JMB

Design and Evolution of Artificial M13 Coat Proteins

Gregory A. Weiss and Sachdev S. Sidhu*

Department of Protein Engineering, Genentech Inc. 1 DNA Way, South San Francisco, CA 94080, USA

Using simple design and selective pressure, we have evolved an artificial M13 bacteriophage coat protein. M13 coat proteins first reside in the bacterial inner membrane and subsequently surround the DNA core of the assembled virus. The artificial coat protein (ACP) was designed and evolved to mimic both functions of the natural M13 coat proteins, but with an inverted orientation. ACP is a non-functional coat protein because it is not required for the production of phage particles. Instead, it incorporates into a phage coat which still requires all the natural coat proteins for structural integrity. In contrast with other M13 coat proteins, which can display polypeptides as amino-terminal fusions, ACP permits the carboxy-terminal display of large polypeptides. The results suggest that viruses can co-opt host membrane proteins to acquire new coat proteins and thus new functions. In particular, M13 bacteriophage can be engineered for new functions, such as carboxy-terminal phage display.

© 2000 Academic Press

Keywords: protein engineering; viral evolution; phage display; protein libraries; M13 bacteriophage

*Corresponding author

Introduction

Viruses consist of genetic material encapsulated within a coat at least partially composed of protein. Coat proteins protect the genetic material and also mediate host cell recognition and infection. It has been proposed that coat proteins descend from host membrane proteins (Dornburg, 1997). For example, an intimate relationship between host and viral proteins has been implicated in the pathogenicity of *Vibrio cholerae*, which depends on a single protein functioning as both a bacterial membrane protein and a bacteriophage coat protein (Karaolis *et al.*, 1999). The promiscuous incorporation of non-wild-type coat proteins has also been exploited for the purposes of phage display (Smith, 1985). The important consequences of this fascinating natural process led us to develop an

M13 bacteriophage model system to test requirements for protein incorporation into the viral coat.

M13 is a filamentous bacteriophage: a worm-like virus approximately 1 μm long with a 10 nm diameter that infects only *Escherichia coli*. The viral particle consists of a single-stranded, closed circular DNA core surrounded by a protein coat (Marvin, 1998 and references therein). Prior to virus assembly, the coat proteins are fixed in the bacterial membrane by transmembrane domains. During assembly, viral DNA is extruded through the membrane and concomitantly enveloped by coat proteins (Webster, 1996) (Figure 1). The ends of the assembled virus are capped by four minor coat proteins (approximately five copies each per virion), and the length of the filament is covered by several thousand copies of the major coat protein (protein-8, P8).

It has been shown that both P8 and the gene-3 minor coat protein (protein-3, P3) can accommodate polypeptide fusions at their amino termini. The fused polypeptide is displayed on the surface of a phage particle that also encapsulates the cognate gene. This physical linkage between phenotype and genotype has been exploited in the selection technology known as phage display (reviewed by Clackson & Wells, 1994; Smith & Petrenko, 1997). In this powerful approach to protein engineering, ligands with desired binding properties can be isolated from large libraries of

---

Abbreviations used: ACP, artificial coat protein; BSA, bovine serum albumin; ELISA, enzyme linked immunosorbant assay; hGH, human growth hormone; hGHbp, human growth hormone binding protein; IPTG, isopropylthio-β-D-galactoside; ORF, open reading frame; P3, protein-3, the gene-3 minor coat protein of M13 phage; P8, protein-8, the major coat protein of M13 phage. Mutations are designated by the single-letter amino acid code for the wild-type residue followed by its sequence position and the mutant residue.

