# M13 Bacteriophage Display Framework That Allows Sortase-Mediated Modification of Surface-Accessible Phage Proteins

The MIT Faculty has made this article openly available. **Please share** how this access benefits you. Your story matters.

<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>As Published</td>
<td><a href="http://dx.doi.org/10.1021/bc300130z">http://dx.doi.org/10.1021/bc300130z</a></td>
</tr>
<tr>
<td>Publisher</td>
<td>American Chemical Society (ACS)</td>
</tr>
<tr>
<td>Version</td>
<td>Author's final manuscript</td>
</tr>
<tr>
<td>Accessed</td>
<td>Thu Mar 28 02:51:51 EDT 2019</td>
</tr>
<tr>
<td>Citable Link</td>
<td><a href="http://hdl.handle.net/1721.1/80400">http://hdl.handle.net/1721.1/80400</a></td>
</tr>
<tr>
<td>Terms of Use</td>
<td>Article is made available in accordance with the publisher's policy and may be subject to US copyright law. Please refer to the publisher's site for terms of use.</td>
</tr>
<tr>
<td>Detailed Terms</td>
<td></td>
</tr>
</tbody>
</table>
An M13 bacteriophage display framework that allows sortase-mediated modification of surface-accessible phage proteins

Gaelen T. Hess†, Juan J. Cagnolini§, Maximilian W. Popp§, Mark A. Allen*,‖, Stephanie K. Dougan§, Eric Spooner§, Hidde L. Ploegh§, †, *, Angela M. Belcher†, ‡, #, *, Carla P. Guimaraes§, *

†The David H. Koch Institute for Integrative Cancer Research, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
‡Biophysics Program, Harvard University, Cambridge, MA 02138, USA
§Whitehead Institute for Biomedical Research, Cambridge, MA 02142
‖Department of Materials Science and Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
†Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139
#Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

*Corresponding authors:

Hidde L. Ploegh / Carla P. Guimaraes
Whitehead Institute for Biomedical Research
9 Cambridge Center
Cambridge, MA 02142
Phone: (617) 324-2031
Fax: (617) 452-3566
e-mail: ploegh@wi.mit.edu / guimarae@wi.mit.edu

Angela M. Belcher
Massachusetts Institute of Technology
77 Massachusetts Avenue, Room 76-561
Cambridge, MA 02139
(617) 252 1163
e-mail: belcher@mit.edu
Abstract

We exploit bacterial sortases to attach a variety of moieties to the capsid proteins of M13 bacteriophage. We show that pIII, pIX, and pVIII can be functionalized with entities ranging from small molecules (e.g., fluorophores, biotin) to correctly folded proteins (e.g., GFP, antibodies, streptavidin) in a site-specific manner, and with yields that surpass those of any reported using phage display technology. A case in point is modification of pVIII. While a phage vector limits the size of the insert into pVIII to a few amino acids, a phagemid system limits the number of copies actually displayed at the surface of M13. Using sortase-based reactions, a 100-fold increase in the efficiency of display of GFP onto pVIII is achieved. Taking advantage of orthogonal sortases, we can simultaneously target two distinct capsid proteins in the same phage particle and maintain excellent specificity of labeling. As demonstrated in this work, this is a simple and effective method for creating a variety of structures, thus expanding the use of M13 for materials science applications and as a biological tool.
**Introduction**

M13 bacteriophage has a cylindrical shape with a length of 880nm and a diameter of 6nm. It encapsulates a single-strand genome that encodes five different capsid proteins (Fig. 1A). The body of the phage is composed of 2700 copies of pVIII, the major capsid protein. At one end of the virus, there are ~5 copies of both pIII and pVI proteins, and at the other end there are ~5 copies of both pVII and pIX proteins.

The capsid proteins of M13 bacteriophage have been used to express combinatorial peptide libraries or protein variants (ranging from single domains to antibodies) to screen for target ligands in a process known as phage display. This technique has enabled not only identification of peptides with affinity for biological targets such as proteins, cells, and tissues, but also allowed the identification of biomolecules that bind inorganics. These molecules, when expressed on the M13 capsid proteins, can serve as scaffolds for nanowires, structures, and devices. Functionalization of a virion capsid such as M13 is currently accomplished using chemical and/or genetic approaches. However both strategies have limitations. Chemical conjugations are convenient and versatile, but they label motifs found on multiple M13 capsid proteins and oftentimes require non-physiological pH and reducing conditions that compromise the activity of the molecule that is being attached or of the moieties already displayed on other capsid proteins.

