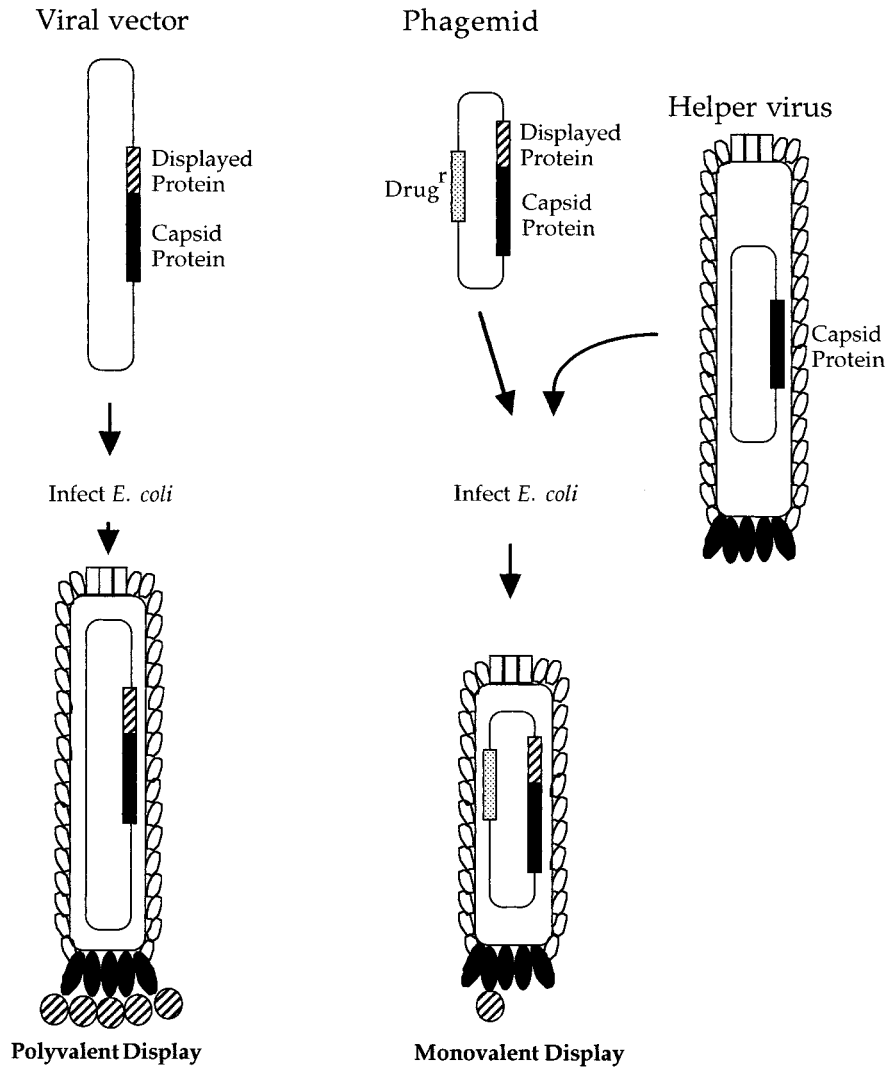

Figure 2. Types of phage-display vectors. Proteins fused to capsid proteins can be displayed in either of two formats. Shown on the left, the protein can be cloned directly into a viral vector as a fusion to a capsid protein resulting in every copy of the capsid protein displaying the fusion (polyvalency). Alternatively, the protein fusion can be constructed in a phagemid vector that carries a copy of the viral capsid gene. When cells containing the phagemid are infected with a helper virus that carries a wild-type copy of the capsid protein, the rescued phagemid DNA is packaged in viral particles that display both wild-type and fusion capsid protein. By adjusting the level of expression, the fusion can be displayed on average as one copy per particle (monovalency). A small number of helper phage are released following infection that also display the protein fusion monovalently. Since the helper phage do not carry a drug resistance marker, they are lost upon subsequent selection for drug resistance.

tion, the ability to select certain domains may be related to valency of the displayed protein. Both of these issues were elegantly addressed with the development of phagemid vectors for protein display.[27]

Phagemid vectors were originally constructed for isolating single-stranded plasmid DNA for sequencing. These vectors contain an M13 origin of replication and a phage packaging site in addition to the elements required for plasmid propagation. When phagemid-containing cells are infected with a helper phage, the phagemid switches to rolling circle replication and forms single-stranded copies that are packaged as virions with the capsid proteins being supplied by the helper. As shown in Figure 2, DNA encoding a protein fused to gene 3 is cloned into a phagemid vector. When cells containing this phagemid are infected with helper phage, both the gp3 fusion and the wild-type copies of gp3 encoded by the helper are expressed. The ratio of these two proteins will dictate what is ultimately displayed by the packaged phagemids. In general, expression of the fusion is kept lower than the wild-type gp3 so that every virion has at least some wild-type gp3. On average, there is only one gp3 fusion per viral particle. Two advantages are accrued from this type of display format: (i) infectivity is not compromised by appending a protein to gp3 since wild-type copies of gp3 are also provided and (ii) the valency of display is reduced which can be advantageous during the selection process.

Phagemid vectors have been constructed for display of fusions to gp8, gp6, gp7, and gp9. In the case of gp8, the phagemid display format results in numerous copies (100−200 per phage) of the displayed fusion. Recently phage-display-based protein design has been used to increase the level of gp8 protein fusions displayed on phage.[28,29] A fusion of human

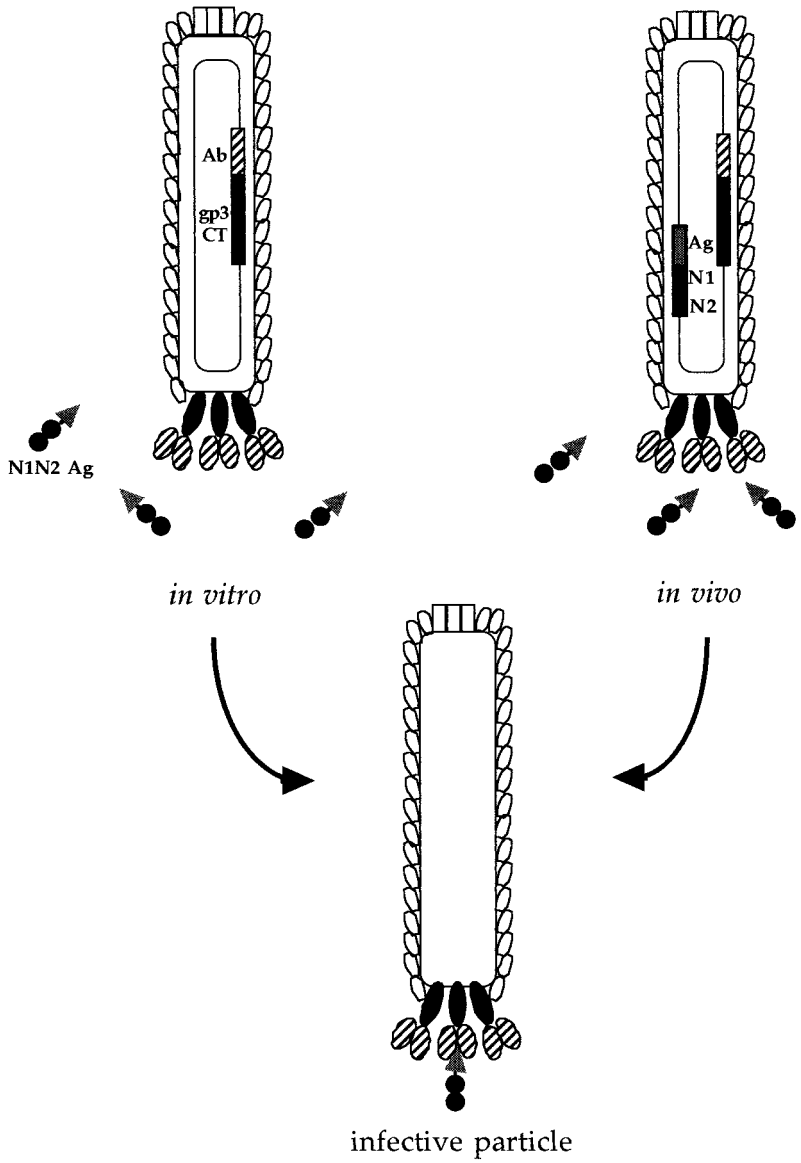

Figure 3. Selectively infective phage (SIP). Phage particles are produced that display an antibody domain (Ab) as a fusion to the carboxy terminal domain (CT) of the capsid protein gp3. Because they lack the amino terminal domains (N1+N2) of gp3 they are noninfective. If these phage are incubated in vitro with N1+N2 fused to an antigen (Ag) recognized by the displayed antibody, the physical association of N1+N2 will regenerate an infective phage. The same can also be accomplished in vivo if the N1+N2−Ag are cloned directly in the viral genome. In this case association of N1+N2−Ag with Ab fused to CT takes place in the periplasmic space following their secretion through the inner membrane. This strategy can be used to select library (antigens) against library (antibodies).

growth hormone (hGH) to gp8 was used to select for gp8 mutations, resulting in higher levels of hGH display. Surprisingly, many mutations can be accommodated without loss of function. Phage were selected that increased the signal of displayed hGH 100-fold. However, a direct correlation between the increase in ELISA signal and increased display of hGH was not demonstrated in these experiments.

