Protein engineering with bacterial display
Patrick S Daugherty

Recent improvements in bacterial surface display systems coupled with efficient selection and screening strategies are propelling bacterial display systems to the forefront of peptide and protein engineering. The ability to analyze and screen very large protein libraries using cell-sorting instrumentation coupled with the ease of manipulating bacteria provides new capabilities for the protein engineering toolbox.

Addresses
Department of Chemical Engineering, Biomolecular Science and Engineering, and Institute for Collaborative Biotechnologies, University of California, Santa Barbara, CA 93106, United States

Corresponding author: Daugherty, Patrick S
(psd@engineering.ucsb.edu)

Introduction

The ability to link a protein’s function to the gene encoding that protein using the so-called ‘display technologies’ has become an essential means to identify proteins with desired properties from large libraries and optimize their properties. Among display methods, cell display methods including bacterial and yeast display (reviewed in this volume) offer a ‘window’ into the library screening process through the use of fluorescence-activated cell sorting (FACS). FACS is particularly useful for library screening since it enables real-time visualization and optimization of library screening [1]. Thus, if properties of interest can be linked to fluorescence, they can be measured quantitatively in real time, enabling high-resolution library separations. Quantitative capabilities are proving to be especially useful for screening naïve or random libraries for new functions and for defining proper screening stringency for directed evolution of affinity and specificity [2].

Interestingly, different display systems and even different scaffolds within the same host can yield different binding solutions for a given target with equivalent libraries. For example, bacterial display peptide libraries yielded a diverse group of binding motifs including those obtained from M13 phage libraries, but not those obtained from T7 phage libraries [3]. Similarly, T7 phage display peptide libraries have been shown to yield a variety of different motifs binding to streptavidin, whereas an equivalent M13 phage library yielded only a subset of these ligands [4]. Thus, each display system may offer differing access to sequence space and some binders may be uniquely identified by a given display system.

Bacterial display systems [5] and their use for enzyme library screening have been reviewed recently [6]. Here, recent bacterial display library methodologies and their use for engineering molecular recognition are highlighted.

Bacterial display scaffolds for library screening

The most frequently used host for bacterial display is Gram-negative bacterium *Escherichia coli*, owing to rapid growth rate, ease of genetic and physical manipulation, and its suitability for making large libraries of up to $10^{11}$ members. A wide variety of different scaffolds, or carrier proteins, have been utilized to present peptides and proteins on the outer surface of *E. coli* [5]. Though, only a few scaffolds have been used for library screening (Table 1).

Of course, for ‘surface display’ the scaffold must be capable of transporting the desired passenger protein to the external surface of *E. coli*. The passenger’s size, folding efficiency, and disulfide content can strongly influence its ability to be secreted across the outer membrane and become localized on the cell surface. Unfortunately, differences in the host strain and expression conditions, surface localization methods employed for characterization, and the passengers themselves make quantitative evaluation or comparison of the passenger limitations for each scaffold problematic [7]. Nevertheless, the sizes of passengers displayed using different scaffolds have been reviewed [5]. However, one should use caution in attempting to generalize scaffold capabilities or limitations based on investigations of one or a few passengers. Instead, carefully designed, systematic studies will be needed to define the capabilities and limitations of each display scaffold. For example, in a systematic evaluation of the passenger limitations for the truncated intimin EaeA [8], five passengers ranging in size from 76 to 287 residues were investigated in different strains and with multiple expression conditions. The largest passenger, β-lactamase, was not displayed under any of the conditions used. Though, a 165-residue β-lactamase inhibitor protein with two disulfides was...
displayed with reduced efficiency in a strain deficient in the periplasmic disulfide oxidoreductase DsbA (dsbA−) that promotes disulfide formation strain, where expression was induced in the presence of 20 mM mercaptoethanol at 37 °C [87]. This study illustrates the complexity of defining passenger limitations for individual scaffolds.

To enable effective affinity-based screening against protein targets, the scaffold should be monomeric to reduce the likelihood that multiple receptor–ligand interactions, or avidity effects, obscure the true affinity of a 1:1 stoichiometric complex. Scaffolds should be randomly distributed and spatially separated on the outer membrane of bacteria, to avoid local clusters of receptors that mediate avidity effects. Additionally, despite the cross-linked architecture of the E. coli outer membrane, some outer membrane proteins such as LamB may be capable of lateral diffusion within the membrane [9], leading to avidity effects for multivalent targets. Lastly, high-level expression of the scaffold alone, in the absence of the passenger protein, should impose minimal metabolic burdens to the host cell to prevent undesired clones from overtaking the library.