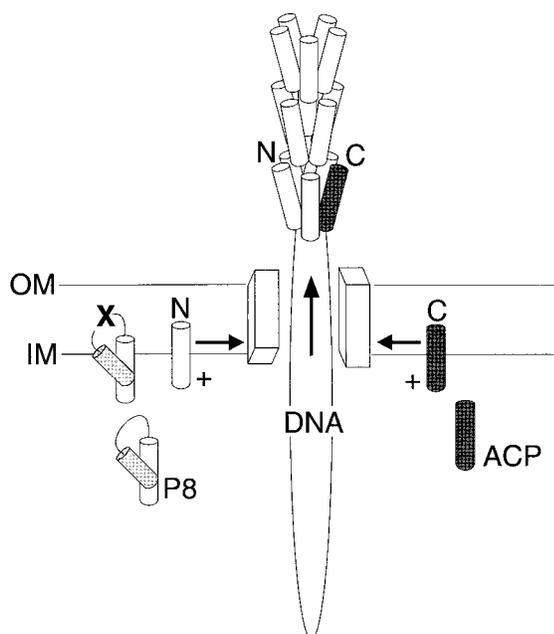E-mail address of the corresponding author: sidhu@gene.com

Figure 1. Assembly of M13 virus (Webster, 1996) and the evolution of an artificial coat protein (ACP). P8 molecules (white cylinders) are synthesized with a proteolytically removed (X) pre-peptide (light gray), necessary for insertion into the inner membrane (IM); thereafter, the P8 amino terminus (N) remains exposed to the periplasm and the positively charged carboxy terminus remains in the cytoplasm. Viral DNA is extruded through a virus-supplied pore (boxes), which bridges the inner membrane and the outer membrane (OM), and several thousand P8 molecules concomitantly encapsulate the DNA. The library of potential artificial coat proteins (dark gray) did not include a pre-peptide, but instead featured a randomized hydrophobic-biased core flanked by the reversed amino acid sequences of the carboxy and amino termini of P8. The resulting sequences resembled bacterial secretion signals which consist of a positively charged amino terminus followed by a hydrophobic domain. Insertion of such sequences into the inner memberane, in a manner analogous to secretion signal insertion, would result in a cytoplasmic amino terminus and a periplasmic carboxy terminus (C). Library members capable of incorporating into the assembling phage particle would then resemble P8 in the reverse orientation, with a surface-exposed carboxy terminus available for fusion protein display.

phage-displayed proteins simply by binding to an immobilized receptor *in vitro*. The sequences of the selected polypeptides can be rapidly determined by sequencing the encapsulated DNA.

Early examples of phage display used single-gene systems in which libraries were fused to the amino terminus of either P3 (Scott & Smith, 1990) or P8 (Greenwood *et al.*, 1991) encoded by the viral genome. The utility of these systems was limited by the fact that fusions resulting in compromised coat protein function could not be efficiently displayed. In particular, the display of large polypep-

tides and proteins was severely limited (Iannolo *et al.*, 1995). This problem was circumvented by the use of two-gene phagemid systems which enabled the packaging of fusion gene products in a phage coat predominantly composed of wild-type proteins supplied *in trans* from a helper phage (Bass *et al.*, 1990). In such systems, the helper phage supplies all the proteins necessary for viral assembly, and thus, the fusion gene product is a non-functional coat protein that incorporates into the phage particle as a non-essential addition to the wild-type coat.

Recently, we have demonstrated that, in a two-gene P8 display system, the P8 gene fused to the displayed protein can accommodate a suprisingly large number of mutations. Furthermore, many of these mutations actually improve packaging, and hence protein display, in a predominantly wild-type coat (Sidhu, *et al.*, 2000b). As an extreme example of sequence divergence, we were able to construct a functional P8 variant with only 50 % identity with the wild-type. Mutational analysis of these P8 variants revealed that only a small subset of the amino-terminal 30 residues are critical for coat incorporation (Weiss, *et al.*, 2000). The fusion-P8 can mutate freely without compromising phage viability, because wild-type P8 supplied by helper phage maintains the integrity of the phage coat. The fusion-P8 variants serve no role in the phage life cycle, yet they can be maintained as non-essential components of the phage particle. Furthermore, they are free to evolve new functions that may prove advantageous under certain conditions (e.g. selective pressure for increased display of heterologous fusions).