An innovative selection format recently developed is selectively infective phage (SIP).[30−32] The principle underlying this technique takes advantage of the modular nature of gp3. Infectivity of phage deleted for domains N1+N2 can be rescued if the phage can become reassociated physically with N1+N2 (Figure 3). For example, if N1 and N2 are expressed as a fusion with an antigen and the CT domain on the phage is fused to an antibody that recognizes the antigen, infectivity is restored through antibody−antigen interaction. This type of interaction is not limited to antibody−antigen but can be done with any two interacting partners with sufficient affinity to hold the domains together. SIP can be done either in vivo or in vitro. In one version both N1+N2 and the CT domain are expressed from a single phage.[32] A library can be constructed where fusions are made either to N1+N2 or to the CT domain. From such a library only members that interact to bring N1+N2 protein segments and CT together are infectious. In vitro experiments using purified N1+N2 protein segments can be done in solution without the need for a solid support or elution steps to capture bound phage. Furthermore, the N1+N2 fusion need not necessarily be a peptide or protein but can be a chemical hapten.[33]

C. Library Construction

One of the most attractive features of phage display is the ability to make large libraries of mutants for a given protein. A judicious choice must be made as to how to construct such a library so that a meaningful search of sequence space will be conducted. The most difficult case is when there is no a priori knowledge of where best to focus mutagenesis, i.e., location of a functional site. In these cases, mutagenesis is frequently done randomly throughout the gene using error-prone PCR. Since multiple random mutations can lead to nonfunctional products, it is necessary to combine error-prone PCR with DNA shuffling to select functional proteins. With DNA shuffling, a pool of mutants are recombined in vitro and mutations with positive effects are selected. Positive clones are then shuffled with wild-type DNA to dilute out mutations that are either neutral or deleterious.[34]

Frequently libraries of mutants are focused to a limited region of the protein so that a small number of amino acid residues can be mutated. The number of residues chosen will dictate the strategy for assembling the library. Despite the relatively large size of libraries that can be constructed by phage display, the randomization of more than six residues rapidly exceeds the ability to construct a complete library with all possible combinations of sequences. A technique that has been applied very effectively to a number of different design problems is the construction of several libraries where each library randomizes only a few amino acids. Selection is carried out with each library, and then the best variants from each library are combined for further selection. The probability that one very large initial library could have contained the sequences that were finally selected is very low. The only potential limitation of this approach is the assumption that each sublibrary can be improved independently of changes elsewhere in the protein.

For randomization of individual amino acid positions, oligonucleotides are synthesized with NNG/C or NNG/T codons that are able to code for all 20 amino acids and one stop codon. More elaborate strategies have been used to try and limit a particular codon to a class of amino acids, i.e., hydrophobic or polar. This is inherently difficult because synthesis of oligonucleotides is usually done on solid supports one nucleotide at a time and the nature of the genetic code makes it difficult to code for one particular class of amino acids. Recently oligonucleotides have been synthesized as codons which can be subsequently assembled to code for any selection of amino acid residues.[35] This method has been used for construction of phage-displayed synthetic human antibodies in which the complementarity determing region (CDR) loops have been randomized using trinucleotide codons. In principle, this method provides a more effective search of sequence space by avoiding sequences that will not form folded structures i.e., polyproline loops.

D. Affinity Selection Methodology

The selection process used to retrieve members with desired properties from a library is as important as the construction of the library. The most commonly used selection is affinity based and usually referred to as biopanning. Typically the selective agent is immobilized on a solid support. The phage library is then incubated with the support to allow binding between the immobilized target and the appropriate phage. After a suitable incubation time, unbound phage are removed and washing steps are done to remove nonspecifically bound phage. Elution of the specifically bound phage can be accomplished by brief incubation at low pH 2.2[36] or alkaline buffers such as 0.1 M triethylamine pH12.[37] The elution is necessarily harsh since the objective is to recover the tightest binding members of the library. The phage particles themselves are impervious to these relatively harsh conditions for a limited period of time and do not lose viability. To enrich for specifically bound phage, particularly in the early rounds of selection, competitive or noncompetitive agents that release only phage specifically associated with the target have been used.

The stringency of selection dictates which phage are selected from the library. Conditions for time, temperature, and number of washes are usually arrived at empirically by the investigator although there has been an attempt to model conditions based on the affinity between phage-displayed peptide and immobilized receptors.[38] Early rounds of selection are usually done under mild conditions followed in subsequent rounds by more stringent rounds of selection. This is to ensure that even low-affinity binders are captured in the early rounds.

Avidity may play an important role in phage-display selection. This results from selection in a nonhomogeneous environment (the selective agent is tethered to a solid support) and the multivalency of the phage. Multivalency does not allow discrimination between high-affinity binders and those with only modest affinities.[27] On the other hand, monovalent display using phagemid vectors allows the two classes of binders to be distinguished. It is often necessary to use polyvalent display to begin affinity evolution when the initial binding affinities are quite low. Polyvalent display may also be important when the initial library is unlikely to contain all possible sequences and therefore may not contain a high-affinity binder.

Avidity effects can also be minimized by selecting binding molecules in solution. This is usually done as a two-step process in which the phage library and ligand are first allowed to bind in solution. A capture step is introduced which quickly brings ligand and any bound phage to a solid support. Commonly the ligand is biotinylated and capture is done via streptavidin coated beads or a solid support.[36]

III. Selection Based on Binding

A. Affinity Maturation

The principle of selection as first demonstrated with phage-displayed peptides is based on searching the library for those members with the highest affinity for the targeted receptor. A logical extension of this principle is the engineering of phage-displayed

proteins to increase the affinity for a target molecule. The first example to integrate phage display with protein design was the selection of mutants of hGH that bind its receptor with high affinity.[39] These important experiments demonstrated the utility of phage display for protein engineering and were the basis for many of the techniques and strategies used to display other proteins.

hGH is an α-helical protein of 191 amino acids that binds growth hormone receptor. Structural analysis of hGH with its receptor coupled with extensive alanine scanning data of hGH itself had identified residues of hGH critical for binding receptor. These residues are found primarily on two α-helices and an extended loop region. Because the number of residues involved in binding far exceeds the ability to build a complete library if all residues were randomized at once, initial experiments subdivided the binding surface into three regions.[39] Three separate libraries were made with each randomizing four codons. A consensus sequence was selected from each library with some of the variants showing modest increases in receptor binding affinity. Variants selected from each of the sublibraries were then combined and further affinity selected.[40] In this manner, the overall affinity of the final variant with 15 amino acid substitutions bound 400-fold tighter to the receptor than wild-type hGH. This work demonstrated the importance of incremental evolutionary steps for improving binding properties.

Selection for higher affinity binders mimics the process by which the immune system selects antibodies through affinity maturation. These strategies that mimic the immune system were quickly adapted to phage-displayed antibody repertories with emphasis placed on reassorting heavy and light chains between rounds of selection to expand the diversities of the libraries. There have been numerous examples where this has proved to be effective for identifying antibodies with affinities in the low nanomolar range. In some cases, shuffling of light and heavy chains may not be sufficient for attaining the desired affinity and a mutagenic step must be introduced.

