ELSEVIER

Journal of Biotechnology 93 (2002) 1–14

JOURNAL OF Biotechnology

www.elsevier.com/locate/jbiotec

Functional selection of phage displayed peptides for facilitated design of fusion tags improving aqueous two-phase partitioning of recombinant proteins

Nina Bandmann [a], James Van Alstine [b,1], Andres Veide [a], Per-Åke Nygren [a,*]

[a] *Department of Biotechnology, Royal Institute of Technology (KTH), SE-106 91 Stockholm, Sweden*
[b] *Department of Chemical Engineering and Technology, Royal Institute of Technology (KTH), SE-100 44 Stockholm, Sweden*

Received 4 January 2001; received in revised form 27 April 2001; accepted 29 June 2001

Abstract

Aqueous two-phase systems allow for the unequal distribution of proteins and other molecules in water-rich solutions containing phase separating polymers or surfactants. One approach to improve the partitioning properties of recombinant proteins is to produce the proteins as fused to certain peptide tags. However, the rational design of such tags has proven difficult since it involves a compromise between multivariate parameters such as partitioning properties, solvent accessibility and production/secretion efficiency. In this work, a novel approach for the identification of suitable peptide tag extensions has been investigated. Using the principles of selection, rather than design, peptide sequences contributing to an improved partitioning have been identified using phage display technology. A 40 million member phagemid library of random nona-peptides, displayed as fusion to the major coat protein pVIII of the filamentous phage M13, was employed in the selection of top-phase partitioning phage particles in a PEG/sodium phosphate system. After multiple cycles of selection by partitioning, peptides with high frequencies of both tyrosine and proline residues were found to be over represented in selected clones. The identified peptide sequences, or derivatives thereof, were subsequently individually analyzed for their partitioning behavior as displayed on phage, as free synthetic peptides and as genetically fused to a recombinant model target protein. The results showed that novel peptide sequences capable of enhancing top-phase partitioning without interfering with protein production and secretion indeed could be identified for the aqueous two-phase system investigated.

Keywords: Aqueous two-phase system; Peptide library; Partitioning; Phage display; Selection; Recombinant proteins; Fusion tag; PEG/sodium phosphate system

1. Introduction

Molecular surface properties causing differences in partitioning behavior in aqueous two-phase systems can be used, for example, in the

* Corresponding author. Tel.: +46-8-55378328; fax: +46-8-55378481.
E-mail address: perake@biochem.kth.se (P.Å. Nygren).
[1] Present address: Amersham Pharmacia Biotech AB, SE-751 82 Uppsala, Sweden.

0168-1656/02/$ - see front matter
PII: S0168-1656(01)00383-2

separation of biomaterials and characterization of macromolecules, viruses and cells (Albertsson, 1986; Walter and Johansson, 1994). Aqueous two-phase systems, characterized by the spontaneous formation of two water-rich phases, can be formed from addition of several different components such as polymers, polymer/salt or detergents (Walter and Johansson, 1994). The use of aqueous two-phase partitioning for protein purification has several advantages such as mild recovery conditions due to the high water content in the two phases, ease of scale up and short processing time. Use of this separation technique in the purification of recombinant proteins, often allows clarification, concentration and target protein purification unit operations to be combined in one primary recovery step (Veide et al., 1983). To achieve high product yield and purity it is desirable that the target protein has a strong partitioning to one of the phases in the system, while contaminating proteins, nucleic acids and cell debris partition into the opposite phase. System-related factors influencing partitioning in the extraction system are pH, temperature and the type and concentration of the phase forming components. In addition, the partitioning behavior of a protein depends strongly on the surface properties of the protein such as charge and hydrophobicity.

Genetic engineering of proteins to alter their partitioning properties has often been performed, either by using peptides or proteins as fusion tags or by direct point mutation of the protein (Köhler et al., 1991a,b; Eiteman et al., 1994; Hassinen et al., 1994; Fan et al., 1998; Berggren et al., 2000a,c; Luther and Glatz, 1994). Such work has shown, for example, that the presence of preferentially surface exposed tryptophans in peptides and proteins generally favors partitioning to the PEG-rich top phase both in PEG/salt and polymer–polymer systems (Köhler et al., 1991a,b; Eiteman et al., 1994; Carlsson et al., 1996; Berggren et al., 2000b). The effect of charge is not as clear, but more negatively charged proteins tend to have higher partitioning to the PEG-rich top phase than positively charged proteins in PEG/salt systems (Flanagan et al., 1991; Asenjo et al., 1994; Franco et al., 1996). In polymer–polymer systems the preference of a net charged protein for the different phases can be influenced by the inclusion of a salt in the system composed of an anion and a cation with different affinities for the two phases (Johansson, 1974). However, for the design of peptide extensions, other factors are also to be considered, including interactions between residues in the peptide and the target protein, the efficiency of exposure of different peptide tags to the solution, as well as the influence of the peptide on expression levels, activity and stability of the resulting fusion protein. This complexity makes partitioning behavior hard to predict and complicates design of fusion tags. Therefore, a general method for the identification of peptide tags taking these aspects into account and being applicable to any two-phase system would be valuable.

In phage display technology foreign peptides or proteins are displayed on the surface of filamentous bacteriophage particles such as, Fd or M13 phage (Smith and Petrenko, 1997). In these virions, the phage coat is composed of approximately 2700 monomer units of the 50-residue major coat protein (pVIII), and a few copies of each of the several minor coat proteins, among them pIII, located at the ends (Kay et al., 1996; Russel, 1991). If randomized oligonucleotides are inserted at an appropriate site in the structural gene of one of the coat proteins, a phage peptide library, with virions displaying unique peptides on the surface can be created, containing the corresponding gene in the packed DNA. In this way, a physical link between the peptide and the corresponding coding sequence is established. Using screening procedures, such as, biopanning, peptides that for example interact with target proteins of interest can be selected from such libraries (Devlin et al., 1990; Scott and Smith, 1990; Barbas, 1993; Burritt et al., 1996).

Peptide libraries displayed on phage have been extensively used for selection of ligands specifically binding to target proteins such as antibodies, receptors and enzymes (Cortese et al., 1996). In this work, the potential of using phage display selection technology for the identification of peptide sequences suitable for use as fusion tags to improve the partitioning of recombinant proteins in aqueous two-phase systems, was investigated.

The use of a selection strategy rather than a design approach should increase the speed with which peptide sequences suitable for use in different aqueous two-phase systems can be identified. In addition, selection systems based on in vivo production and display of peptide sequences should potentially allow simultaneous selection of sequences fulfilling a 'package' of requirements including proteolytic stability, efficient production and secretion, all highly important properties of peptides used for tagging of recombinant proteins for large scale production.