These results suggest that the components of a viral phage coat need not be restricted to the gene products of the viral genome. Thus, our findings support the hypothesis that, for viruses that assemble at the host membrane, membrane proteins capable of stable association with the virus particle could evolve to form new coat proteins (Dornburg, 1997). New coat proteins could arise from host membrane proteins or even from completely synthetic sequences. We now report the ultimate test of this hypothesis: the design and evolution of an artificial coat protein that is packaged into the M13 bacteriophage coat during viral assembly.

Results and Discussion

A library of potential coat proteins

We constructed a library of potential coat proteins. The library design was based on the sequence of P8, because P8 forms the vast majority of the phage coat. P8 is synthesized with a 23-residue pre-peptide, which is necessary for insertion into the host cell inner membrane. After cleavage of the prepeptide, mature P8 is a 50-residue α-helix consisting of 20 amino-terminal residues which reside in the periplasm, followed by a 19-residue hydrophobic, transmembrane domain, and an 11-

residue carboxy-terminal cytoplasmic domain (Webster, 1996). Integration of P8 into the viral coat requires interactions between the positively charged cytoplasmic domain and the negatively charged DNA, and also, interactions between the central hydrophobic domains of adjacent P8 molecules. These interactions also maintain the integrity of the assembled virus (Marvin, 1998). P8 carboxy termini are buried close to the DNA core, while the central hydrophobic domains of adjacent P8 molecules pack together to provide structural support. P8 amino termini are exposed on the surface of the phage particle (Marvin *et al.*, 1994).

Interestingly, the basic structure of P8, a positively charged cytoplasmic domain and a hydrophobic transmembrane domain, is found in many membrane proteins (Gennis, 1989). Furthermore, most secreted proteins are transiently anchored to the membrane by similarly structured peptides (secretion peptides) which direct the secretion process (von Heijne, 1986). We reasoned that these structural similarities may predispose such proteins for incorporation into the viral coat, and that the main determinant for incorporation may be the transmembrane domain. Mutations in the hydrophobic transmembrane domain that introduce favorable interactions with the coat would then provide a mechanism for the evolution of new coat proteins. To test this hypothesis, we constructed a library of genes (>10^{10} unique members) encoding potential membrane proteins with randomized hydrophobic domains. The library of genes was placed in a phagemid vector that could be packaged into phage particles upon co-infection with a helper phage. From this library, we selected members capable of functioning as additional M13 coat proteins.

The library of potential membrane proteins borrowed from the reversed sequence of mature P8. By reversed sequence, we mean the amino acid sequence obtained by reading a protein sequence from the carboxy terminus to the amino terminus, such that the first residue becomes the last residue and the last residue becomes the first. Reversing the sequence of the P8 carboxy-terminal domain resulted in a positively charged amino-terminal domain. This was followed by a variable stretch consisting of 19 randomized residues biased towards hydrophobic side-chains. The structure terminated with the reversed sequence of the P8 amino-terminal domain. We do not claim that the reversed sequence of P8 provided any inherent interactions with the phage coat; the approach was simply used to generate a library with members resembling bacterial secretion peptides: a positively charged amino-terminal region followed by a hydrophobic stretch (Gennis, 1989). Thus, we reasoned that some library members would function as secretion signals and insert into the bacterial membrane, with amino termini in the cytoplasm and carboxy termini in the periplasm. We further reasoned that amongst these, a few

sequences might incorporate into assembling virus particles by virtue of favorable interactions with the viral coat and the phage DNA.

Coat protein selection for peptide display

To isolate rare individuals capable of functioning as reverse-oriented coat proteins, we applied selective pressure based on the concept of phage display. A short peptide epitope (polyHis tag) was fused to the carboxy termini of the potential membrane proteins, phage particles encapsulating phagemid DNA were produced by co-infection with helper phage, and phage displaying the polyHis tag were isolated from the library by binding to an anti-polyHis antibody.