Given the wealth of antibody structural data available it might appear straightforward where to localize mutagenesis for affinity improvement. This has proven to be a far more complex problem, and extensive literature concerning in vitro affinity maturation exists. Described here are a few illustrative examples of the types of approaches that have been used. For a more comprehensive discussion, the reader is referred to a number of reviews.[41,42] The most general approach is to mutagenize the entire variable domain by growing the phage-expressed antibodies in a bacterial mutator strain.[43] This approach makes no a priori assumptions as to which mutations might improve affinity. The drawback is the vast amount of sequence space that may be required to find such rare mutations. Most other approaches have excluded the framework and focused on the CDR loops. Even this limitation does not allow construction of a phage-display library to cover the entire sequence space. One solution is to randomize a single CDR and select for improvement.[44,45] When randomizing more than one CDR loop it is possible to optimize in either a sequential or parallel manner.[46] In one study CDR loops were mutagenized independently and then incorporated in the parental Fab fragment (parallel) or individual CDRs were mutagenized in sequential fashion. While both methods showed improvements in binding affinity, mutations were not strictly additive using the parallel approach. This suggests underlying context effects must play a role so that each CDR loop cannot be viewed as an independent unit. The most focused mutagenesis approach involves creation of hot spot libraries which target DNA sequences in the variable regions most prone to hypermutation.[47] Using modest sized libraries of $\sim 10^4$ individuals, antibodies with improved affinity have been selected.

While antibodies have been the major focus of affinity maturation, other proteins have also been improved using phage display. Among these are human interleukin 6,[48] SH3 domains,[49] and β-lactamase inhibitory protein.[50] These experiments have been guided by structure and use a rational approach for construction of the phage-displayed library.

B. DNA-Binding Proteins

The development of proteins with new DNA-binding specificities is a long-standing goal in protein engineering. Numerous structures of DNA−protein complexes demonstrate that a variety of structural motifs are utilized for DNA recognition. This picture is further complicated by the apparent lack of a recognition code where there is a unique correspondence between amino acid and a specific base.

The discovery of zinc finger domains served as an ideal starting point from which the problem of DNA recognition might be solved. Zinc fingers are small structural units of ∼30 amino acids. Each unit or finger assumes a ββα fold stabilized by hydrophobic interactions and the chelation of a zinc molecule. A single zinc finger contacts three contiguous bases in the major groove via residues projecting from the α helix. In nature zinc fingers are found as multiples of the finger module in order to extend recognition over multiples of three bases. Considerable effort has been made in tabulating naturally occurring zinc fingers and their cognate DNA-binding sequences in order to try and deduce a code for DNA recognition.

Several laboratories have used phage display to select for zinc fingers with new DNA recognition specificities.[51−55] The basis for this work was the display of Zif268, a zinc finger of known DNA specificity and for which a crystal structure was available. Phage displaying Zif268 selectively bound a biotinylated oligonucleotide containing the recognition sequence which could subsequently be captured with streptavidin. The goal of the experiments was to change one of the codons of the recognition sequence and randomize the corresponding finger that contacts that particular codon. The displayed libraries that were constructed represented two approaches, one where only positions known to contact the base were randomized[51,52] and one where additional amino acid positions adjacent to those that contact the base.[53,55] The second library was built to

address local context effects that may influence recognition and binding. Novel zinc finger domains were isolated that bound a new codon sequence more effectively than the wild-type codon sequence. In addition, specific amino acids appeared to play specific roles in DNA binding as suggested by structural data. Nevertheless, despite using very large libraries, there were codons for which tight binding zinc fingers could not be found.

The ultimate goal of this work is to design a zinc finger protein that will recognize any specific DNA sequence. Much of the early work was based on the assumption that each finger module acted independently of the adjacent fingers and that modules could be mixed and matched to recognize a specific DNA sequence. Inspection of the crystal structures reveals that there are many interactions between neighboring modules indicating that context is important. More recently phage-displayed zinc fingers have been selected using a modified method in which selection is done in sequential fashion, first altering one finger and then altering the next finger.[56,57] In this sequential manner mutants are selected not only on the basis of their DNA recognition, but also by their contacts with neighboring fingers. Alternatively, two fingers can be randomized simultaneously.[58] Other studies have focused on zinc fingers that recognize the 5′GNN3′ family of codons[59,60] or DNA containing 5-methylcytosine.[61] Further refinement of these new zinc fingers has come to rely on molecular modeling and site-directed mutagenesis experiments. In the absence of a simple recognition code, selection via phage display will continue to be critical for design of zinc fingers with new binding specificities.

One other DNA-binding protein that has been displayed on phage is Cro the λ bacteriophage repressor.[62] Since Cro binds DNA as a dimer, a single-chain construct was made by tethering two monomer coding sequences together with a linker sequence. Phage displaying the single chain Cro bound specifically to its operator. Mutants were isolated that bound a new operator sequence, although binding affinities were not reported.

C. Scaffold Proteins

The examples discussed thus far have dealt with modifications of preexisting functions that were being engineered. A different approach has been to modify proteins to introduce a completely novel function. Much of this work has centered on trying to identify surrogate antibody molecules that bind selectively to a specific ligand. The method involves selecting a protein framework and then randomizing a portion, typically a loop, that could serve as a structural binding motif. A number of proteins have been used as scaffolds including cytochrome b562,[63] Tendamistat,[64] Zn fingers,[65] the Z domain of protein A,[66] the Kunitz domain,[67] bovine pancreatic trypsin inhibitor,[68] fibronectin type III domain,[69] knottins,[70] and cytotoxic T lymphocyte-associated antigen 4.[71] While all of these scaffolds were shown to bind to selected target molecules, it appears unlikely that any one of them will represent a universal library similar to antibodies. This probably results from limitations of the protein architecture being used since these scaffolds only partially imitate natural antibody architecture. One interesting application of protein scaffolds is the minibody, a minimized design of an immunoglobulin variable domain. Phage-displayed minibodies that bind IL-6 have been selected.[72−74]

An interesting departure from the usual scaffold proteins that are restricted to presentation of one or two variable loops is the protein lipocalin.[75] The lipocalins are a family of β barrel proteins that bind small molecules in a pocket at the center of the barrel structure. The lipocalin from *Pieris brassicae* was displayed on phage, and a library was constructed in which 16 residues at the center of the binding site were varied. The library was then used to select for variants that bind fluorescein. Lipocalin molecules with K_d's in the low nanomolar range were selected and shown to be very specific for the target fluorescein molecule. This work has been extended to selection for binding to digoxigenin where variants with modest affinity were selected and then evolved to yield high-affinity mutants by a second round of focused mutagenesis.[76]

IV. Phage Display and Folded Proteins

A. Understanding Protein Folding Using Phage Display

Evolution of binding affinities by phage display is a natural extension of the principles of biopanning. In contrast, the utility of phage display for studying protein folding was not initially obvious. The ability to select for properly folded proteins would be invaluable, particularly in the area of de novo design where phage display could serve as a fine-tuning device to help move from either a loosely folded or a molten globule state to a more nativelike structure. To date, phage-display technology has not been directly applied to de novo designed proteins but has seen increasing use as a tool for understanding protein folding and stability. While all of these studies have been done with native proteins, the underlying principles also apply to designed proteins.

One of the first reports that indicated phage display would be a useful tool for protein folding studies was the display of the IgG binding B1 domain of *Streptococcus*.[77] Phage displaying the B1 domain give a phenotypically smaller plaque morphology than phage that do not display the B1 domain. Occasionally plaques were found that appear larger, similar to wild-type phage. Sequencing of the B1 domain from larger plaque phage invariably indicated the presence of mutations that would destablize the protein. Some were deletions that abolished binding to IgG altogether, while others were more subtle amino acid substitutions that still retained IgG binding. A number of these mutant B1 domains were characterized physically and shown to be thermodynamically destabilized. The connection between plaque morphology and protein stability remains obscure, but the study was important for showing that a range of folding stabilities could be obtained for a protein domain displayed on phage. A somewhat more direct approach to the issue of protein stability was taken

using the related IgG-binding domain L1.[78] This approach represents the inverse of the strategy discussed in the previous section for enhanced binding. Instead of mutating residues directly involved in ligand binding, residues outside of the binding site were mutated. By holding the binding site residues fixed, one can select for those proteins that retain the ability to bind ligand since binding requires a folded protein. Libraries were constructed in two β strands at the beginning of the domain where 14 positions were randomized. From a library of 10^6 mutants, 1 out of 200 were still able to bind, suggesting considerable plasticity in sequences tolerated for correct folding. While it is possible that most of the residues involved in the randomization were not critical for folding, thereby leading to a large solution set, the study illustrated the power of phage display for rapidly sorting through large numbers of mutants to identify those that are folded. More extensive mutagenesis of protein L was carried out where two β-turns, a helix, and two β strands were randomized.[79] Individuals from this library were selected for IgG binding and then characterized. Once again, functional clones were found with numerous amino acid substitutions. Interestingly, many of the proteins were able to fold as fast or faster than the wild-type protein even though they were destabilized overall. This suggests that on the time scale of the selection process stability rather than the kinetics of folding is the driving force of the selection.