2. Materials and methods

2.1. Bacterial strains, vectors and phage library

The *Escherichia coli* strain RRIΔM15 (*supE*) (Ruther, 1982) was used for all phage and cloning work and for protein production. The phagemid vector pC89 and the phage peptide library have been described earlier by Felici (Felici et al., 1991) and was kindly supplied by Instituto di Richerche di Biologia Molecolare P. Angeletti SpA (IRBM), Italy. For secreted expression of recombinant proteins the plasmid pEZZmp18 (Ståhl et al., 1989) was used, resulting in fusion to a divalent IgG-binding fusion partner ZZ (Nilsson et al., 1987).

2.2. DNA constructs and DNA sequencing

The cloning work was performed according to standard methods (Sambrook et al., 1989). Restriction endonucleases, T4 DNA-ligase and ligase buffer obtained from New England Biolabs Inc, MA, were used as recommended. Construction of the fusion protein encoding vectors pEZZ-cutinase-wt and pEZZ-cutinase-$(WP)_4$ was described previously (Bandmann et al., 2000). DNA linkers were formed by hybridization of oligonucleotides (Interactiva, Germany), encoding the different peptide tags (P5:32): 5′-TCGACCCGTACG-CTTACCAGTACCCGTCTTCTTAACTGCA-3′ and 5′-GTTAAGA AGACGGGTACTGGTAA-GCGTACGGG-3′; (P5:32-R): 5′-TCGAC CC-GCGTGC TCGTCAGCGTCCGTCTTCTTAA-CTGCA-3′ and 5′-GTTAAGAA GACGGACG-CTGACGAGCACGCGGG-3′; $(YP)_4$: 5′-TCGA-CTACCCGTA CCCGTACCCGTACCCGTA-ACTGCA-3′ and 5′-GTTACGGGTACG GGT-ACGGGTACGGGTAG-3′; $(Y)_4$: 5′-TCGACTA-CTACTACTACTAACTGCA-3′ and 5′-GTTA-GTAGTAGTAGTAG-3′ and $(Y)_8$: 5′-TCGAC-TACTACTACTACTACTACTACTACTAACT-GCA-3′ and 5′-GTTAGTAG TAGTAGTAG-TAGTAGTAGTAT-3′. The linkers were subsequently inserted by ligation into the vector pEZZ-cutinase-wt, digested with the enzymes *Sal*I and *Pst*I. Positive clones of pEZZ-cutinase-P5:32, -P5:32-R, -$(YP)_4$, -$(Y)_4$ and -$(Y)_8$ were identified by PCR screening. In addition, nucleotide sequences encoding peptide inserts, both in the vectors encoding ZZ-cutinase-tagged fusion proteins and in the phagemid vectors, were verified by cycle DNA sequencing using BigDye™-terminators (Amersham Pharmacia Biotech, Sweden), according to the supplier's instructions, and fragment analysis on an ABI PRISM™ 377 DNA Sequencer (Perkin Elmer, CA). Protein production and IgG-affinity chromatography purification were performed as described elsewhere (Bandmann et al., 2000).

2.3. Preparation of phage stocks

For first round selections, the original peptide phage library was used as supplied. Phage stocks from clones selected in the aqueous two-phase selection rounds were prepared using standard procedures including PEG precipitation (Kay et al., 1996), reproducibly yielding titers of 10^{12}–10^{13} colony forming units (cfu) per ml. The 'wild type' phage stock was prepared using cells transformed with the pC89 phagemid vector (Felici et al., 1991), containing no peptide encoding sequence insert, thus resulting in phage particles not displaying modified pVIII variants.

2.4. Aqueous two-phase system partitioning

PEG 4000 (molecular weight distribution of 3500–4500), NaH_2PO_4 and Na_2HPO_4 were obtained from Merck, Germany. In the following,

concentrations of the phase forming components are given as % weight/weight (w/w). A batch phase system of 20 g was made by weighing 40% PEG 4000 and sodium phosphate (base/acid molar ratio of 1.42, pH ≈ 7.0) stock solutions in appropriate amounts. The top and bottom phases were subsequently separated and stored at 4 °C. Around 2 ml aliquots of two-phase systems used for phage selections were prepared by adding 950 μl of top and bottom-phases, respectively, and 100 μl phage containing water solution to a test tube. This resulted in an aqueous two-phase system with an 11/10 top/bottom phase volume ratio. The total number of original library phage particles added to the two-phase system for a first selection round, was approximately 10^{10} cfu. The system was mixed by inversion for approximately 10 min and left to separate for approximately 15 min. The phage particle concentration in the two separate phases was determined by withdrawing aliquots of 15 μl from the top and bottom phase, respectively, with subsequent reinfection of *E. coli* followed by plating on selective media (ampicillin) to allow counting of colony forming units. The remainder of the top phase (700–850 μl) was transferred to 950 μl of fresh bottom phase for a subsequent round of partitioning. The loss of top phase volume due to the enumeration procedure was compensated for by addition of a corresponding volume of fresh top phase. About 100 μl distilled water was added to give the final volume of 2 ml. This procedure was repeated in all consecutive partitioning cycles.

For one-cycle partitioning of free peptides and fusion proteins, the same composition of aqueous two-phase system and partitioning procedure were used. Resulting K values were calculated from at least triple partitioning experiments. For photometrical measurements, samples were collected from top and bottom phases and added to an equal volume of opposite pure phase and to one volume of distilled water. This converts the system to a homogeneous phase and gives a constant background. The total peptide and fusion protein concentrations in the two-phase systems were approximately 100–500 and 5–100 μg ml^{-1}, respectively.

2.5. Peptide and fusion protein analysis

The synthetic peptides P5:32, P5:32-R, $(YP)_4$ and $(WP)_4$ with a purity of >95% were purchased from SynPep Corporation, CA. The partitioning ratios of the peptides in the two-phase systems were determined by measuring the absorbance at 280 nm (P5:32, $(YP)_4$ and $(WP)_4$) or at 220 nm (P5:32-R) using a Cary 50 Bio spectrophotometer (Varian Incorporated, CA). Para-nitrophenyl butyrate (pNPB) and sodium taurodeoxycholate (TDOC) used in the cutinase enzyme assay were purchased from Sigma, Sweden. To a mixture of 955 μl assay buffer (10 mM Tris–HCl, 10 mM NaCl and 50 mM TDOC, pH 8.0) and 20 μl of substrate solution (50 mM pNPB in acetonitrile), 25 μl of sample was added. The components were immediately mixed and analyzed. The hydrolysis of pNPB was spectrophotometrically monitored at 25 ± 0.1 °C using a Cary 50 Bio spectrophotometer (Varian Incorporated, CA), following the formation of the para-nitrophenol (pNP) at 400 nm. The activity yields of the cutinase variants and the yields of oligopeptides in the phase systems relative to added amount was 100 ± 10 and 100 ± 15% respectively.