Six artificial coat proteins (ACPs) were found that displayed the polyHis tag at levels detectable above background (Figure 2). The library design fixed the amino-terminal and carboxy-terminal domains, but diverse central sequences were found. The lack of homology between selected sequences was expected, as the potential diversity of the randomized region ($\approx 10^{17}$) was sparsely sampled by the library diversity (>10^{10}).
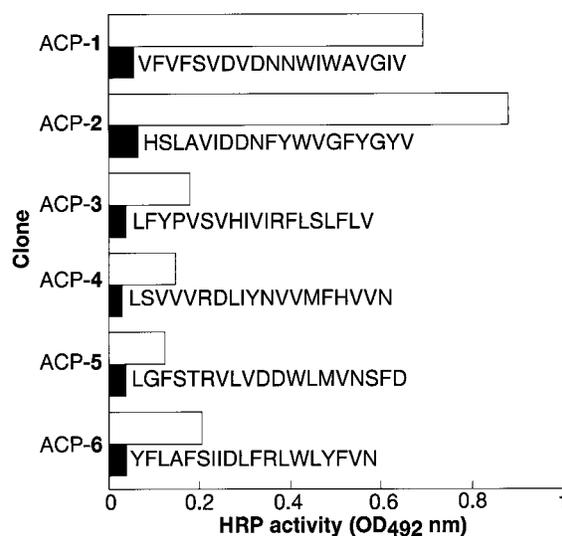

Figure 2. PolyHis tag display by artificial coat proteins (ACPs) selected from the first generation library. In these phage ELISAs (Sidhu *et al.*, 2000a), virus particles in solution (10^{13} particles/ml) were captured with immobilized binding target and quantified using a colorimetric reaction. Specific binding was measured by capture with an anti-His$_4$ antibody (white) while non-specific background was measured by capture with BSA (black). The intensity of the absorbance at 492 nm is proportional to the number of particles bound. The protein sequences were as follows: MSKSTFKKFLK-(X)$_{19}$-ETASAQLSNFAAKAPDDGEAAAHHHHHHA, where (X)$_{19}$ indicates the 19-residue sequence unique to each selectant (shown).

Coat protein selection for large protein display

To obtain a protein that incorporated more efficiently into the phage coat, we evolved one of these selectants (ACP-1) for large protein display. Polypeptide display by fusion to P8 is extremely sensitive to fusion protein length, with large proteins displayed at levels far below those observed for small peptides (Malik *et al.*, 1996). Similarly, while ACP-1 displays the small polyHis tag (Figure 2), a large 20 kDa protein (human growth hormone, hGH) is not displayed at detectable levels (Figure 3(a)). We constructed several libraries varying discrete regions of ACP-1 and fused their carboxy termini to an hGH variant (hGHsm) (Lowman & Wells, 1993) with a very high level of affinity for the extracellular domain of the hGH receptor (hGHbp) (Fuh *et al.*, 1993). The libraries were pooled together and, from this large population ($\approx 10^{11}$ unique members), we isolated phage displaying hGHsm by selecting for phage binding to immobilized hGHbp. A selected coat protein (Figure 3(b)) capable of hGHsm display (ACP-7) was fused to wild-type hGH. In contrast with ACP-1, the ACP-7 construct displayed carboxy-terminal-fused hGH, albeit at levels 30-fold lower than those achieved with the standard amino-terminal fusion to P8 (Figure 3(a)). Since hGH fused to P8 is displayed at monovalent levels (Sidhu *et al.*, 2000b), the incorporation of the hGH-ACP-7 fusion into the phage coat is still rather inefficient, with an incorporation rate less than one fusion protein per phage particle. However, the display levels were sufficient to enable selection

from a library of non-displaying phage, and further improvements may still be possible with further rounds of mutagenesis and selection.