One of the features that makes phage display such a powerful tool is the control that can be exercised over selection conditions. An example of this is a study to look at the contribution of turn sequences to the thermodynamic stability of a folded protein.[80] In this case two parameters were controlled by the investigator, the overall thermodynamic stability of the starting protein scaffold and the temperature at which selection for function (binding = folded) was done. Using the B1 domain as a protein scaffold, a series of turn libraries were constructed consisting of two, three, or four amino acid residues. All three libraries were cloned in a wild-type framework and in two mutant domains that were folded but destablizied to varying degrees. Selection for IgG binding was then carried out at room temperature or at increasingly higher temperatures. As might be expected in the wild-type framework, the majority of clones from each library were able to bind IgG. However, as the stringency of selection was increased, either by destabilizing the scaffold or by increasing the temperature at which binding took place, the percentage of functional clones dropped precipitously. The sequences obtained for the most stable turns reflect the statistical preferences obtained from the structural database, suggesting that the experiments are able to mimic some of the evolutionary pressures for thermodynamically stable turns. Similar experiments were also done on two turns in protein L.[81] Here kinetic analysis of some of the mutants suggests that each turn may play a different role and that the context of where the turn is located is probably important. These strategies could easily be applied to de novo designed proteins, so that rather than trying to evolve the protein in one step, the evolutionary process could be done in modular fashion. For example, libraries of turn sequences or helices could be built and selected in a stepwise fashion for increased thermostablity.

Another example of using phage display to select for protein stability was done using the prodomain of subtilisin BPN′.[82] The prodomain by itself is unfolded, and only upon binding subtilisin does it fold into a stable conformation. Since increased stability of the folded prodomain will lead to increased affinity for subtilisin, the prodomain was displayed on phage and mutants with improved binding were selected. Two structural motifs within the prodomain were selected for randomization: four residues within the interior and four residues comprising a turn. Results of these selections indicated that only a limited number of amino acid combinations confer maximum stability and are similar to the results obtained with the destabilized forms of the B1 domain where only a limited set of residues in the turn are able to confer thermodynamic stability.

Despite the large library sizes that can be constructed by phage display, most libraries represent only a tiny fraction of sequence space. Thus, it is unlikely that a functional protein could be evolved from such libraries. A number of attempts have been made to evolve proteins from a reduced amino acid code and to restrict the potential number of amino acids at any given position.[83] This strategy limits the amount of sequence space to be explored. While the initial experiments were not done in a phage-display format, the technology readily lends itself to such experiments. Along these lines the 57 amino acid src SH3 domain which recognizes a proline-rich peptide sequence was displayed on phage.[84] Libraries were then constructed where residues involved in binding the peptide were fixed but all other residues were varied. The varied residues were randomized with only five possible amino acid substitutions. Even this drastic reduction would require a library much greater in size than is technically feasible. Borrowing from the strategy used to select proteins with higher binding affinity, three separate libraries representing a third of the SH3 domain were constructed. Following selection for binding to the peptide ligand from these sublibraries, the selected clones were recombined to generate proteins with mutations throughout the domain. Astonishingly, variants were identified that had 40 out of 45 positions reduced to the five amino acid simplification code. This work has important implications for the evolution of protein folding. It is doubtful that the results could have been obtained without the ability to select among millions of variants using phage-display technology.

The examples discussed thus far are based on selecting binders from a population of nonbinding and presumably unfolded proteins. An important advance would be a method to select for proteins with increased stability from a pool of already folded variants. Due to the robust nature of the phage particle, selections can be done under conditions that will denature mostly stable proteins but still maintain phage viability. Starting with a well folded and

stable scFv, a library of random mutants in the scFv were displayed on phage and selected for binding at increasingly higher temperatures or increasing concentrations of guanidinium chloride.[85] Because scFv molecules tend to unfold irreversibly at high temperatures, the selections using temperature proved effective in identifying mutants with higher stability. Selection in the presence of guanidinium chloride proved less effective. This is probably because the binding selection itself cannot be done at high concentrations of guanidinium chloride and unfolding is reversed as the guanidinium concentration is lowered. Nevertheless, phage are able to survive under conditions that would prove too harsh for most proteins, making this a promising avenue for selecting higher protein stability.

Ultimately for the technology to have wider applicability in the area of protein folding it would be useful to have methods that do not necessitate the requirement for the protein to have an intrinsic binding function. There are several reasons why one might want selection to be independent of binding. Reliance on binding restricts the residues one can mutagenize since those residues directly involved in binding cannot be altered. If certain residues are absolutely required for folding, the experiments would be further limited. Finally, if protein function does not lend itself to binding an immobilized ligand or substrate, selection is not possible.

Two similar methods have been devised that uncouple the requirement for function from the stability of the protein. Both methods rely on the fact that unfolded or destabilized proteins are more sensitive to protease digestion than folded proteins. In the simpler of the two methods, a polyhistidine tag is placed at the amino terminus of the displayed protein.[86] All members of the library are first captured on a Ni^{2+} support. Once captured, phage are incubated in the presence of a protease and those that display an unfolded protein will be cleaved and released from the resin. Those that are protease resistant are retained and can be eluted with imidazole. Since phage are relatively inert to most proteases, there is little loss in viability. This method has been used effectively to select mutants of ubquitin with a repacked hydrophobic core.[87]

A second method, although very similar in principle, also takes advantage of the phage structure required for infectivity.[88,89] To utilize this modular system for phage infection, vectors were constructed so that DNA encoding a protein domain could be inserted between N2 and the carboxyl domain (CT) of gp3. If the inserted domain is unfolded, phage displaying the fusion will be sensitive to proteolytic cleavage, removing domains N1 and N2 and rendering the particle noninfectious. One of the nice features of this approach is that the phage are treated in solution, allowing the investigator to have better control over the conditions for proteolysis. The stringency of selection for folded domains can be varied by the time, temperature, and ionic conditions under which proteolysis is done. Two different vector systems have been developed for this type of display. In one case, the gene for the fusion protein is part of the phage genome so that five copies are ultimately displayed on the surface of the viral particle.[89] For this particle to become noninfectious, all five copies of the fusion protein would have to be proteolyzed. This might have the effect of allowing less stable domains to not be discriminated against during the selection process. An approach that is potentially more sensitive would be to use monovalent display so that on average only one copy of the fusion is displayed. Monovalent display in this case is not as straightforward as the usual protein display since the helper phage used in the rescue will supply wild-type gp3 which is protease insensitive. To circumvent this problem the helper phage was reengineered to introduce a protease cleavage site in the wild-type copy of gp3.[88] Both of these systems have been tested with proteins of know stability, Barnase, RNaseT1, and a destabilized RNaseT1, and have been shown to enrich for stable protease resistant forms of these proteins. One problem encountered appears to be the inherent instability of the insertion between N2 and CT, particularly in the case where this is incorporated directly in the phage genome.[89] Whether deletion of the insert provides some growth advantage to phage is unclear, but it is a serious problem since these phage will be protease resistant and selected. To avoid protease-insensitive deletions overtaking the population, selected variants are PCR amplified after each round of selection and recloned. This method has been used to select WW domains with higher binding affinities for peptides.[90] While intermediate concentrations of proteases led to the selection of binders with the highest affinity, when examined these WW domains lacked cooperative thermal unfolding. This suggests that in the absence of ligand these domains may adopt loosely folded states. Since the selection with protease was done in the presence of ligand, this would not discriminate against domains that were loosely folded in the absence of ligand.

B. Applications for Engineering Proteins

As phage-display technologies have evolved, the methods have become a valuable tool for protein engineering. Since most selection schemes rely on stably folded proteins, the technology can be readily applied when making even significant structural changes in the protein.

One area where this technology has proven useful is in the engineering of single-chain antibody molecules (scFv). ScFvs consist of variable heavy and light chain domains joined with a canonical linker sequence. This sequence may not necessarily represent the optimal sequence for stability. For example, a catalytic scFv prepared from a full-length antibody was inactive for binding to the target hapten when the variable domains were connected with a canonical linker.[91] To find the optimal linker a random linker library was constructed and selection for binding hapten was performed. Many of the variants showed significant increases in binding activity. While the sequences that were obtained were diverse, it was apparent that certain features were maintained. Similar experiments were also done using the SIP

method described earlier. In this case, the scFv with the canonical linker was stable but the investigators wanted to find a nonrepetitive replacement linker with comparable stability.[92]

Structural features other than the linker region of scFvs have been studied. In the framework of the immunoglobulin variable domain, a kink is found in the first strand of the β sheet that varies in sequence and in length depending on whether it is from a κ, λ or heavy chain variable domain. Two libraries were constructed to mimic the structures found in κ and λ chains, and the SIP method was used to select variants with increased thermostability.[93]

Phage selection has also been used to select for a scFv that is stable in the absence of disulfide bonds.[94] In contrast to previous studies where libraries were screened, mutations were introduced over the entire scFv since there was little data to suggest which mutations would most increase stability. A combination of DNA shuffling and stepwise selection allowed the isolation of mutants that were as thermostable as wild-type scFv containing the disulfide bond.