The surface display of passenger proteins on E. coli can be achieved by genetic fusion with various ‘scaffold’ proteins targeted to the outer membrane as well as those assembled into flagella and fimbrial structures. Bacterial display scaffolds can be broadly grouped into those allowing N-terminal, C-terminal, and insertional fusions (Figure 1). Among these approaches, display of peptides on bacterial flagella using the ‘FliTrx’ system has been used most often given the commercial availability of this system [10] (Table 1). With FliTrx, peptides are presented as constrained insertions within the active site loop of E. coli thioredoxin, which is in turn inserted into a surface-exposed region of the abundant, repeating flagellar protein FliC [11]. Insertion libraries such as FliTrx and those created in outer membrane proteins (e.g. OmpA, OmpC, OmpX, and FhuA) are well suited for mapping antibody and protein-binding epitopes, and selecting initial low affinity binders toward challenging targets for subsequent affinity maturation. Though peptides identified from insertion libraries are typically scaffold-dependent [38].

Display of passenger polypeptides as N-terminal fusions with a surface-exposed N-terminus of the display scaffold can be accomplished via fusion to autotransporter proteins. Autotransporters used for library screening include the IgA protease from Neisseria gonorrhoeae, E. coli AIDA-I [12], or EstA from Pseudomonas aeruginosa [13]. Although autotransporters are thought to translocate unfolded passengers, recent studies suggest that autotransporters can also translocate various folded passengers [7]. Even so, use of a can improve the display of some disulfide-containing proteins [8,14].

Display via the scaffold’s C-terminus may be beneficial to enhance the diversity of peptide libraries since stop codons arising from common randomization schemes and nonintended errors (primer deletions or PCR errors) can yield functional binders without truncating the carrier protein. C-terminal display libraries have been generated and screened using intimins (EaeA), invasins, and the LppOmpA vector. Again, efficient C-terminal display of some proteins via EaeA appears to require maintenance.

Table 1

<table>
<thead>
<tr>
<th>Display scaffold (NCBI Accession #)</th>
<th>Displayed library</th>
<th>Target(s)</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td>OmpA (NP_415477)</td>
<td>Insertional fusion; semirandom 9 a.a. peptide, random 12 a.a./15 a.a. peptide</td>
<td>HIV GP120, CRP, T7 mAb, HSA, streptavidin, tumor cell, red blood cells</td>
<td>[17,35,37,40]</td>
</tr>
<tr>
<td>OmpX (POA917)</td>
<td>Insertional fusion; random 15 a.a. peptide</td>
<td>Monoclonal antibodies</td>
<td>[18]</td>
</tr>
<tr>
<td>FhuA (NP_539207)</td>
<td>Insertional fusion; Staphylococcus aureus gene fragments</td>
<td>Tumor cells</td>
<td>[36]</td>
</tr>
<tr>
<td>FliTrx</td>
<td>Insertional fusion constrained in thioredoxin; 12 a.a. constrained peptide</td>
<td>Monoclonal antibodies, metal oxides, tumor cell lines, amyloid β-peptide, pooled serum IgG</td>
<td>[10,16,19*,38,41,42]</td>
</tr>
<tr>
<td>FimH</td>
<td>Insertional fusion constrained in fimbrial adhesion protein; semirandom 9 a.a. peptide</td>
<td>Zinc ions</td>
<td>[43]</td>
</tr>
<tr>
<td>Autotransporters: AIDA-I (CAA46156), IgA protease (CAAA0270)</td>
<td>N-terminal fusion; random 15 a.a. peptide, semirandom 28 residue cystine knot peptide</td>
<td>Cathepsin G, trypsin</td>
<td>[21,44]</td>
</tr>
<tr>
<td>CPXa (derived from POA917)</td>
<td>N-terminal fusion; 7 a.a. constrained peptide, random 15 a.a. peptide</td>
<td>Streptavidin, tumor cells</td>
<td>[3*,29**,37]</td>
</tr>
<tr>
<td>Lpp (1–9) OmpA (46–159)</td>
<td>C-terminal fusion; antidigoxin scFv antibody libraries</td>
<td>Digoxin-fluorescent conjugate</td>
<td>[45–47]</td>
</tr>
<tr>
<td>Invasin (1–625) (P11922)</td>
<td>C-terminal fusion; random 10 a.a. peptide</td>
<td>Human cell lines</td>
<td>[48]</td>
</tr>
</tbody>
</table>