The fusion proteins were analyzed by SDS-PAGE on a 10–20% gradient gel (NOVEX, CA) under reducing conditions (Laemmli, 1970). Coomassie Brilliant Blue R-350 (Pharmacia Biotech, Sweden) was used for staining.

Electrospray mass spectra were recorded on a Quadrupole-time-of-flight instrument (Morris et al., 1996) (Q-TOF, Micromass, Manchester, England).

3. Results

3.1. Phage partitioning

Phage display selection technology was investigated as a means to identify peptide sequences useful for tagging recombinant proteins to increase their partitioning to the PEG-rich top phase in a PEG4000/Na-phosphate aqueous two-phase system. Based on the hypothesis that phage particles individually decorated with multiple

copies of a foreign peptide should show peptide sequence-dependent phase preferences, phage clones with desired partition characteristics should be possible to isolate and identify. As starting material for selection of top phase partitioning phage particles, a 40 million member phagemid library of random nona-peptides displayed as fusions to the major coat protein pVIII of the *E. coli* filamentous phage M13 (Fig. 1), was used (Felici et al., 1991). This display system is designed to result in multiple rather than single-copy display of pVIII peptide fusions on the phage surface, interspersed between wt and pVIII copies, which should contribute to an increased effect on aqueous two-phase system partitioning.

Phage particle partition experiments were performed in a standard aqueous two-phase system consisting of 11% PEG and 8.5% sodium phosphate (base/acid molar ratio of 1.42) with a tie line length of 21.5. For phage particles, a partitioning coefficient K_{phage} was defined as:

$$K_{phage} = \frac{\text{phage concentration in top phase}}{\text{phage concentration in bottom phase}}$$

where, the phage concentrations were determined by counting of colony forming units (cfu) after plating of *E. coli* cells on selective media following infection with aliquots of top and bottom phase samples, respectively.

Initial partition experiments with wild type M13 phages resulted in a K_{phage} value of 1×10^{-3}, indicating that non-decorated phage particles showed a strong bottom phase preference. This low K value obtained for wild type M13 phage suggested that recombinant phage particles isolated from the top phase should correspond to decoration with peptides of strong top phase preference.

Different formats for multiple-step aqueous two-phase selections (Fig. 2) were investigated. These differed with respect to whether or not phage populations isolated from the top phase were amplified in *E. coli* (infection followed by production of new phage stock solution) between selection rounds. In selection series including such amplification (format 1, Fig. 2), non-conclusive results were routinely obtained, with no observed increases in phage pool K values in successive rounds and no patterns regarding peptide sequences of analyzed clones (data not shown). In contrast, omitting the amplification step (format 2, Fig. 2) resulted in increasingly higher partition values (confirmed by several repeated experiments) for phage pools over successive rounds, and observable trends in peptide sequences. In addition, experiments performed with phage stock solutions obtained from individual clones corresponding to the defined peptides with expected high K values (see below), have shown that in order to be able to observe partitioning differences between phage clones, more than one round of consecutive partitioning without intervening

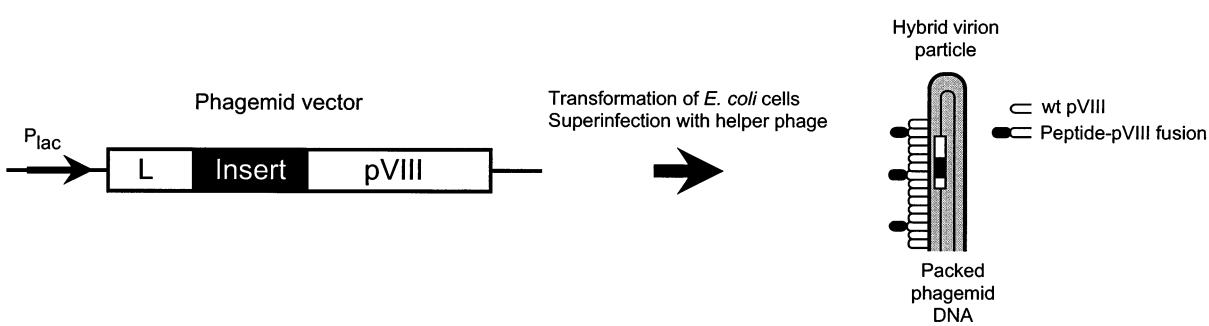

Fig. 1. Schematic description of the *lac* operon controlled phagemid expression cassette and the corresponding hybrid virions used for selections of top phase partitioning phage particles (Felici et al., 1991). Peptide-pVIII fusions are synthesized with a 23-residue leader peptide (L), later processed during secretion. The use of a phagemid system results in multiple copy display of the foreign peptide interspersed with copies of the wild type pVIII protein. The packed phagemid DNA contains the gene encoding the particular peptide displayed.