An elegant example of selection for stability is the minimization of the Z binding domain of protein A.[95] The Z domain is a three-helix bundle of 59 residues that binds the F_c portion of the IgG immunoglobulin with a K_d of 10 nM. All of the residues involved in binding are located in two of the helices. Although no residues in the third helix are directly involved in binding, removal of that helix destabilizes the structure and greatly reduces the binding affinity for F_c. To engineer a minimal Z domain, a two-helix derivative containing only 33 amino acids was displayed on phage and selected for binding to Fc. Overall seven different libraries were constructed to address different issues of binding and stability. Consensus residues from each library were then combined and reselected for the highest affinity binders. In this stepwise fashion, a minimal two-helix domain was constructed with an Fc binding affinity that was only a factor of 4 less than the starting wild-type Z domain. Two important themes are demonstrated in these experiments. First, instead of constructing enormous libraries where the entire protein is randomized, a solution was obtained using smaller focused libraries, thereby avoiding the necessity of preparing extremely large libraries. In addition, structural information can be used to minimize the number of libraries and mutants required to obtain a useful solution.

V. Phage Display and Catalytic Function

A. Selection Using Transition-State Analogues

One important goal of protein design is to improve or endow new catalytic activities on proteins using a variety of techniques including phage display. Because of the complex nature of catalysis, progress to date has been relatively modest. With the advent of phage display, some of the limitations imposed on exploring larger areas of sequence space could be overcome so that large libraries could be screened for new catalytic functions. The potential of phage display for selecting novel enzymes was demonstrated when a number of enzymes including bacterial alkaline phosphatase,[96] trypsin,[97] Staphylococcal nuclease (SNase),[98,99] and β-lactamase[100,101] were functionally displayed. While most mutational studies have centered on the enzyme active site, it is also possible to modify enzyme activity by changing a cofactor binding site. For example, phage display has been used to optimize the Zn binding site on carbonic anhydrase II.[102]

Much of the initial work with phage-displayed enzymes employed a strategy of selection based on binding to transition state analogues. To alter the specificity of human glutathione transferase (hGST), a phage-displayed library of hGST variants was constructed in which residues comprising the substrate-binding pocket were randomized.[103] The library was then used to select for binding to different transition-state analogues. After several rounds of selection, variants were identified that bound the transition-state analogue with higher affinity than the wild-type enzyme. Additional characterization of the mutants indicated that higher affinity binding to the transition-state analogue did not necessarily correlate with reaction rate enhancements. Likewise, using a transition-state analogue that more closely resembled the high-energy intermediate of the reaction was useful for selecting variants that bound with high affinity but did not improve catalysis.[104]

An alternative approach to reengineering natural enzymes is evolution of catalytic function in antibody domains. Initially this was done by immunizing mice with transition-state analogues and selecting for antibodies that bind the transition-state analogue. With the demonstration that antibody repertoires could be displayed on phage, the technology was applied to select catalytic antibodies. Antibody libraries were prepared from mice that had been immunized with a transition-state analogue. Antibodies with nanomolar affinities for the immunogen were selected from such libraries, but none were capable of substrate hydrolysis at a rate comparable to a natural enzyme.[105]

Another attempt to improve catalysis was to start with an antibody preselected for binding to a particular transition-state analogue and proceed to improve it by affinity maturation using phage display.[106] In this example, a humanized version of a murine-derived catalytic antibody that catalyzes the hydrolysis of amino acid phenyl esters was the starting point for three display libraries. Each library consisted of five or six randomized amino acid positions at or near the hapten binding site. The libraries were then biopanned to select for higher affinity binding to the hapten. Modest improvements in binding were identified, but still none showed increased catalytic activity. A similar approach was used for a murine antibody that catalyzes regiospecific deprotection of acylated carbohydrates.[107] On the basis of molecular modeling of the antigen combining site, six residues were selected for randomization. The resulting library was panned, and antibodies were selected that bound the transition-state analogue with higher affinity. One antibody showed a 12-fold enhancement of catalytic activity over the

parental antibody. From these studies it appears that while it is relatively easy to find high-affinity binders, it is very rare to select catalytic activity comparable to a natural enzyme. The modest gains in catalytic efficiency demonstrated to date are probably a reflection of the strategy of selection by binding transition-state analogues. To improve catalysis it will be necessary to devise schemes to select enzymes based on catalytic activity rather than binding. While this is more difficult to implement, it provides a more direct route for selecting catalysis.

B. Selection Using Suicide Inhibitors

One direct route for selection based on catalysis is the use of mechanism-based suicide inhibitors that result in covalent attachment of the inhibitor to the enzyme. To adapt this to phage selection, suicide inhibitors have been synthesized to include an affinity tag attached via a cleavable linker. Such bifunctional molecules have been synthesized using a β-lactamase inhibitor linked to biotin via a disulfide linkage. Using this substrate, phage displaying β-lactamase were enriched from nondisplaying phage by capture on streptavidin beads followed by release with reducing agent such as dithiothreitol.[100,101] This work demonstrated that phage displaying enzymatically active β-lactamase could be enriched over phage displaying inactive β-lactamase but did not address selection of variants with differing activities. More recently, mutations in the active site Ω-loop of β-lactamase have been studied.[108] A phage-displayed library in which the Ω-loop was randomized was used to select enzymes that were rapidly acylated using a biotinylated penicillin derivative. This work showed that the final result is dependent on both the concentration and the kinetics of the reaction. Suicide inhibitors have also been used to alter the substrate specificities of enzymes. Phage displaying subtilisin 309 (savinase) were selected using a series of biotinylated phosphonylating inhibitors that were then recovered on streptavidin beads.[109] Native savinase has limited reactivity when a charged residue is in the P4 position of the substrate. Using the appropriate inhibitor, mutants were isolated that had improved catalytic activity for substrates containing a basic residue in the P4 position.

Similarly, a number of phage-displayed catalytic antibodies have been selected based on reactivity with a substrate analogue that becomes covalently attached to the displayed enzyme. For example, a semi-synthetic antibody library (CDR3 of heavy and light chains randomized) was biopanned using a phenethylpyridyl disulfide. Phage were trapped by a disulfide exchange reaction that linked the substrate to a free cysteine in the CDR loop.[110] One selected antibody efficiently hydrolyzed the corresponding thioester. A more elaborate strategy was used to select for catalytic antibodies that hydrolyze glycosidic bonds.[111] Hybridoma lines were first generated following immunization of mice with a transition-state analogue. These lines were pooled, and a combinatorial phage-displayed antibody library was generated by PCR amplification of heavy and light chains. Phage were then selected using an immobilized hapten that became covalently attached after catalysis. Demonstration that the selected antibodies could indeed cleave glycosidic bonds was shown by the conversion of chromogenic substrates in cells expressing the antibody.

C. Selection Linked to Catalysis

All of the above-mentioned methods rely on having a mechanism-based suicide inhibitor for selection. Such reagents are often not available, so alternative methods are required. A number of strategies have been reported that do not rely on such inhibitors and may be more generally applied to a wider range of enzymes. These methods are designed to provide a substrate that upon conversion to product results in the selective release or capture of the phage as a result of catalysis. This technique was very effectively demonstrated with phage displaying subtiligase.[112] Subtiligase is a double mutant of subtilisin BPN′ that catalyzes the ligation of C-terminal activated peptides to the N-terminal α-amine of an acceptor protein or peptide. To be useful for selection it was necessary for subtliligase to catalyze a self-ligation. Thus, the amino terminus of subtiligase was first extended so that it could reach into the active site of the enzyme. A biotinylated peptide that served as substrate donor was then added to the modified subtiligase phage. Successful ligation to the amino terminus of subtiligase allowed catalytically active phage to be captured with streptavidin-coated beads. In this way an optimal sequence for the amino terminal extension of subtiligase was selected. Six different libraries were then generated in which a total of 25 residues comprising the active site of the enzyme were mutated. Mutants were isolated that showed increased catalytic activity, improved stability, and resistance to oxidation, suggesting that selection for catalytic activity encompasses a variety of parameters.