* Sequences encoding these scaffolds can be identified from the original references given.
of the passenger in an unfolded conformation for export [8*]. The ice nucleation protein (INP) scaffold [15] might also enable screening of C-terminal display libraries for binders, but has been used primarily to screen enzyme libraries.

A scaffold presenting both N-terminal and C-terminal on the cell surface was recently engineered by circular permutation of the smallest member of the outer membrane protein family, OmpX (Figure 1) [3*]. The circularly permuted OmpX (CPX) scaffold enables normalization of protein display levels by fluorescence labeling of a C-terminal affinity tag. Alternatively, the adjacent termini could be used to present heterodimeric proteins. Directed evolution has been used to further improve this scaffold for presenting N-terminal passengers of up to at least 100 amino acids in length (J. Rice & PSD, unpublished).

**Screening for binding ligands**

Affinity-based screening of cell surface display libraries generally requires use of FACS, since use of magnetic selection (MACS) alone or panning processes such as that used with the FLiTrx system lead to avidity interactions that interfere with affinity screening (Figure 2). Similarly, multivalent targets can lead to avidity effects if display levels are high or the target is sufficiently large to cross-link multiple surface displayed ligands. Even when using FACS, screening protocols that rely upon secondary labeling using multivalent antibodies or streptavidin/neutravidin must be designed to allow the affinity of receptor–ligand interaction of interest to be effectively measured. Finally, when screening on the basis of dissociation kinetics using FACS the use of a large excess of a nonfluorescent competitor (i.e. nonbiotinylated target protein) may be necessary to prevent re-binding effects that can enhance apparent affinities measured for cell displayed ligands [3*].

**Antibody epitope mapping**

Determination of an antibody’s binding specificity using peptide libraries, or epitope mapping, has been demonstrated using several bacterial display scaffolds using both selections and screening via FACS. As a means to demonstrate the multispecificity capability of antibodies, FLiTrx has been used to identify peptide mimotopes for an antihapten IgE [16]. Linear peptides derived from screening possessed barely detectable affinity, while peptides constrained within thioredoxin possessed dissociation constants of roughly 10 μM. These low affinities were attributed to the multivalent binding that occurs during the panning process [16]. Sequential MACS and FACS has been used to screen a large library of $5 \times 10^{10}$ random 15-mer insertions into OmpA against the anti-T7 tag antibody [17] yielding a six-residue consensus sequence. Using a similar library constructed as insertions within OmpX, the epitopes of two monoclonal antibodies were mapped by performing two cycles of library enrichment in a dime-sized microfluidic device [18].

Bacterial display was also used recently to identify the dominant specificities of the circulating antibody repertoire, wherein peptides recognizing disease-specific anti-
bodies are isolated. Such serum ‘antibody fingerprinting’ studies can provide insights into mechanisms of pathogenesis, as well as provide reagents that could potentially be used to create improved diagnostic tests. Using the FlitTrx library, alternating selections were performed for peptides that bind to pooled serum antibodies from patients with celiac disease, and not to antibodies from patients without active disease. A peptide ligand was identified with similarity with a rotavirus major coat protein epitope, generating the hypothesis that rotavirus infections could serve as the antigenic trigger for this autoimmune disease [19]. An array of mimotopes generated in a similar fashion using phage was shown to provide enhanced specificity and sensitivity in diagnosing prostate cancer when compared to the use of the prostate-specific antigen (PSA) [20]. The coupling of bacterial display peptide libraries with quantitative screening by FACS will likely provide a means to enhance specificity of ligand panels for the disease state, since specificity can be measured directly using FACS.