Genetic engineering of phage allows the encoded protein/peptide to be displayed precisely, but it has intrinsic restrictions. Two classes of vectors are available for genetic phage display: phagemid and phage. A phagemid allows expression of large fusions with any of
the five M13 phage capsid proteins, but these fusions are incorporated at low efficiency\textsuperscript{17-21}. In a phage vector, the M13 bacteriophage genome is modified directly. As a result, every copy of the recombinant capsid protein incorporated into the virus displays the modified protein. However, this strategy does not support display of large moieties\textsuperscript{22-24}. pVIII allows the display of a larger number of recombinant molecules per phage particle, but it also has the strictest size limitation in phage vector display. pVIII peptide libraries are mostly limited to sizes of up to 10 amino acids, as phage with longer insertions rarely assemble\textsuperscript{25-26}. Insertions of 6-20 amino acids onto pVIII are possible using phagemid, but their display is inefficient with less than 25\% of the copies of pVIII containing the desired fusion product\textsuperscript{20}. Incorporation of proteins is even less efficient on pVIII: a 23kDa protein is displayed, on average, on less than a single copy of the pVIII fusion per phage particle using a phagemid vector\textsuperscript{18}. Phage display methods on the pVIII have been able to increase the binding affinity of phage displaying a moiety\textsuperscript{23}, but the displayed copy number of the moiety has not been determined. Large moieties of at least 23kDa have been genetically fused to all four minor capsid proteins using a phagemid vector\textsuperscript{22, 27-28}, but only pIII has been extensively used in the phage vector system\textsuperscript{39}. However, viability of the resultant phage fusions does not guarantee that the recombinant peptide/protein of interest displays its native structure. Both the environment where phage assembles and the phage coat protein to which the protein of interest is fused may interfere with proper folding\textsuperscript{30}. This is particularly critical for enzymes and antibodies, as they might not be functional when incorporated into the phage structure.

To expand the versatility of M13 as a display platform, we devised a strategy based on sortase-mediated chemo-enzymatic reactions to covalently attach a variety of moieties to the N-
terminus of pIII, pIX, and pVIII. We obtained a vast improvement over the published data in the copy number of displayed peptides and proteins, particularly on pVIII.

Sortase A enzymes allow modification of proteins by enzymatic ligation with a wide range of functional groups (including biotin, fluorophores, and other proteins) at the C-terminus, N-terminus, or at both termini of the protein of interest\textsuperscript{31-35}. The widely used sortase A from \textit{Staphylococcus aureus} (Srt\textsubscript{A} aureus) recognizes substrates that contain an LPXTG sequence\textsuperscript{36-38}, whereas sortase A from \textit{Streptococcus pyogenes} (Srt\textsubscript{A} pyogenes) recognizes substrates with an LPXTA motif\textsuperscript{33,39}. The sortase enzymes cleave between the threonine and glycine or alanine residue, respectively, to yield a covalent acyl-enzyme intermediate that is resolved by nucleophilic attack of a suitably exposed amine, namely oligoglycine or oligoalanine-containing peptides\textsuperscript{39} in the case of Srt\textsubscript{A} aureus or Srt\textsubscript{A} pyogenes, respectively (Fig. 1B). Using these two orthogonal sortase A enzymes\textsuperscript{33}, we established conditions to site-specifically attach two different moieties onto two capsid proteins in a single phage particle.

The sortase labeling method has several advantages over genetic and chemical methods. First, the reaction is site-specific, as none of the coat proteins naturally display the required motifs to participate in sortase-mediated reactions. Second, these motifs are small and, therefore, can be easily inserted into the phage genome, maximizing the number of potential attachment sites. Third, a protein to be displayed on phage by means of sortase can be properly folded separate from the assembly of phage. The site-specific nature of the reaction fixes the orientation of the displayed protein. Fourth, the reactions are performed under physiological conditions. Fifth, sortase reactions afford attachment of a wide range of molecules, including those that cannot be genetically encoded such as fluorophores and biotin.
This work establishes the component parts to build phage structures that have new material and biological applications. We provide two examples: the creation of a new lampbrush structure by fusing different phage particles through pIII/pVIII and a fluorescently labeled phage containing a cell-targeting moiety to stain and to FACS sort cells.