Two novel methods that link catalysis and selection without requiring direct self-labeling of the enzyme have been reported. In the first method the enzyme is displayed in typical monovalent fashion using a phagemid−helper phage system.[113] However, instead of the helper phage providing wild-type copies, gp3 was modified to display an acidic peptide containing a free cysteine (Figure 4A). This displayed peptide was used to bind and covalently attach a basic peptide that also contains a free cysteine and an affinity-tagged substrate. In the model system described, staphylococcal nuclease (SNase) is the displayed enzyme and the oligonucleotide substrate is tagged with biotin. Following rescue with the modified helper phage, the substrate-modified basic peptide was added and phage were captured on streptavidin coated beads. Catalysis was initiated by addition of Ca^{2+} with cleavage of the oligonucleotide releasing phage from the solid support. The number of phage released via catalysis, compared with control phage displaying no catalytic activity, was modest. This differential may be sufficient to select a rare phage displaying catalytic activity from a library of predominately catalytically inactive members. However, selection for catalytic improvement where all mem-

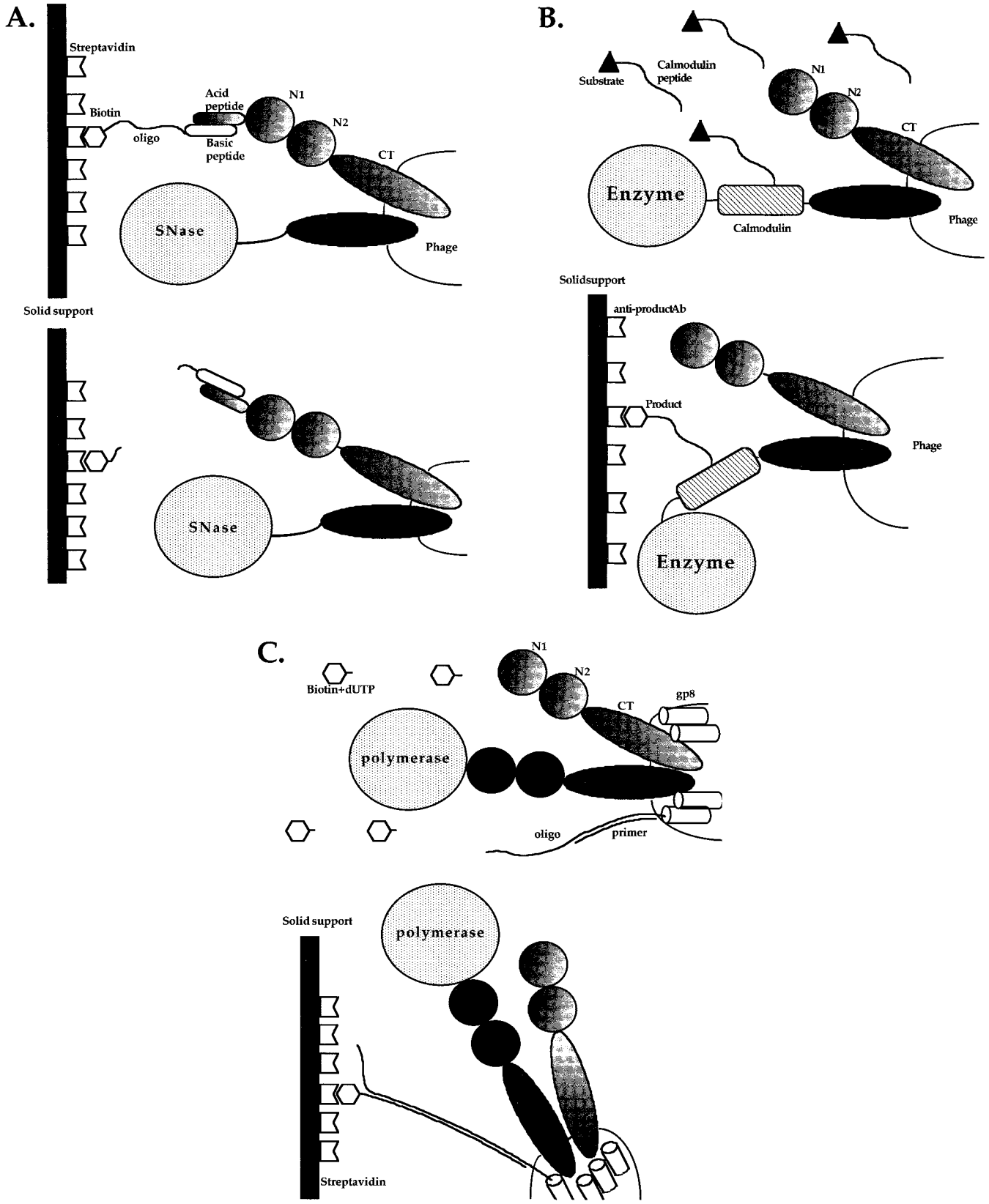

Figure 4. Methods for linking catalysis and selection. (A) A phagemid particle is generated that displays a copy of the enzyme (SNase) as a fusion to gp3. A modified helper virus is used for rescue that supplies copies of gp3 with a fusion of an acid peptide at its amino terminus. A basic peptide linked to a biotinylated oligonucleotide is added in vitro. The acid and basic peptide bind to each other allowing the phage to be captured on a solid support coated with streptavidin. Initiation of the enzymatic reaction by the addition of Ca^{2+} results in cleavage of the oligonucleotide by the enzyme and a release of the phage particle. (B) A phagemid particle is produced that has an enzyme and calmodulin fused to the amino terminus of gp3. A calmodulin peptide linked to substrate is added in vitro. Following conversion of substrate to product, phage are captured with an antibody specific for product. (C) Phagemid particles displaying a DNA polymerase have an oligonucleotide chemically cross-linked to gp8. Following annealing of a primer a biotinylated dUTP substrate is added to initiate the enzymatic reaction. Incorporation of the substrate allows capture of the phage on a solid support by streptavidin.

bers of the library retain some catalytic function has yet to be tested.

A conceptually similar method was developed using calmodulin and its high-affinity binding peptide[114] (Figure 4B). Rather than modifying the gp3 of the helper phage, this system makes a tripartite fusion of enzyme−calmodulin−gp3. The basic amino acid peptide described in the first method is replaced with the calmodulin peptide, which has 2 pM affinity for calmodulin. The enzyme substrate is chemically linked to the calmodulin peptide, which is then allowed to bind calmodulin. In this scheme the enzymatic reaction is allowed to proceed in solution; enzymatically active phage are captured using an antibody that recognizes the reaction product. A model system using displayed glutathione-*S*-transferase shows a moderate enrichment over catalytically inactive control phage. As with the previous system, actual experiments to select for catalytic improvement have not been reported.

If catalysis is allowed to take place in solution, a potential difficulty can occur should the enzyme acts in trans. Substrate molecules on neighboring phage may then be modified, and phenotype and genotype would no longer be linked i.e., enzymatic product would not necessarily be associated with active enzyme. This may not be a serious problem since substrate tethered to the phage with enzyme will be in high local concentration while the concentration of enzyme overall in solution (represented as phage particles) will be quite low. A variation of the methods described above may circumvent this problem (Figure 4C). In this method a phagemid system was constructed where DNA polymerase was displayed as a gp3 fusion and the DNA template was cross-linked to gp8.[115] To remove any trans product of catalysis, the helper phage gp3 was modified to insert a trypsin cleavage site between the N2 and CT domains. Since in any phagemid rescue experiment only 10% or less display the gp3 fusion, many phage in the population would have no polymerase but still would have substrate attached. By treating the phage lysate with trypsin prior to the polymerase reaction, any particle displaying only substrate would be rendered noninfective, thus eliminating phage acted upon by trans catalysis.

VI. Conclusions

Ten years have elapsed since the first demonstration that proteins could be displayed on the surface of phage.[27,116] During these years we have witnessed an explosive growth of this technology to the point where phage display has become an indispensable tool for protein engineering. Initially the technology was applied to improving the binding properties of various proteins to ligands but was quickly adapted to address issues in protein folding and catalysis. This has been made possible by numerous innovations. Undoubtedly there will be further improvements in library construction, vector design, and selection methodology that will further enhance phage display. Most importantly, phage display has helped make in vitro evolution of proteins a reality. In the future, protein design will be aided by the combined use of phage display and computational methods.