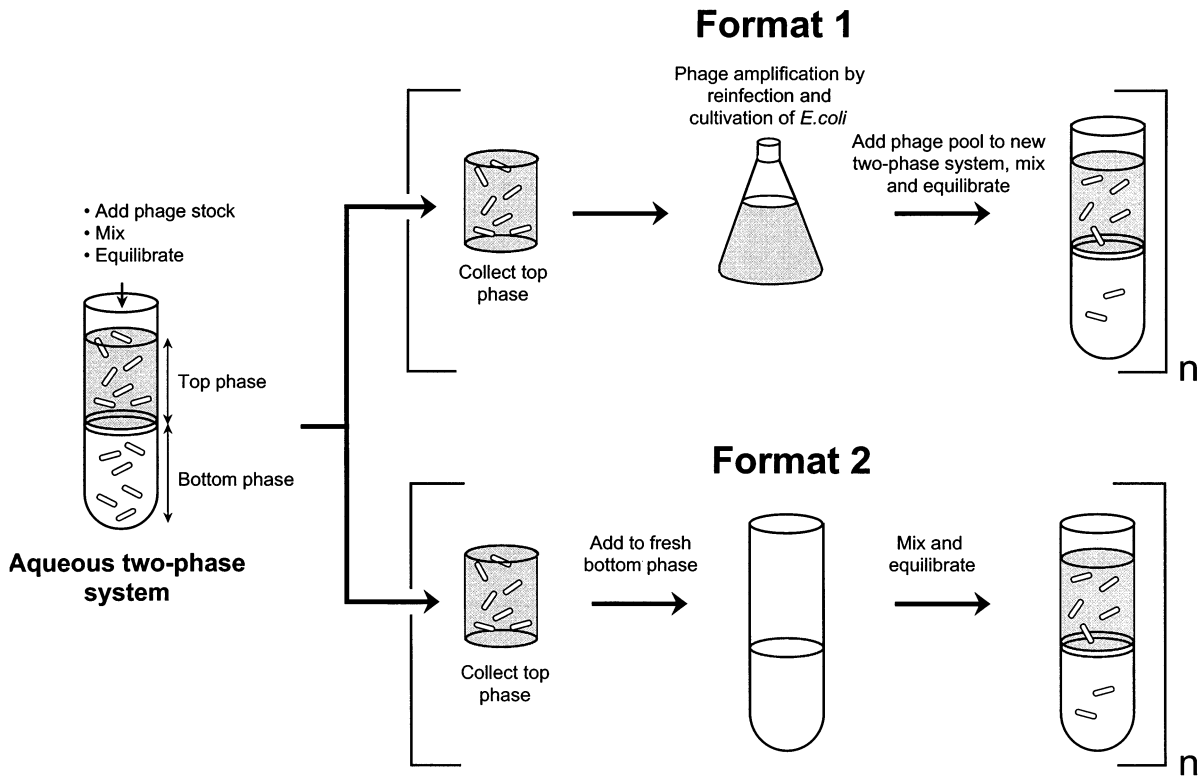

Fig. 2. Schematic description of the two formats used for multi-step (n) selection of peptide-decorated phage particles with desired phase partitioning properties. A phage stock solution containing phage particles genetically decorated with individual peptide sequences is added to an aqueous two-phase system. After mixing and equilibration the top phase is collected and in format 1 used to reinfect *E. coli* for amplification of the phage population. An aliquot of the obtained phage stock is subsequently added to a fresh two-phase system for a new round of partitioning. In format 2, the collected top phase is directly added to a fresh bottom phase for a new round of partitioning.

amplification was generally necessary (see below and discussion).

In one series of aqueous two-phase selections of the phage peptide library involving five rounds of consecutive phage partitioning, where phage particles were isolated from the top phase fraction and added directly to a fresh bottom phase for a consecutive round (format 2), significant increases in phage population K_{phage} values were observed (Table 1). To rule out possible influence on K_{phage} values by a decreasing total phage concentration, control partitioning was performed at different phage particle concentrations without any such effect observed (data not shown).

In total, 276 peptide-encoding phagemid inserts were subjected to DNA sequencing, including individual clones of the starting library (76 clones) and phages isolated from the top phases in each cycle of selection (25 clones from the first four cycles and 100 clones from the fifth cycle). This allowed for an analysis of deduced peptide sequences for identification of traits in top phase-

Table 1
Phage pool K values determined during the selection

Cycle number	K
1	0.01
2	0.01
3	0.06
4	0.6
5	13

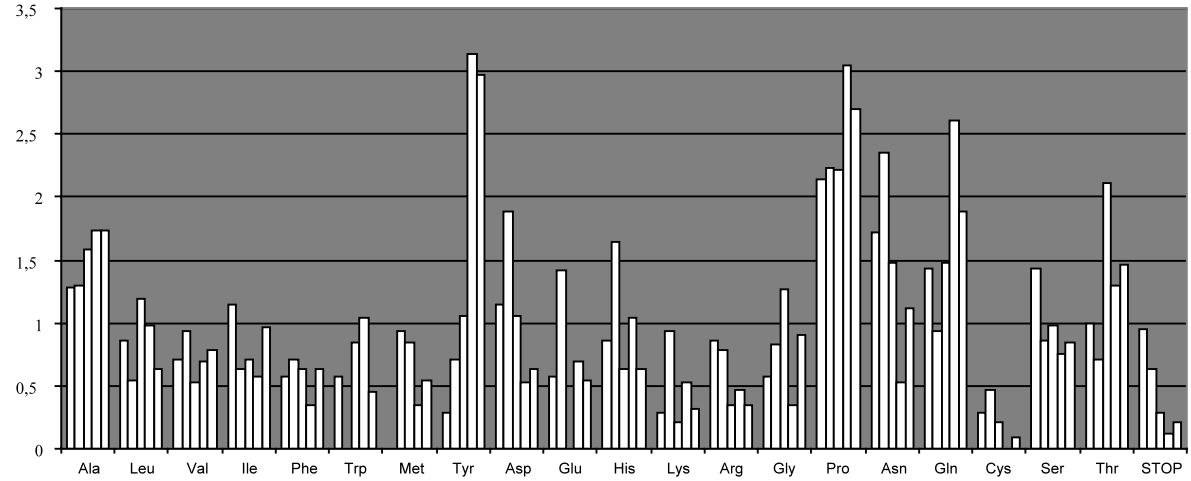

Fig. 3. Histogram plot of amino acid appearance frequencies in aqueous two-phase system selected peptides compared to their respective theoretical frequencies according to the library design. Values above or under one (1) thus correspond to over and under representation, respectively. The values for the respective amino acids are shown from left to right for the five cycles of aqueous two-phase phage selections, using format 2. STOP corresponds to any of the three stop codons.

preferring peptide variants. From the 100 clones isolated after the fifth cycle (representing 10% of the total output) 79 were found to contain peptide encoding inserts without stop codons. An alignment of these sequences showed that none of the variants were represented more than once (data not shown).

The peptide sequence data was further analyzed statistically by a round-by-round comparison of frequencies with which the different amino acids appeared in top-phase selected clones to expected frequencies calculated from the genetic variation used for library construction (Felici et al., 1991). Only the first eight N-terminal amino acids of the peptide inserts were included in this analysis since their corresponding nucleotide sequences had been completely randomized in the library construction, in contrast to the codons for the ninth amino acid which all contained a fixed guanine nucleotide (G) in the third position. In the histogram plot (Fig. 3) of the data, values above or under one (1) correspond to an over or under representation, respectively. In the peptide sequence inserts for clones isolated after the fifth round of phage partitioning, both tyrosine and proline showed more than a 2.5-fold over representation. Whereas, the proline frequency was high during the entire experiment, the frequency of tyrosine increased from values below one in early cycles to values around three in later cycles. Of the identified tyrosine residues, approximately one third was positioned in the first (i.e. N-terminal) position in the peptides (data not shown). Other amino acids showing high frequencies in later cycles included alanine, glutamine and threonine, although with lower values than for tyrosine and proline and with somewhat inconsistent cycle number variation (glutamine and threonine). Interestingly, both lysine and arginine residues appeared at low frequencies in peptide inserts of phage clones isolated from the top phases in later cycles.