**Experimental Procedures**

**Generation of the M13 phage constructs**

The oligonucleotides used to design the different phage constructs are compiled in Table S1. The G5-pIII phage was engineered by inserting the G5pIIIC and G5pIIINC annealed oligonucleotides into the M13KE vector (New England Biolabs), previously digested with EagI and Acc65I restriction enzymes. To construct the A2G4-pVIII phage, the M13SK vector was digested with PstI and BamHI restriction enzymes and the A2G4pVIIIC and A2G4pVIIINC annealed oligonucleotides were inserted. To engineer the G5HA-pIX construct, the 983 vector was used. This vector was created by refactoring the M13SK vector so the pIX and pVII genes are not overlapping. Upon digestion of this vector with SfiI, the annealed G5HApIXC and G5HApIXNC oligonucleotides were inserted. The G5-pIII-A2-pVIII phage construct was created using a modified M13SK vector\textsuperscript{40}, which has a DSPHTEL sequence on pVIII and a biotin acceptor peptide (GLQDIFEAQKIEWHE) on pIII. Five N-terminal glycines were added to pIII following the above strategy described for G5-pIII phage. The resultant vector was then modified at the N-terminus of pVIII using the QuikChange II site-directed mutagenesis kit (Stratagene) and the pVIIAADSPH oligonucleotide pair. All the generated phage vectors were transformed into the XL-1 Blue bacterial strain, plated in agar top on LB agar plates containing
1mM IPTG, 40µg/mL X-Gal, and 30µg/mL tetracycline. Plaques were selected and DNA was isolated and sequenced to check for the insertion.

For phage amplification, the E. coli strain ER2738 (New England Biolabs) in LB media supplemented with 30µg/mL tetracycline, was infected with phage for at least 12hrs at 37°C. The cultures were centrifuged at 12000g for 20min and the phage was precipitated from the supernatant at 4°C with the addition of 1/5 of the supernatant volume of 20% PEG8000/2.5M NaCl solution. Upon centrifugation at 13500g for 20min, the pellet was resuspended in 25mM Tris, 150mM NaCl, pH 7.0-7.4 (TBS). For further purification, this resuspension was subjected to two rounds of centrifugation/precipitation. The final phage concentration averaged between 10^{13}-10^{14} plaque forming units (pfu) per mL as determined by UV-vis spectrometry^{41}.

**Sortase-mediated reactions**

SrtA_{pyogenes} and SrtA_{aureus} were expressed and purified as described^{33,42}. Sortase reactions were performed as indicated in the figures. A typical sortase reaction with SrtA_{aureus} included 200nM phage, 50µM SrtA_{aureus}, and 50µM substrate for small peptides or 20µM for proteins. The reactions were incubated for 3hrs at 37°C (for small peptides) or at room temperature (for proteins) in TBS with 10mM CaCl₂. SrtA_{pyogenes}-mediated reactions included 8nM phage, 50µM SrtA_{pyogenes}, and 20µM substrate, incubated for 3hr at 37°C in TBS. Where indicated, phage was purified by PEG 8000/NaCl precipitation after diluting the reactions with TBS such that the substrate concentration was below 600nM.

For the flow cytometry experiments, the G₅-pIII-A₂-pVIII phage construct was labeled with K(TAMRA)-LPETAA on pVIII. The resultant labeled phage was purified by PEG8000/NaCl precipitation, resuspended in TBS, and split into three parts. One part remained
unlabeled, and the other two were labeled with either VHH7.LPETG or anti-GFP.LPETG on pIII. As assessed by the anti-pIII antibody, a yield of 2.5 antibody molecules per virion was achieved in both cases.

The yield of the sortase-mediated biotinylation reactions was determined using biotinylated GFP as a standard. This was prepared labeling GFP - comprising a LPETG at its C-terminus - with a biotin group using SrtA\textsubscript{aureus} (GFP.LPETGGGK(biotin))\textsuperscript{42}. Known amounts of the purified GFP.LPETGGGK(biotin) standard and varying volumes of the phage labeling reactions were loaded onto the same SDS-PAGE gel and analyzed by immunoblot using streptavidin-HRP (GE Healthcare). The signal obtained in the phage labeling reactions was compared with the signal derived from the GFP.LPETGGGK(biotin) calibration curve allowing us to infer the amount of phage protein labeled in the reaction. To calculate the labeling efficiency, the amount of labeled protein was divided by the amount of total phage protein loaded into the gel. The phage concentration was determined by UV-vis spectrometry and it was assumed that there were 2700 copies of pVIII, 5 copies of pIII, and 5 copies of pIX per phage particle.