**Protein-binding ligands**

The identification of protein-binding ligands is probably one of the most commonly cited uses of phage display. Bacterial display is only beginning to yield high affinity protein-binding ligands from diverse repertoires, an application for which bacterial display may be particularly well suited. Peptides were identified from a large library of $5 \times 10^{16}$ random 15-mer insertions into OmpA binding to several protein targets including human serum albumin, human C-reactive protein, HIV GP120, and streptavidin [17]. Consensus sequences were apparent for all targets investigated.

The autotransporter AIDA-I has also been used recently in the screening of a library of variants of a linear peptide inhibitor of cathepsin G [21]. A library of $1 \times 10^5$ clones was generated by partial randomization of parent peptide, and screened using FACS. Several variants were isolated, though all had reduced inhibition activity when compared to the parental inhibitor. Nevertheless, given the utility of autotransporters in the display of diverse proteins [7,22,23], methodological improvements in library construction will likely yield improved protein variants.

To directly compare loop-constrained and nonconstrained peptide libraries, two large libraries of the form X$_2$CX$_7$CX$_2$ were constructed as insertions into OmpX or N-terminal fusions to the N-terminus of the CPX scaffold. Both libraries were screened using MACS and FACS for streptavidin-binding ligands [3]. N-terminal display enabled improved affinity resolution of clones during screening. Both libraries yielded known streptavidin-binding motifs HPQ/HPM identified using M13 and T7 phage, and mRNA display, libraries [4,24]. However, using bacterial display, additional streptavidin-binding motifs were identified unique to each scaffold. Interestingly, the unconstrained (CPX) library yielded two novel motifs having higher affinities than the best peptides obtained from the constrained library (OmpX) [3]. Both motifs retained high affinity when grafted to the terminus of a monomeric fluorescent protein.

The potential utility the Gram-positive host, *Staphylococcus carnosus* in library screening has been established recently [25,26]. The authors demonstrated that FACS can be used to enrich rare high affinity binders from an excess of weak binders, and that co-display of a second binding domain allows for expression normalization [26]. Despite the low transformation efficiency of this host, optimized protocols [27] should allow for the creation of libraries of $\sim 10^8$ variants. Given their thick cell wall, such hosts may be particularly useful in screens involving harsh selection pressure (e.g. pH, proteases) to increase stringency.

**Peptide substrate identification**

Bacterial display of peptide substrates, analogous to substrate phage [28], using CPX enabled the identification of optimal peptide substrates for proteolytic enzymes using cellular libraries of peptide substrates (CLiPS) [29]. In
this approach, a candidate substrate with flanking flexible linkers is inserted between an N-terminally displayed affinity ligand, and the scaffold. Single-cell conversion measurements via FACS enable screening on the basis of cleavage kinetics and direct ranking of identified substrates on the cell surface. Proteolytic substrates were identified and sorted by means of their absence of fluorescence after incubation with enteropeptidase. Screening yielded substrates with ~10-fold greater activity than the substrate DDDDK frequently used in protein purification.

Antibody library screening

While there is substantial interest in using bacterial display to screen antibody libraries, screening for protein-specific antibody fragments has proven challenging. To enable screening of anti-protein antibodies, a system was developed allowing for anchored periplasmic expression (APEX) [30,31]. Single-chain antibodies were targeted into the periplasm via the secretory pathway, and anchored to the periplasmic face of the inner membrane via a fatty acylated lipoprotein sequence. Removal of the outer membrane using lysozyme enabled labeling with fluorescently tagged antigen and screening with FACS. Antibody encoding genes were recovered from the sorted pool via PCR, subcloned into the expression vector, and the process was repeated for enrichment. The APEX system can also be used to express target antigens directly in conjunction with the library [32]. A similar approach enabled screening of libraries of whole IgG antibodies anchored, via a membrane tethered Fc binding protein, in the periplasm of E. coli [33**].

Continued evolution of screening strategies will foster new applications for cell display methods. The use of multiple fluorescence parameters during screening by FACS can allow for screening for peptides and antibodies with well-defined specificity or crossreactivity. For example, cell display in conjunction with two-color FACS enabled the evolution of antibodies that crossreact with two different botulinum neurotoxin subtypes in a yeast system [34**]. Multiparameter screening strategies could be used to obtain binding to multiple targets for various therapeutic applications.