The relative frequency of stop codons in the peptide sequences decreased significantly throughout the partitioning rounds. Due to the N-terminal positioning of the peptides in fusion to the phage coat protein VIII, such phage clones do not display any foreign peptides and thus have protein coats phenotypically corresponding to wild type M13 phage. During the DNA sequencing, phagemids devoid of peptide encoding inserts were also observed (re-ligated vectors). Such variants decreased from about 60% in the starting library to approximately 10% in the top phase of

Table 2
Listing of dive selected peptide sequences isolated after the fifth aqueous two-phase selection round

Peptide identity	Sequence (N–C terminal)[a]
P5:6	**Y P Y P** I S A S **P**
P5:20	**Y P P** I **Y** T A **P** N
P5:26	**Y P** I A **P** T L **Y P**
P5:32	**P Y** A **Y** Q **Y P** S S
P5:92	**Y P P** T **P** I **Y** A T

[a] Tyrosines and prolines are shown in bold face, with tyrosine–proline dipeptide motifs underlined.

the final round (data not shown). Both these observations are in accordance with the earlier finding that wt-phenotype phage shows a strong preference for the bottom phase in the PEG/sodium phosphate system used.

Based on the results from the statistical analysis, five of the most tyrosine and proline rich sequences corresponding to phage clones isolated after the fifth aqueous two-phase selection cycle were chosen as candidates for further studies (Table 2). To investigate if any partitioning differences could be observed between these phage clones and wild type phage, separate multiple-round aqueous two-phase selections were performed with individually prepared phage stocks. The results summarized in Table 3 show that all five phage clones had significantly higher K values (top/bottom phage titer ratios) than the wild type phage clone. Noteworthy, significant differences between the different clones did not appear before the third cycle of partitioning (see discussion).

No significant adsorption of phage particles to the interface, neither for the wild type phage, the phage library nor the individual phage clones was observed under the conditions used (data not shown).

3.2. Partitioning of free peptides and fusion proteins

To make a first investigation if the partitioning properties in the PEG/sodium phosphate system were dependent on context of the display on the phage or were inherent properties of the peptides themselves, partitioning experiments were performed with four free peptides produced through chemical peptide synthesis. The same two-phase system composition was used as in the phage selections. For this experiment the most significantly partition altering peptide (clone P5:32) was synthesized as selected from the library. To elucidate if its properties were dependent on the tyrosine residues and to examine the role of arginine (counter-selected for in the selections, see Fig. 3), a variant of this peptide (P5:32-R) in which the tyrosines were replaced with arginines was included in the study. In addition, a $(YP)_4$ peptide corresponding to a four-repeat of the frequently observed motif YP (Table 2) was included. For comparison, a tryptophan rich peptide $(WP)_4$, which earlier has shown significant top-phase partitioning properties in this type of aqueous two-phase systems (Bandmann et al., 2000), was also used in the study.

In the results summarized in Table 4, a significant top phase partitioning was observed for the free P5:32 peptide, reflecting the effect seen in the phage partitioning. The variant peptide in which the tyrosines had been replaced with arginines

Table 3
Individual partitioning data for wild type M13 phage and five selected phage clones

Round	Partitioning data (K)					
	Wild type M13	P5:6	P5:20	P5:26	P5:32	P5:92
1	0.071	0.0039	0.0043	0.00046	0.0043	0.0021
2	nd	0.036	0.0029	0.00019	0.012	0.0027
3	0.0012	0.74	0.016	0.87	0.22	0.062
4	0.0082	14	1.1	4.4	3.0	1.8
5	0.0020	4.8	2.3	1.0	55	4.4

Table 4
Partitioning data for four synthetic peptides

Peptide	K
P5:32	4.6 ± 0.42
P5:32-R	0.36 ± 0.17
$(YP)_4$	13 ± 1.9
$(WP)_4$	59 ± 15

showed a markedly lower top phase partitioning. This supports the notion of the importance of tyrosines for top phase partitioning and the preference for the bottom phase of arginines in this aqueous two-phase system. Interestingly, the $(YP)_4$ peptide showed a higher K value than the P5:32 peptide, suggesting a strong influence of the YP dipeptide motif. As expected, the $(WP)_4$ peptide showed strong affinity for the top phase.

To further investigate partitioning properties of the YP motif when fused to a larger recombinant protein, five recombinant fusion proteins were constructed. The *Fusarium solani pisi* lipase cutinase (Carvalho et al., 1999), used as a model target protein in an earlier study (Bandmann et al., 2000) was genetically extended with sequences corresponding to the C-terminal peptide tags P5:32, P5:32-R, $(YP)_4$, $(Y)_4$ and $(Y)_8$. The $(Y)_4$ and $(Y)_8$ tags were included to examine the individual roles of proline and tyrosines, respectively, in the $(YP)_4$ tag. The five modified cutinase enzymes were produced as fused to a 14 kDa N-terminal IgG-binding affinity fusion partner ZZ, enabling one-step affinity recovery of fusion proteins. Two earlier investigated ZZ-cutinase fusion constructs, ZZ-cutinase-wt and ZZ-cutinase-$(WP)_4$ representing C-terminally untagged and $(WP)_4$-tagged cutinase, respectively (Bandmann et al., 2000), were also included in the study (Fig. 4).

Earlier observations based on secreted production in both yeast and *E. coli* of C-terminally tagged cutinase had shown that the expression yields of recombinant proteins are influenced by the peptide tag composition, with a negative effect from tryptophan residues (Bandmann et al., 2000 and unpublished). Here, the seven cutinase variants were produced in over-night shake flask cultures as secreted to the periplasm of *E. coli* cells and purified by IgG-affinity chromatography with recovery yields ranging from 2 to 30 mg l^{-1}. The ZZ-cutinase-$(Y)_4$ and ZZ-cutinase-$(YP)_4$ fusion proteins were recovered under these conditions at four-fold higher levels than the ZZ-cutinase-$(WP)_4$ reference, and the ZZ-cutinase-P5:32 and ZZ-cutinase-$(Y)_8$ fusion proteins showed two-fold higher recovery yields (data not shown). Analysis by SDS-PAGE suggested that all ZZ-cutinase fusion proteins were of apparent full-length sizes and with no signs of proteolytic degradation. Mass spectrometry analysis of two of the purified fusion proteins (ZZ-Cut-P5:32 and ZZ-Cut-$(Y)_8$) confirmed that both of these products had measured molecular weights in accordance with the presence of undergraded C-terminal tags (data not shown).