ACP-7 is composed of only 50 amino acids, yet it supports the display of large carboxy-terminal fusions. ACP-7 is even simpler than the natural major coat protein, because whereas P8 requires a prepeptide for function, ACP-7 incorporates into the viral coat without the aid of any auxiliary sequence. The function of ACP-7 can be explained by considering its structure in relation to the requirements of the secretion pathway (Boyd & Beckwith, 1990). ACP-7 was selected from a library spanning the ten residues immediately following the selected sequence of ACP-1. The junction of the two selected sequences creates an uninterrupted 13-residue hydrophobic stretch, which could span the *Escherichia coli* inner membrane (Figure 3(b)). It should be noted that in the second generation libraries, codons were randomized to encode all 20 natural amino acid residues, and thus, the hydrophobic domain arose from an unbiased naïve library. Furthermore, neither ACP-1 nor any of the other first generation selectants contains a hydrophobic stretch of this length. Thus, the improved function of ACP-7 relative to the first generation selectants is likely due to this hydrophobic domain. This feature, together with the positively charged amino-terminal domain, constitutes a potential bacterial secretion signal. If ACP-7 passed through the secretion pathway, it would resemble the membrane-associated form of P8 in reverse orientation. The central hydrophobic stretch would be
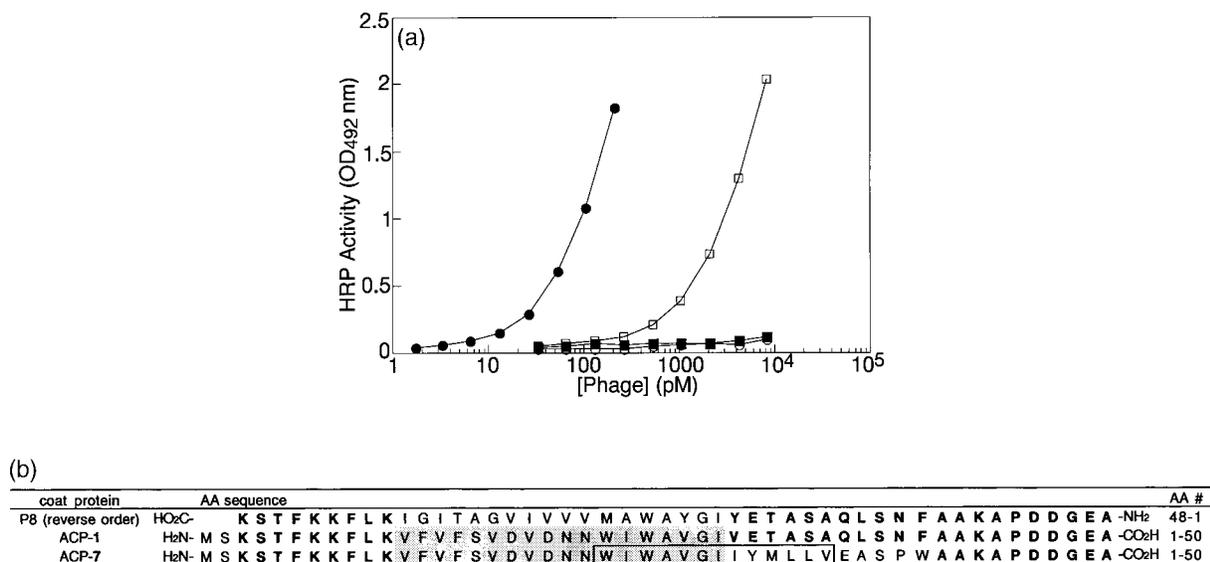

Figure 3. Protein display by ACP-7, as measured by phage ELISA. (a) Levels of protein display by hGH fusion to the carboxy terminus of ACP-**7** (□) compare favorably to display by hGH fusion to the amino terminus of P8 (●). Also shown, carboxy-terminal hGH fusion to ACP-**1** (○) and, as a negative control, P8 without the hGH fusion (■). (b) Selectants from the first generation (ACP-1) and second generation (ACP-7) library share homology (bold text) with the reversed amino acid (AA) sequence of P8, as designed. The second generation library retained some ACP-1 sequence (gray). Following second generation selection, a putative transmembrane domain (box) was obtained in ACP-7.

imbedded in the inner membrane with the amino terminus in the cytoplasm and the carboxy terminus in the periplasm (Figure 1). Incorporation of ACP-7 into the viral coat would then result in a final orientation opposite to that of P8, with the carboxy terminus exposed on the surface of the phage particle.