VII. Acknowledgments

The author thanks Dr. K. T. O'Neil for many helpful discussions and critical reading of the manuscript.

VIII. References

(1) Bryson, J. W.; Betz, S. F.; Lu, H. S.; Suich, D. J.; Zhou, H. X.; O'Neil, K. T.; DeGrado, W. F. *Science* **1995**, *270*, 935.
(2) Street, A. G.; Mayo, S. L. *Struct. Fold Des.* **1999**, *7*, R105.
(3) Smith, G. P. *Science* **1985**, *228*, 1315.
(4) Hanes, J.; Plückthun, A. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 4937.
(5) Roberts, R. W.; Szostak, J. W. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 12297.
(6) Webster, R. E. In *Phage Display of Peptides and Proteins*; Kay, B. K., Winter, J., McCafferty, J., Eds.; Academic Press: New York, 1996.
(7) Holliger, P.; Riechmann, L. *Structure* **1997**, *5*, 265.
(8) Lubkowski, J.; Hennecke, F.; Plückthun, A,; Wlodawer, A. *Nat. Struct. Biol.* **1998**, *5*, 140.
(9) Holliger, P.; Riechmann, L.; Williams, R. L. *J. Mol. Biol.* **1999**, *288*, 649.
(10) Gray, C. W.; Brown, R. S.; Marvin, D. A. *J. Mol. Biol.* **1981**, *146*, 621.
(11) Jacobson, A. *J. Virol.* **1972**, *10*, 835.
(12) Sun, T. P.; Webster, R. E. *J. Bacteriol.* **1986**, *165*, 107.
(13) Sun, T. P.; Webster, R. E. *J. Bacteriol.* **1987**, *169*, 2667.
(14) Reichmann, L.; Holliger. *Cell* **1997**, *90*, 351.
(15) Armstrong, J.; Perham, R. N.; Walker, J. E. *FEBS Lett.* **1981**, *135*, 167.
(16) Nelson, F. K.; Friedman, S. M.; Smith, G. P. *Virology* **1981**, *108*, 338.
(17) Crissman, J. W.; Smith, G. P. *Virology* **1984**, *132*, 445.
(18) Fuh, G.; Sidhu, S. S. *FEBS Lett* **2000**, *480*, 231.
(19) Glucksman, M. J.; Bhattacharjee, S.; Makowski, L. *J. Mol. Biol.* **1992**, *226*, 455.
(20) Kishchenko, G.; Batliwala, H.; Makowski, L. *J. Mol. Biol.* **1994**, *241*, 208.
(21) Greenwood, J.; Willis, A. E.; Perham, R. N. *J. Mol. Biol.* **1991**, *220*, 821.
(22) Felici, F.; Castagnoli, L.; Musacchio, A.; Jappelli, R.; Cesareni, G. *J. Mol. Biol.* **1991**, *222*, 301.
(23) Iannolo, G.; Minenkova, O.; Petruzzelli, R.; Cesareni, G, *J. Mol. Biol.* **1995**, *248*, 835.
(24) Petrenko, V. A.; Smith, G. P.; Gong, X.; Quinn, T. *Protein Eng.* **1996**, *9*, 797.
(25) Jespers, L. S.; Messens, J. H.; De Keyser, A.; Eeckhout, D.; Van Den Brande, I.; Gansemans, Y. G.; Lauwereys, M. J.; Vlasuk, G. P.; Stanssens, P. E. *Bio/Technology* **1995**, *13*, 378.
(26) Gao, C.; Mao, S.; Lo, C. H.; Wirsching, P.; Lerner, R. A.; Janda, K. D. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 6025.
(27) Bass, S.; Greene, R.; Wells, J. A. *Proteins* **1990**, *8*, 309.
(28) Sidhu, S. S.; Weiss, G. A.; Wells, J. A. *J. Mol. Biol.* **2000**, *296*, 487.
(29) Weiss, G. A.; Wells, J. A.; Sidhu, S. S. *Protein Sci.* **2000**, *9*, 647.
(30) Dueñas, M,; Borrebaeck, C. A. K. *Bio/Technology* **1994**, *12*, 999.
(31) Gramatikoff, K.; Georgiev, O.; Schaffner, W. *Nucleic Acid Res.* **1994**, *22*, 5761.
(32) Krebber, C.; Spada, S.; Desplancq, D.; Plückthun, A. *FEBS Lett.* **1995**, *377*, 227.
(33) Gao, C.; Lin, C.-H.; Lo, C.-H. L.; Mao, S.; Wirsching, P.; Lerner, R. A.; Janda, K. D. *Proc. Natl. Acad. Sci. U.S.A.* **1997**, *94*, 11777.
(34) Stemmer, W. P. *Nature* **1994**, *370*, 389.
(35) Knappik, A.; Ge, L.; Honegger, A.; Pack, P.; Fischer, M.; Wellnhofer, G.; Hoess, A.; Wölle, J.; Plückthun, A.; Virnekas, B. *J. Mol. Biol.* **2000**, *296*, 57.
(36) Parmley, S. F.; Smith, G. P. *Gene* **1988**, *73*, 305.
(37) Harrison, J. L.; Williams, S. C.; Winter, G.; Nissim, A. *Methods Enzymol.* **1996**, *267*, 83.
(38) Levitan, B. *J. Mol. Biol.* **1998**, *277*, 893.
(39) Lowman, H. B.; Bass, S. H.; Simpson, N.; Wells, J. A. *Biochemistry* **1991**, *30*, 10832.
(40) Lowman, H. B.; Wells, J. A. *J. Mol. Biol.* **1993**, *234*, 564.
(41) Winter, G.; Griffiths, A. D.; Hawkins, R. E.; Hoogenboom, H. R. *Annu. Rev. Immunol.* **1994**, *12*, 433.
(42) Hoogenboom, H. R.; de Bruine, A. P.; Hufton, S. E.; Hoet, R. M.; Arends, J. W.; Roovers, R. C. *Immunotechnology* **1998**, *4*, 1.
(43) Low, N. M.; Holliger, P.; Winter, G. *J. Mol. Biol.* **1996**, *260*, 359.
(44) Short, M. K.; Jeffrey, P. D.; Kwong, R. F.; Margolies, M. N. *J. Biol. Chem.* **1995**, *270*, 28541.