Partitioning experiments of purified fusion proteins were performed using the same PEG/sodium phosphate system composition as in the phage selections. K values for the different variants were determined through cutinase activity measurements in top and bottom phases, respectively. In addition, for all constructs, the partitioning improvement (PI):

$$\mathrm{PI} = \frac{K_{\text{ZZ-Cut-tag}}}{K_{\text{ZZ-Cut-wt}}}$$

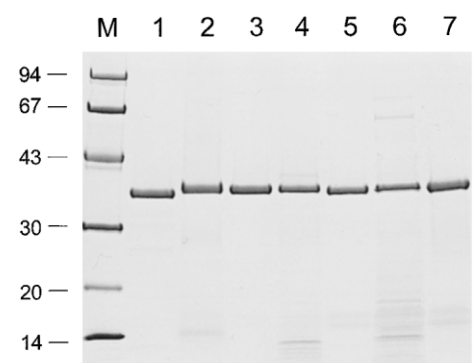

Fig. 4. SDS-PAGE analysis of the *E. coli* produced and IgG-affinity chromatography purified ZZ-cutinase variants evaluated in this study. *M*: molecular weight marker given in kDa. Lane 1: ZZ-cutinase-wt; lane 2: ZZ-cutinase-P5:32; lane 3: ZZ-cutinase-P5:32-R; lane 4: ZZ-cutinase-$(YP)_4$; lane 5: ZZ-cutinase-$(Y)_4$; lane 6: ZZ-cutinase-$(Y)_8$; lane 7: ZZ-cutinase-$(WP)_4$.

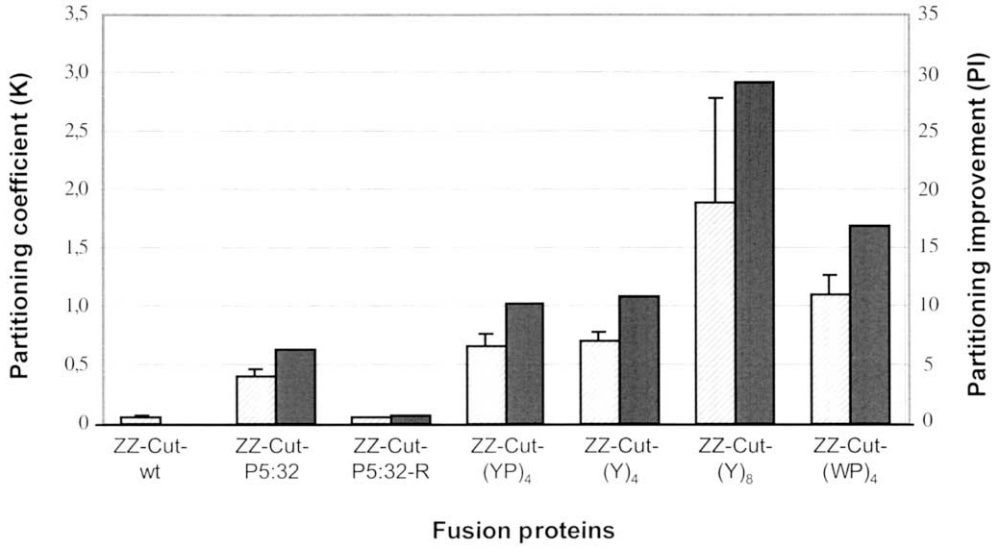

Fig. 5. Histogram plot showing calculated partitioning improvement and tag efficiency, respectively, for the different peptide-tagged ZZ-cutinase fusion proteins.

and the tag efficiency (TEF) (Berggren et al., 2000a):

$$\mathrm{TEF} = \frac{\log K_{\text{tagged prot.}} - \log K_{\text{prot.}}}{\log K_{\text{tag}}}$$

was calculated. The results, summarized in Fig. 5, showed that all the tagged fusion proteins, except for the construct containing the P5:32-R peptide, showed improved top phase partitioning, compared to the reference protein ZZ-cutinase-wt, with PI values in the range 6.3–29.

The use of the $(YP)_4$ tag resulted in a ten-fold partitioning improvement (PI = 10) compared to the untagged ZZ-Cut-wt construct ($K = 0.065$). Interestingly, also the $(Y)_4$ tagged construct showed a similar PI value (PI = 11). The largest effect was observed with the $(Y)_8$ tag with a resulting PI value of 29, which is even higher than the PI value of 17 observed for the $(WP)_4$ tagged reference construct. Interestingly, a comparison of tag efficiencies shows that the highest value was observed for the ZZ-Cut-P5:32 fusion protein (TEF = 1.2), containing the 'native' P5:32 peptide sequence as selected from the library (Table 5).

4. Discussion

This study shows the possibility to use phage display selection technology for identification of peptide sequences conferring improved partitioning behavior in aqueous two-phase systems. This novel application of the technology corresponds to an evaluation of interactions between surface-engineered phage particles and the surrounding solution, in contrast to the most common use involving the identification of ligand-target affinity pairs (Smith and Petrenko, 1997).

The format used for enrichment of top phase preferring phage particles did not include any amplification of the selected phage populations

Table 5
Tag efficiency factors for different tags when analyzed as fused to the recombinant ZZ-cutinase fusion protein

Fusion protein	TEF
ZZ-Cut-P5:32	1.2
ZZ-Cut-P5:32-R	0.18
ZZ-Cut-$(YP)_4$	0.91
ZZ-Cut-$(WP)_4$	0.69

between partitioning rounds. Consequently, no 'traditional' convergence into a population of iteratively and clonally amplified variants could be expected from the approach, as generally observed for routine multiple-round affinity based selections (Nord et al., 1997). Here, the use of a selection protocol with the intervening phage stock amplification between rounds (format 1) resulted in decreasing phage pool K values during successive rounds. However, if this step was omitted (format 2) phage pool K values showed progressively higher values throughout the procedure. One explanation as to why phage pool K values did not increase using format 1, could have its basis in the observation that when phage clones selected for strong top phase partitioning behavior (from format 2) were partitioned as individual phage preparations, significant partitioning effects could only be detected after several consecutive extraction rounds (Table 3). A possible explanation for this could be that a major part of the phage particles in a newly prepared phage pool (as in each round using format 1) are not at all or poorly decorated with peptide fusions (wild type phenotype). Consequently, any effects from decorated variants are efficiently diluted into a wild type phenotype background during initial rounds of partitioning until these particles represent a sufficiently large fraction of the total phage population. In addition, the PEG precipitation steps employed in the phage stock preparations in format 1 could obviously result in a loss of phage clones remaining in solution under the conditions used. The use of 'raw' phage stocks or alternative phage concentration procedures such as ultra-filtration should circumvent this potential problem. The use of a whole-phage single protein VIII gene system, with peptides genetically fused to all copies of the major coat protein, would potentially eliminate the problem with a large background of non-decorated phage particles. However, for display of peptides longer than six residues this could drastically reduce the number of variant peptides sequences possible to display, due to problems concerning functional phage coat assembly solely from peptide-extended pVIII copies (Greenwood et al., 1991; Iannolo et al., 1995).

The analysis of the 79 sense peptide sequences corresponding to the clones isolated from cycle five showed that none of the nonapeptide sequences were represented more than once. Although 10^{10} phage particles of the 4×10^7 member library was used as input, this is not entirely surprising considering accumulated volume losses owing to marginal pipetting and that the used format 2 did not include any phage amplification steps that could lead to clonal expansion. In addition, only approximately 10% of the total phage output were included in the analysis.