Similarities between ACP-7 and natural bacteriophage proteins

The design and evolution of ACP-7 depended on the use of a phagemid system in which all the wild-type phage coat proteins were supplied by a helper phage. As a consequence, non-functional coat proteins could be selected, provided they incorporated into the wild-type coat without significantly impairing the assembly process. While ACP-7 was evolved from completely synthetic sequences, there are natural precedences for several aspects of the design strategy. M13 protein-1 is an inner membrane protein that interacts with both P8 and the viral DNA during phage assembly. While protein-1 is not present in the assembled phage particle, it resembles ACP-7 because it contains a cytoplasmic amino terminus with a positively-charged pattern similar to the carboxy terminus of P8 (Rapoza & Webster, 1995). The bacteriophage Pf3 major coat protein resembles M13 P8 both in its membrane-associated form and also as a component of the phage coat. However, the Pf3 coat protein is similar to ACP-7 in that it inserts into the bacterial membrane without the aid of a pre-peptide (unlike P8 which requires a pre-peptide). Furthermore, the membrane orientation of the Pf3 coat protein can be reversed simply by reversing the charges on the intracellular and extracellular domains (Keifer *et al.*, 1997). The resulting inverted protein is strikingly similar to ACP-7: it has a positively-charged, cytoplasmic amino terminus and a negatively-charged, periplasmic carboxy terminus. In light of our results, it may be worth investigating whether M13 protein-1 or the reversed Pf3 coat protein could be evolved to incorporate into the phage coat.

Implications of artificial coat protein evolution

In summary, we have evolved new proteins that can be incorporated into the M13 phage particle. The evolved ACP-7 lacks significant homology with any known protein (based on a search of Genbank), yet it incorporates into the M13 coat. The selected region of ACP-7 is only 29 residues in length, and it includes a significant hydrophobic stretch that could act as both a transmembrane domain and guide incorporation into the viral coat. Our results suggest that the requirements for incorporation into the viral coat may be quite lax. Host membrane proteins could evolve into viral coat proteins by virtue of mutated transmembrane domains that interact favorably with the assembling coat. Such a process could enable viruses to acquire new functions.

Our results have significant implications for protein engineering in general and phage display in particular. In a phagemid system, where a helper phage provides all the wild-type coat proteins, additional proteins can be incorporated into the coat for the purposes of phage display (Bass *et al.*, 1990). For the first time, we demonstrate that these additional proteins need not be restricted to variants of the natural coat proteins. Instead, artificial coat proteins can be selected and evolved for applications that are unsuitable for the natural coat proteins. Our evolved carboxy-terminal display platform should prove useful for studying the binding specificities of proteins that recognize ligands with free carboxy termini, including members of the highly diverse PDZ domain family which mediate numerous signaling pathways (Fanning & Anderson, 1999). Also, carboxy-terminal phage display formats are essential for the functional display of cDNA libraries in which the presence of natural stop codons precludes the use of amino-terminal fusions (Crameri & Suter, 1993).

Materials and Methods

Selection of viral coat proteins for polyHis tag display

A phagemid pS1207a was constructed using standard molecular biology techniques. pS1207a is identical to a previously described IPTG-inducible phagemid (Sidhu *et al.*, 2000a), except that the open reading frame (ORF) under the control of the IPTG-inducible P_{tac} promoter has been deleted and replaced by a new ORF, encoding the following polypeptide: MSKSTFKKFLKETA-SAQLSNFAAKAPDDGEAAAHHHHHHA. This ORF was designed as follows. The first two residues were chosen to allow good initiation of translation. This dipeptide was followed by the reverse sequence of residues 40-48 of mature P8, the reversed sequence of P8 residues 1-20 and a polyHis tag (AAHHHHHHA). A method was used to construct a library (Sidhu *et al.*, 2000a), which inserted the following hydrophobic-biased degenerate codons between codons 11 and 12 of the above-described ORF: 5′-NWTNKTNWTNYTNYTN KTNWTNWT NWTNWTNWTNKG NYTNKGNYTNW CNKTNWTNWT-3′ (N = A/C/G/T, W = A/T, K = G/ T, Y = C/T). The library contained 8.3×10^{10} unique members. Phage from the library were propagated, through packaging by VCSM13 helper phage (Stratagene), and cycled through rounds of binding selection (Sidhu *et al.*, 2000a) with an anti-His$_4$ antibody (Qiagen) as the capture target. After three or four rounds of selection, individual clones were assayed for polyHis tag display using a phage ELISA (Sidhu *et al.*, 2000a) with either the anti-His$_4$ antibody or bovine serum albumin (BSA) as target.