(45) Thompson, J.; Pope, T.; Tung, J. S.; Chan, C.; Hollis, G.; Mark, G.; Johnson, K. S. *J. Mol. Biol.* **1996**, *256*, 77.
(46) Yang, W. P.; Green, K.; Pinzsweeney, S.; Briones, A. T.; Burton, D. R.; Barbas, C. F., III *J. Mol. Biol.* **1995**, *254*, 392.
(47) Chowdhury, P. S.; Pastan, I. *Nat. Biotechnol.* **1999**, *17*, 568.
(48) Cabibbo, A.; Sporeno, E.; Toniatti, C.; Altamura, S.; Savino, R.; Paonessa, G.; Ciliberto, G. *Gene* **1995**, *167*, 41.
(49) Hiipakka, M.; Poikonen, K.; Saksela, K. *J. Mol. Biol.* **1999**, *293*, 1097.
(50) Huang, W.; Zhang, Z.; Palzkill, T. *J. Biol. Chem.* **2000**, *275*, 14964.
(51) Rebar, E. J.; Pabo, C. O. *Science* **1994**, *263*, 671.
(52) Jamieson, A. C.; Kim, S.-H.; Wells, J. A. *Biochemistry* **1994**, *33*, 5689.
(53) Choo, Y.; Klug, A. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 11163.
(54) Choo, Y.; Klug, A. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 11168.
(55) Wu, H.; Yang, W.-P.; Barbas, C. F., III *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 344.
(56) Greisman, H. A.; Pabo, C. O. *Science* **1997**, *275*, 657.
(57) Wolfe, S. A.; Greisman, H. A.; Ramm, E. I.; Pabo, C. O. *J. Mol. Biol.* **1999**, *285*, 1917.
(58) Isalan, M.; Klug, A.; Choo, Y. *Biochemistry* **1998**, *37*, 12026.
(59) Segal, D. J.; Dreier, B.; Beerli, R. R.; Barbas, C. F., III *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 2758.
(60) Dreier, B.; Segal, D. J.; Barbas, C. F., III *J. Mol. Biol.* **2000**, *303*, 489.
(61) Isalan, M.; Choo, Y. *J. Mol. Biol.* **2000**, *295*, 471.
(62) Nilsson, M. T. I.; Mossing, M. C.; Widersten, M. *Protein Eng*, **2000**, *13*, 519.
(63) Ku, J.; Schultz, P. G. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 6552.
(64) McConnell, S. J.; Hoess, R. H. *J. Mol. Biol.* **1995**, *250*, 460.
(65) Bianchi, E.; Folgori, A.; Wallace, A.; Nicotra, M.; Acali,; Phalipon, A.; Barbato, G.; Bazzo, R.; Cortese, R.; Felici, F.; Pessi, A. *J. Mol. Biol.* **1995**, *247*, 154.
(66) Nord, K.; Nilsson, J.; Nilsson, B.; Uhlen, M.; Nygren, P.-A. *Protein Eng.* **1995**, *8*, 601.
(67) Dennis, M. S.; Lazarus, R. A. *J. Biol. Chem.* **1994**, *269*, 22129.
(68) Roberts, B. L.; Markland, W.; Ley, A. C.; Kent, R. B.; White, D. W.; Guterman, S. K.; Ladner, R. C. *Proc. Natl. Acad. Sci. U.S.A.* **1992**, *89*, 2429.
(69) Koide, A.; Bailey, C. W.; Huang, X.; Koide, S. *J. Mol. Biol.* **1998**, *284*, 1141.
(70) Smith, G. P.; Patel, S. U.; Windass, J. D.; Thornton, J. M.; Winter, G.; Griffiths, A. D. *J. Mol Biol.* **1998**, *277*, 317.
(71) Hufton, S. E.; van Neer, N.; van den Beuken, T.; Desmet, J.; Sablon, E.; Hoogenboom, H. R. *FEBS Lett.* **2000**, *475*, 225.
(72) Venturini, S.; Martin, F. *Protein Pept. Lett.* **1994**, *1*, 70.
(73) Martin, F.; Toniatti, C.; Salvati, A. L,; Venturini, S.; Ciliberto, G.; Cortese, R.; Sollazzo, M. *EMBO J.* **1994**, *13*, 5303.
(74) Martin, F.; Toniatti, C.; Salvati, A. L,; Ciliberto, G.; Cortese, R.; Sollazzo, M. *J. Mol. Biol.* **1996**, *255*, 86.
(75) Beste, G.; Schmidt, F. S.; Stibora, T.; Skerra, A. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 1898.
(76) Schlehuber, S.; Beste, G.; Skerra, A. *J. Mol. Biol.* **2000**, *297*, 1105.
(77) O'Neil, K. T.; Hoess, R. H.; Raleigh, D. P.; DeGrado, W. F. *Proteins* **1995**, *21*, 11.
(78) Gu, H.; Yi, Q.; Bray, S. T.; Riddle, D. S.; Shiau, A. K.; Baker, D. *Protein Sci.* **1995**, *4*, 1108.
(79) Kim, D. E.; Gu, H.; Baker, D. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 4982.
(80) Zhou, H. X.; Hoess, R. H.; DeGrado, W. F. *Nat. Struct. Biol.* **1996**, *3*, 446.
(81) Gu, H.; Kim, D.; Baker, D. *J. Mol. Biol.* **1997**, *274*, 588.
(82) Ruan, B.; Hoskins, J.; Wang, L.; Bryan, P. N. *Protein Sci.* **1998**, *7*, 2345.
(83) Kamtekar, S.; Schiffer, J. M.; Xiong, H.; Babik, J. M.; Hecht, M. *Science* **1995**, *262*, 1680.
(84) Riddle, D. S.; Santiago, J. V.; Bray-Hill, S. T.; Doshi, N.; Grantcharova, V. P.; Yi, Q.; Baker, D. *Nat. Struct. Biol.* **1997**, *4*, 805.
(85) Jung, S.; Honegger, A.; Plückthun, A. *J. Mol. Biol.* **1999**, *294*, 163.
(86) Finucane, M. D.; Tuna, M.; Lees, J. H.; Woolfson, D. N. *Biochemistry* **1999**, *38*, 11604.
(87) Finucane, M. D.; Woolfson, D. N. *Biochemistry* **1999**, *38*, 11613.
(88) Kristensen, P.; Winter, G. *Folding Des.* **1998**, *3*, 321.
(89) Sieber, V.; Pluckthun, A.; Schmid, F. X. *Nat. Biotechnol.* **1998**, *16*, 955.
(90) Dalby, P. A.; Hoess, R. H.; DeGrado, W. F. *Protein Sci.* **2000**, *9*, 1.
(91) Tang, Y.; Jiang, N.; Parakh, C.; Hilvert, D. *J. Biol. Chem.* **1996**, *271*, 15682.
(92) Hennecke, F.; Krebber, C.; Plückthun, A. *Protein Eng.* **1998**, *11*, 405.
(93) Spada, S.; Honegger, A.; Plückthun, A. *J. Mol. Biol.* **1998**, *283*, 395.
(94) Proba, K.; Wörn, A.; Honegger, A.; Plückthun, A. *J. Mol. Biol.* **1998**, *275*, 245.
(95) Braisted, A. C.; Wells, J. A. *Proc. Natl. Acad. Sci. U.S.A.* **1996**, *93*, 5688.
(96) McCafferty, J.; Jackson, R. H.; Chiswell, D. J. *Protein Eng.* **1991**, *4*, 955.
(97) Corey, D. R.; Shiau, A. K.; Yang, Q.; Janowski, B. A.; Craik, C. S. *Gene* **1993**, *128*, 129.
(98) Ku, J.; Schultz, P. G. *Bioorg. Med. Chem.* **1994**, *2*, 1413.
(99) Light, J.; Lerner, R. A. *Bioorg. Med. Chem.* **1995**, *3*, 955.
(100) Soumillion, P.; Jespers, L.; Bouchet, M.; Marchand-Brynaert, J.; Winter, G.; Fastrez, J. *J. Mol. Biol.* **1994**, *237*, 415.
(101) Soumillion, P.; Jespers, L.; Bouchet, M.; Marchand-Brynaert, J.; Sartiaux, P.; Fastrez, J. *Appl. Biochem. Biotechnol.* **1994**, *47*, 175.
(102) Hunt, J. A.; Fierke, C. A. *J. Biol. Chem.* **1997**, *272*, 20364.
(103) Widersten, M.; Mannervik, B. *J. Mol. Biol.* **1995**, *250*, 115.
(104) Hansson, L.; Widersten, M.: Mannervik, B. *Biochemistry* **1997**, *36*, 11252.
(105) McCafferty, J.; Fitzgerald, K. J.; Earnshaw, J.; Chiswell, D. J.; Link, J.; Smith, R.; Kenten, *J. Appl. Biochem. Biotechnol.* **1994**, *47*, 157.
(106) Baca, M.; Scanlan, T. S.; Stephenson, R. C.; Wells, J. A. *Proc. Natl. Acad. Sci. U.S.A.* **1997**,*94*, 10063.
(107) Fujii, I.; Fukuyama, S.; Iwabuchi, Y.; Tanimura, R. *Nat. Biotechnol.* **1998**, *16*, 463.
(108) Vanwetswinkel, S.; Avalle, B.; Fastrez, J. *J. Mol. Biol.* **2000**, *295*, 527.
(109) Legendre, D.; Laraki, N.; Gräslund, T.; Bjornvad, M. E.; Bouchet, M.; Nygren, P.-A.; Borchert, T. V.; Fastrez, J. *J. Mol. Biol.* **2000**, *296*, 87.
(110) Janda, K. D.; Lo, C.-H. L.; Li, T.; Barbas, C. F., III; Wirsching, P.; Lerner, R. A. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 2532.
(111) Janda, K. D.; Lo, L.-C.; Lo, C.-H. L.; Sim, M.-M.; Wang, R.; Wong, C.-H.; Lerner, R. A. *Science* **1997**, *275*, 945.
(112) Atwell, S.; Wells, J. A. *Proc. Natl. Acad. Sci. U.S.A.* **1999**, *96*, 9497.
(113) Pedersen, H.; Hölder, S.; Sutherlin, D. P.; Schwitter, U.; King, D. S.; Schultz, P. G. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 10523.
(114) Demartis, S.; Huber, A.; Viti, F.; Lozzi, L.; Giovannoni, L.; Neri, P.; Winter, G.; Neri, D. *J. Mol. Biol.* **1999** *286*, 617.
(115) Jestin, J.-L.; Kristensen, P.; Winter, G. *Angew. Chem., Int. Ed. Engl.* **1999**, *38*, 1124.
(116) McCafferty, J.; Griffiths, A. D.; Winter, G.; Chiswell, D. J. *Nature* **1990**, *348*, 552.

CR000056B