Assuming that the relative order of amino acids in the peptides were of second importance relative to a frequent presence of certain amino acids, the outcome of the statistical analysis was used to select sequences for further studies. The amino acid representation analysis of the peptide sequences isolated in the different partitioning cycles (Fig. 3) showed a significant enrichment for tyrosine residues. This is interesting considering the aromatic nature of this amino acid and the earlier observed effect on partitioning of the also aromatic tryptophan residue. Earlier work in this field has focused on the use of tryptophan residues for inclusion in peptide tags to obtain an improved PEG-phase partitioning. Interestingly, in our study, tryptophans were only sparsely represented in selected peptide sequences, corresponding to both early and late rounds of partitioning. It is possible that although tryptophans are capable of conferring PEG-rich phase affinity other factors, related to protein secretion, solubility and virion assembly, negatively influence their suitability for being frequently represented in recombinant peptide sequences destined for translocation and phage surface display. In studies of octa- or decapeptide libraries with display on all copies of pVIII, without any other selection pressure than growth and infectivity, the relative frequency of tryptophans in peptide inserts was lower in proven viable phage clones relative to clones analyzed before phage particle production (Iannolo et al., 1995, 1997). In addition the tryphtophan rich peptide $(WP)_4$ showed, in an earlier study (Bandmann et al., 2000), to negatively affect recombinant production of the model target protein cutinase (both wild type and affinity-tagged versions) when included as C-terminal peptide tag. These results could accentuate

an advantage of using biological systems for selection, which can serve as 'filters' for the exclusion of sequences not compatible with recombinant expression routes. However, it is possible that certain details/motifs in selected sequences are only relevant in the context of the particular selection system used, and may not find relevance in the design and performance of the final tag. In contrast to the gradually increasing frequency of tyrosine residues, the frequency of proline residues was high in all five partitioning cycles. It was interesting to find that the $(Y)_4$ and $(YP)_4$ tags both were able to improve by ten-fold the partitioning of the ZZ-cutinase target protein. These results suggest that the presence of tyrosines in phage displayed peptides contributed to phage top phase partitioning and that the high prevalence of prolines in several selected tags, leading to the design and testing of the $(YP)_4$ tag, seems to have been important mainly in the context of functional/efficient display of the peptides on the phage surface and not associated with increased partitioning. In fact, a high prevalence of prolines in displayed peptides has earlier been suggested to be associated with a growth advantage, possibly through a resulting compact coat structure favorable during the secretion process (Iannolo et al., 1997).

The pronounced difference in phase preference observed for the free synthetic P5:32 and P5:32-R peptides, which was also retained when the corresponding sequences were genetically fused to the ZZ-cutinase model protein, supported the proposed influences of tyrosine and basic amino acid residues, respectively, on partitioning in this system. Although not associated with the highest peptide partition value observed, the P5:32 tag with a sequence as selected from the library had the highest tag efficiency value, indicative of efficient exposure of the tag to the solution. This suggests that the presentation of the relatively few tyrosines in this 'native' tag is more efficient than in the other tags tested. Considering that the 40 million member-size of the library only represents a small fraction of all possible nonamer peptides, it is possible that selections using a fully sampled library could result in the identification of tag sequences with partitioning properties superior to those of a stretch of consecutive tyrosines. Nevertheless, the $(Y)_8$ peptide tag showing the highest partitioning values in this study should be an interesting candidate for future protein engineering studies aimed at improving the partitioning of target proteins. Interestingly, while this work was in progress, tyrosine-rich peptides were suggested as interesting candidates for tagging of target proteins to improve partitioning properties (Berggren et al., 2000c). This conclusion was based on solubility studies in water solutions and on partitioning properties in an ethylene oxide-propylene oxide (EOPO)/dextran system of a set of cutinase surface mutants and free peptides (Berggren et al., 2000c).

In summary, the results suggest that peptide library and selection technology could be a valuable tool for identification of peptide sequences with desired partitioning properties. The presented data shows the unequal distribution in an aqueous two-phase system of phage particles differing in surface properties and the possibility to use phage display for selection of such properties. The five individual phage clones chosen for such consecutive partitioning studies had significantly higher K values (Table 3) than the wild type phage clone, which indicates that they were selected in the aqueous two-phase system by virtue of the foreign peptide sequences displayed on their respective coats.

This work was a first investigation of this novel application of phage display. To refine this method, future work could for example include the use of larger libraries, presentation formats of different valencies and a selection procedure allowing phage amplification between rounds. The described principles should potentially be applicable also on proteins for more general studies of surface properties.

It should be noted that although top phase partitioning variants were of primary interest here, also bottom phase preferring peptides could find value in aqueous two-phase system applications. Furthermore, it should be interesting to apply a multiple step extraction procedure using a counter current distribution apparatus to increase the efficiency and resolution during the selection with the format described in this work (Al-

bertsson, 1965). Although, phage display technology was used here, also other display principles could be of interest (Li, 2000).

Acknowledgements

We wish to thank Instituto di Richerche di Biologia Molecolare P. Angeletti SpA (IRBM), Rome, Italy, for kindly supplying us with the pC89 phagemid vector and the phagemid library. This work was supported by grants from the EU project EU-BIO-4-CT96-0435, the Swedish Centre for Bioprocess Technology (CBioPT) at KTH, and the Swedish Natural Science Research Council (NFR).

References

Albertsson, P.-Å., 1965. Thin layer countercurrent distribution. Anal. Biochem. 11, 121–125.

Albertsson, P.-Å., 1986. Partition of Cell Particles and Macromolecules.

Asenjo, J.A., Schmidt, A.S., Hachem, F., Andrews, B.A., 1994. Model for predicting the partition behavior of proteins in aqueous two-phase systems. J. Chromatogr. A 668, 47–54.

Bandmann, N., Collet, E., Leijen, J., Uhlen, M., Veide, A., Nygren, P.A., 2000. Genetic engineering of the *Fusarium solani pisi* lipase cutinase for enhanced partitioning in PEG-phosphate aqueous two-phase systems. J. Biotechnol. 79 (2), 161–172.

Barbas, C.F.d., 1993. Recent advances in phage display. Curr. Opin. Biotechnol. 4 (5), 526–530.

Berggren, K., Tjerneld, F., Veide, A., 2000a. Peptide fusion tags with tryptophan and charged residues for control of protein partitioning in PEG potassium phosphate aqueous two-phase systems. Bioseparation 9, 69–80.

Berggren, K., Nilsson, A., Johansson, G., Bandmann, N., Nygren, P.-Å., Tjerneld, F., 2000b. Partitioning of peptides and recombinant protein-peptide fusions in thermo-separating aqueous two-phase systems: effect of peptide primary structure. J. Chromatogr. B 743, 295–306.