Selection of a coat protein for hGH display

The method described by (Kunkel *et al.*, 1987) was used to insert a *Nsi*I restriction site followed by a *Xba*I restriction site into a phagemid encoding ACP-1,

between the regions encoding ACP-1 and the polyHis flag. After *Nsi*I and *Xba*I digestion of the resultant phagemid, a similarly digested DNA fragment encoding hGHsm (hGH-H1.4b/H4b.4A/M1.5B/L1.6E) (Lowman & Wells, 1993) was inserted. The resulting phagemid (pS1239b) contains an ORF encoding ACP-1 followed by a tetrapeptide linker (AADA) and hGHsm. To obtain a ACP variant capable of displaying hGHsm as a carboxy-terminal fusion, libraries were constructed to vary the sequence of ACP-1 encoded by pS1239b, using a described method (Sidhu *et al.*, 2000a). ACP-1 was divided into six zones with each zone containing contiguous residues, as follows: residues 2 to 7; 6 to 11; 12 to 21; 21 to 30; 31 to 40; 41 to 50. Libraries replaced all codons within each zone with an equal number of degenerate codons (NNS, where N = A/C/G/T, S = G/C) encoding all 20 natural amino acid residues. In addition, a library was designed to insert 14 degenerate codons (VVC, where V = A/C/G) encoding the residues A, R, N, D, G, H, P, S, or T into the middle of the tetrapeptide linker connecting ACP-1 to hGHsm. The diversities of these libraries were as follows: linker library, 2.8×10^{10}; zones, in the above order: 2.5×10^{10}; 2.5×10^{10}; 2.6×10^{10}; 2.4×10^{10}; 2.4×10^{10}; 2.3×10^{10}. Phage from all libraries were pooled (total diversity = 1.7×10^{11}) and cycled through rounds of binding selection with hGHbp as the binding target. After two rounds of selection, 24 individual clones were analyzed for hGH display using a phage ELISA (Sidhu *et al.*, 2000a). A single positive clone was isolated and the new coat protein was designated ACP-7.

To ensure that the ACP-7 sequence was necessary and sufficient for hGH display on M13 phage, we used site-directed mutagenesis to convert ACP-1 into ACP-7. The display of hGH on phage particles was detected using a phage ELISA (Sidhu *et al.*, 2000a).

Acknowledgments

References

Bass, S., Greene, R. & Wells, J. A. (1990). Hormone phage: an enrichment method for variant proteins with altered binding properties. *Proteins: Struct. Funct. Genet.* **8**, 309-314.

Boyd, D. & Beckwith, J. (1990). The role of charged amino acids in the localization of secreted and membrane proteins. *Cell,* **62**, 1031-1033.

Clackson, T. & Wells, J. A. (1994). *In vitro* selection from protein and peptide libraries. *Trends Biotechnol.* **12**, 173-184.

Crameri, R. & Suter, M. (1993). Display of biologically active proteins on the surface of filamentous phages: a cDNA cloning system for selection of functional gene products linked to the genetic information responsible for their production. *Gene,* **137**, 69-75.

Dornburg, R. (1997). From the natural evolution to the genetic manipulation of the host-range of retroviruses. *Biol. Chem.* **378**, 457-468.

Fanning, A. S. & Anderson, J. M. (1999). PDZ domains: fundamental building blocks in the organization of protein complexes at the plasma membrane. *J. Clin. Invest.* **103**, 767-772.

Fuh, G., Mulkerrin, M. G., Bass, S., McFarland, N., Brochier, M., Bourell, J. H., Light, D. R. & Wells, J. A. (1990). The human growth hormone receptor. Secretion from *Escherichia coli* and disulfide bonding pattern of the extracellular binding domain. *J. Biol. Chem.* **265**, 3111-3115.