Cell-targeting peptides

Bacterial display systems offer the prospect of simple means to generate cell-specific affinity ligands. Several different bacterial display systems (Invasin, FhuA, OmpA, FliTrx, and CPX) have enabled the identification of peptide ligands that bind or mediate bacterial entry into a variety of mammalian cell types [35–37]. Cell display enables recovery of binding clones from the target cells by regrowth, and use of FACS to screen for target cell binding using intrinsically fluorescent display libraries [37].

Most studies aiming to identify cell-targeting peptides have relied upon selection by ‘panning’ of bacterial libraries on whole cells. For example, the FliTrx system was used to identify peptides specific for a prostate carcinoma cell line [38]. Using four rounds of selection by panning on adherent cells followed by binder recovery by vortexing resulting in the identification of five unique clones. These peptides shared a consensus sequence of up to eight amino acid residues, suggesting that they targeted a single receptor. One synthetically prepared peptide retained binding activity in a cell culture assay, and was stable in serum for 24 hours.

To enable quantitative bacterial display library screening via FACS for cell-specific ligands, a bicistronic display system was developed encoding a random peptide library display at the N-terminus of CPX with the second cistron encoding a fluorescent protein [37]. Thus, binding of the fluorescent, peptide-displaying bacterial cells to the target tumor cells enabled quantification of binding and clone isolation using FACS. Using this system, a library of $4 \times 10^9$ members was screened by FACS after enrichment by panning. Identified ligands mediated 80-fold increased bacterial cell binding to the target breast tumor cells, relative to ‘normal’ cell lines. Similarly, a fluorescent bacterial display peptide library was used to identify peptides that mediate binding to human and rat erythrocytes [35]. Peptide-functionalized particles strongly adhered to erythrocytes.

The use of bacterial display provides a convenient means to identify ligands that mediate entry or invasion of mammalian cells [36]. Such peptides may be useful for the delivery of genes, proteins, and particles into mammalian cells. A gentamycin protection assay (GPA) was used to identify clones from a library of insertions into FhuA that invaded the target cells. Both RGD-containing and novel peptides were obtained mediating a roughly 100-fold increase in bacterial invasion into human cervical cancer cells, as measured by the GPA. While soluble versions of the peptides identified using bacterial display typically have low monovalent affinity (10–100 µM) [37], it may be possible to improve ligand affinity by decreasing surface display levels or using more stringent washing.

Conclusions

Bacterial display methodologies are poised to make important contributions in peptide and polypeptide engineering. Combining the simplicity of bacterial cell manipulation, with large libraries, and quantitative screening via FACS, these approaches have potential to streamline the identification and optimization of binding ligands. Bacterial display in conjunction with multi-parameter FACS should be particularly useful for the screening and optimization of a variety of emerging nonantibody therapeutics derived from small scaffolds [39].

References and recommended reading

Papers of particular interest, published within the annual period of review, have been highlighted as:

• of special interest
•• of outstanding interest


A biterminal bacterial display scaffold was generated by circular permutation of outer membrane protein X to result in both C-terminal and N-terminal exposed on the cell surface. Using this scaffold for N-terminal display and the parental OmpX for insertion display, equivalent libraries were screened against streptavidin for comparison.


Comparison of the cell surface exposure of different passenger proteins established an inverse correlation between the propensity to form a stable folded structure and the ability to translocate across the outer membrane and be displayed.


The FLTRX peptide display was used to identify peptides binding to serum antibodies from patients with celiac disease. A peptide with similarity to rotavirus coat protein was identified.


Phage display was used to select peptides reacting to serum antibodies from patients with prostate cancer. Construction of an array using whole phage, and training via pattern recognition, enabled prostate cancer diagnosis with greater sensitivity and specificity than by using prostate-specific antigen (PSA).


Display of peptide substrates and an N-terminal affinity ligand enabled screening of substrate libraries via FACS to identify rapidly cleaved substrates for two proteases. The approach provided a means to identify optimal protease substrates on the basis of quantitative cleavage kinetics.


Using the APEx system, full-length IgG antibodies were assembled in the bacterial periplasm and anchored to the inner membrane enabling screening of libraries via FACS. The system provides a direct route to generate full-length antibodies from large libraries.


Yeast cell display and two-color FACS was used to evolve antibodies that crossreact with two different botulinum neurotoxins. The evolved antibody fragments exhibited similar affinities for two different neurotoxin subtypes.