Berggren, K., Egmond, M.R., Tjerneld, F., 2000c. Substitutions of surface amino acid residues of cutinase probed by aqueous two-phase partitioning. Biochem. Biophys. Acta 1481 (2), 317–327.

Burritt, J.B., Bond, C.W., Doss, K.W., Jesaitis, A.J., 1996. Filamentous phage display of oligopeptide libraries. Anal. Biochem. 238 (1), 1–13.

Carlsson, M., Berggren, K., Linse, P., Veide, A., Tjerneld, F., 1996. Effects of fused tryptophan rich peptides to a recombinant protein A domain on the partitioning in polyethylene glycol-dextran and Ucon-dextran aqueous two-phase systems. J. Chromatogr. A 756, 107–117.

Carvalho, C.M., Aires-Barros, M.R., Cabral, J.M., 1999. Cutinase: From molecular level to bioprocess development. Biotechnol. Bioeng. 66 (1), 17–34.

Cortese, R., Monaci, P., Luzzago, A., Santini, C., Bartoli, F., Cortese, I., Fortugno, P., Galfre, G., Nicosia, A., Felici, F., 1996. Selection of biologically active peptides by phage display of random peptide libraries. Curr. Opin. Biotechnol. 7 (6), 616–621.

Devlin, J.J., Panganiban, L.C., Devlin, P.E., 1990. Random peptide libraries: a source of specific protein binding molecules. Science 249 (4967), 404–406.

Eiteman, M.A., Hassinen, C., Veide, A., 1994. A mathematical model to predict the partition of peptides and peptide-modified proteins in aqueous two-phase systems. Biotechnol. Prog. 10, 513–519.

Fan, W., Bakir, U., Glatz, C.E., 1998. Contribution of protein charge to partitioning in aqueous two-phase systems (in process citation). Biotechnol. Bioeng. 59 (4), 461–470.

Felici, F., Castagnoli, L., Musacchio, A., Jappelli, R., Cesareni, G., 1991. Selection of antibody ligands from a large library of oligopeptides expressed on a multivalent exposition vector. J. Mol. Biol. 222 (2), 301–310.

Flanagan, J.A., Huddleston, J.G., Lyddiatt, A., 1991. Applications of aqueous two-phase partition in PEG-phosphate systems. Bioseparation 2 (2), 43–61.

Franco, T.T., Andrews, A.T., Asenjo, J.A., 1996. Use of chemically modified proteins to study the effect of a single protein property on partitioning in aqueous two-phase systems: effect of surface hydrophobicity. Biotechnol. Bioeng. 49, 300–308.

Greenwood, J., Willis, A.E., Perham, R.N., 1991. Multiple display of foreign peptides on a filamentous bacteriophage. Peptides from *Plasmodium falciparum* circumsporozoite protein as antigens. J. Mol. Biol. 220 (4), 821–827.

Hassinen, C., Kohler, K., Veide, A., 1994. Polyethylene glycol-potassium phosphate aqueous two-phase systems. Insertion of short peptide units into a protein and its effects on partitioning. J. Chromatogr. A 668 (1), 121–128.

Iannolo, G., Minenkova, O., Gonfloni, S., Castagnoli, L., Cesareni, G., 1997. Construction, exploitation and evolution of a new peptide library displayed at high density by fusion to the major coat protein of filamentous phage. Biol. Chem. 378 (6), 517–521.

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 (4), 835–844.

Johansson, G., 1974. Effects of salts on the partition of proteins in aqueous polymeric two-phase systems. Acta Chem. Scand. Ser. B B28, 873–882.

Kay, B., Winter, J., McCafferty, J., 1996. Phage Display of Peptides and Proteins: a Laboratory Manual. Academic Press, Inc, San Diego, CA.

Köhler, K., Ljungquist, C., Kondo, A., Veide, A., Nilsson, B., 1991a. Engineering proteins to enhance their partition coefficients in aqueous two-phase systems. Biotechnology (NY) 9 (7), 642–646.

Köhler, K., Veide, A., Enfors, S.-O., 1991b. Partitioning of B-galactosidase fusion proteins in PEG/potassium phosphate aqueous two-phase systems. Enzyme Microb. Technol. 13, 204–209.

Laemmli, U.K., 1970. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 227 (259), 680–685.

Li, M., 2000. Applications of display technology in protein analysis. Nat. Biotechol. 18, 1251–1256.

Luther, J.R., Glatz, C.E., 1994. Genetically engineered charge modifications to enhance protein separation in aqueous two-phase systems: electrochemical partitioning. Biotechnol. Bioeng. 44, 147–153.

Morris, H.R., Paxton, T., Dell, A., Langhorne, J., Berg, M., Bordoli, R.S., Hoyes, J., Bateman, R.H., 1996. Rapid Commun. Mass Spectrom. 10, 889.

Nilsson, B., Moks, T., Jansson, B., Abrahmsen, L., Elmblad, A., Holmgren, E., Henrichson, C., Jones, T.A., Uhlen, M., 1987. A synthetic IgG-binding domain based on staphylococcal protein A. Protein Eng. 1 (2), 107–113.

Nord, K., Gunneriusson, E., Ringdahl, J., Stahl, S., Uhlen, M., Nygren, P.A., 1997. Binding proteins selected from combinatorial libraries of an α-helical bacterial receptor domain. Nat. Biotechnol. 15 (8), 772–777.

Russel, M., 1991. Filamentous phage assembly. Mol. Microbiol. 5 (7), 1607–1613.

Ruther, U., 1982. pUR 250 allows rapid chemical sequencing of both DNA strands of its inserts. Nucl. Acids Res. 10 (19), 5765–5772.

Sambrook, J., Fritsch, E.F., Maniatis, T., 1989. Molecular Cloning: a Laboratory Manual.

Scott, J.K., Smith, G.P., 1990. Searching for peptide ligands with an epitope library. Science 249 (4967), 386–390.

Smith, G.P., Petrenko, V.A., 1997. Phage Display. Chem. Rev. 97, 391–410.

Ståhl, S., Sjolander, A., Nygre, P.A., Berzins, K., Perlmann, P., Uhlen, M., 1989. A dual expression system for the generation, analysis and purification of antibodies to a repeated sequence of the *Plasmodium falciparum* antigen Pf155/RESA. J. Immunol. Methods 124 (1), 43–52.

Veide, A., Smeds, A.L., Enfors, S.O., 1983. A process for large-scale isolation of B-galactosidase from *E. coli* in an aqueous two-phase system. Biotechnol. Bioeng. XXV, 1789–1800.

Walter, H., Johansson, G. 1994. Methods in Enzymology. Aqueous Two-Phase Systems.