Gennis, R. B. (1989). Membrane biogenesis. In *Biomembranes: Molecular Structure and Function* (Cantor, C. R., ed.), pp. 370-417, Springer-Verlag, New York.

Greenwood, J., Willis, A. E. & Perham, R. N. (1991). Multiple display of foreign peptides on a filamentous bacteriophage. *J. Mol. Biol.* **220**, 821-827.

Guan, C., Li, P., Riggs, P. D. & Inouye, H. (1987). Vectors that facilitate the expression and purification of foreign peptides in *Escherichia coli* by fusion to maltose-binding protein. *Gene,* **67**, 21-30.

Iannolo, G., Minenkova, O., Petruzzelli, R. & Cesareni, G. (1995). Modifying filamentous phage capsid: limits in the size of the major capsid protein. *J. Mol. Biol.* **248**, 835-844.

Karaolis, D. K. R., Somara, S., Maneval, D. R., Jr., Johnson, J. A. & Kaper, J. B. (1999). A bacteriophage encoding a pathogenicity island, a type-IV pilus and a phage receptor in cholera bacteria. *Nature,* **399**, 375-379.

Keifer, D., Hu, X., Dalbey, R. & Kuhn, A. (1997). Negatively charged amino acid residues play an active role in orienting the Sec-independent Pf3 coat protein in the *Eschericha coli* inner membrane. *EMBO J.* **16**, 2197-2204.

Kunkel, T. A., Roberts, J. D. & Zakour, R. A. (1987). Rapid and efficient site-specific mutagenesis without phenotypic selection. *Methods Enzymol.* **154**, 367-382.

Lowman, H. B. & Wells, J. A. (1993). Affinity maturation of human growth hormone by monovalent phage display. *J. Mol. Biol.* **234**, 564-578.

Malik, P., Terry, T. D., Gowda, L. R., Langara, A., Petukhov, S. A., Symmons, M. F., Welsh, L. C., Marvin, D. A. & Perham, R. N. (1996). Role of capsid structure and membrane protein processing in determining the size and copy number of peptides displayed on the major coat protein of filamentous bacteriophage. *J. Mol. Biol.* **260**, 9-21.

Marvin, D. A. (1998). Filamentous phage structure, infection and assembly. *Curr. Opin. Struct. Biol.* **8**, 150-158.

Marvin, D. A., Hale, R. D., Nave, C. & Citterich, M. H. (1994). Molecular models and structural comparisons of native and mutant class I filamentous bacteriophages. Ff (fd, f1, M13), If1 and IKe. *J. Mol. Biol.* **235**, 260-286.

Rapoza, M. P. & Webster, R. E. (1995). The products of gene I and the overlapping in-frame gene XI are required for filamentous phage assembly. *J. Mol. Biol.* **248**, 627-638.

Scott, J. K. & Smith, G. P. (1990). Searching for peptide ligands with an epitope library. *Science,* **249**, 386-390.

Sidhu, S. S., Lowman, H. B. & Wells, J. A. (2000a). Phage display for selection of novel binding peptides. *Methods Enzymol.* In the press.

Sidhu, S. S., Weiss, G. A. & Wells, J. A. (2000b). High copy display of large proteins on phage for functional selections. *J. Mol. Biol.* **296**, 487-495.

Smith, G. P. (1985). Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. *Science,* **228**, 1315-1317.

Smith, G. P. & Petrenko, V. A. (1997). Phage display. *Chem. Rev.* **97**, 391-410.

von Heijne, G. (1986). A new method for predicting signal sequence cleavage sites. *Nucl. Acids Res.* **14**, 4683-4690.

Webster, R. E. (1996). Biology of the filamentous bacteriophage. In *Phage Display of Peptides and Proteins* (Kay, B. K., Winter, J. & McCafferty, J., eds), pp. 1-20, Academic Press, San Diego, USA.

Weiss, G. A., Sidhu, S. S. & Wells, J. A. (2000). Mutational analysis of the major coat protein of M13 identifies residues that control protein display. *Protein Sci.* **9**, 647-654.

Edited by G. Von Heijne