To determine the yield of GFP-pVIII phage labeling, unincorporated GFP and sortase was removed from phage by PEG8000/NaCl precipitation. Varying volumes of GFP-pVIII phage and known amounts of GFP were loaded onto the same SDS-PAGE gel and analyzed by immunoblot using an anti-GFP-HRP antibody (Santa Cruz Biotechnology). The signal of the GFP-pVIII fusion protein was compared to the signal of the GFP calibration curve as described for the biotinylation reactions. For GFP-pIII and GFP-pIX labeling, the signal of the fusion protein was compared to the input amount of pIII or pIX as detected by anti-pIII (New England Biolabs) or anti-HA (Roche) antibodies, respectively. For GFP-pIII, the input signal consisted of
only intact pIII molecules and lower molecular weight anti-pIII reactive proteins were not included. These proteins can be attributed to proteolyzed pIII. Because the anti-pIII antibody recognizes the C-terminus of the protein, these fragments cannot be labeled using SrtA_sureus. In all cases the blots were scanned and densitometric analysis was performed using the ImageJ program (National Institutes of Health). The labeling yield was averaged over three independent reactions with three aliquots from each reaction analyzed. The standard deviation of the reactions was calculated from the averages of the three independent reactions.

**Dynamic light scattering (DLS)**

DLS measurements were obtained with a Beckman Delsa-Nano C Particle Analyzer (Beckman Coulter Inc). Phage mixtures were diluted to \( \sim 10^{11} \text{pfu/mL} \) in 1mL of water and loaded into a cuvette. Samples from each experiment were measured in triplicate and the results were averaged by cumulant analysis. Autocorrelation functions were used as a direct comparison of aggregation because aggregates have a slower Brownian motion causing the signal correlation to be delayed to longer relaxation times.

**Atomic force microscopy (AFM)**

Phage preparations were diluted to a concentration of \( \sim 10^{11} \text{ pfu/mL} \), and 100µL of this mixture were deposited on a freshly cleaved mica disc. AFM images were taken on a Nanoscope IV (Digital Instruments) in air using tapping mode. The tips had spring constants of 20-100N/m driven near their resonant frequency of 200-400kHz (MikroMasch). Scan rates were approximately 1Hz. Images were leveled using a first-order plane fit to remove sample tilt.

**Flow cytometry analysis**
C57BL/6 mice were purchased from Jackson Labs. Animals were housed at the Whitehead Institute for Biomedical Research and were maintained according to guidelines approved by the Massachusetts Institute of Technology (MIT) Committee on Animal Care.

Lymph nodes were isolated from 6-8 week old C57BL/6 mice and crushed through a 40μM cell strainer. Cells were washed once with PBS, resuspended at 2x10^7 cells per mL, aliquoted at ~1x10^6 cells per sample, and incubated with staining agents in 5% milk in PBS for 1hr at room temperature. 10^{11} VHH7 molecules and 10^{11} anti-GFP molecules either directly conjugated to TAMRA using SrtA_{auraeus}, or covalently attached to phage (5x10^{10} phage particles of VHH7-G_5-pIII-TAMRA-A_2-pVIII or anti-GFP-G_5-pIII-TAMRA-A_2-pVIII, see Sortase-mediated reactions section) were incubated with the cells. The same amount of non-targeted fluorescent phage particles (i.e., G_5-pIII-TAMRA-A_2-pVIII) was used as a negative control. B cells were stained with Pacific Blue anti-mouse B220 (BD Pharmingen, clone RA3-6B2). Upon staining, the cells were centrifuged at 170g for 5min, washed with PBS three times, and resuspended in 500μL of PBS. Flow cytometry was performed using a FACSARia (BD). 100,000 events were collected for each sample.

Miscellaneous

Expression and purification of GFP.LPETG.His_6 and GFP.LPETA.His_6, were performed as described^{33}. Identification, characterization, expression, and purification of VHH7.LPETG.His_6 will be published elsewhere. Streptavidin was cloned as a streptavidin.LPETG.HAtag.His_6 fusion protein using the template Addgene 20860^{44}, and expressed as a soluble tetrameric streptavidin^{45}. Purification was performed following the same protocol used for GFP^{33}. Sortase
reactions were analyzed on 4-12% Bis-Tris SDS-PAGE gels with MES running buffer except for Fig. S4 which was analyzed on a 12% Laemmli SDS-PAGE gel.

The K(biotin)-LPETGG, K(biotin)-LPETAA, K(TAMRA)-LPETAA, and GGGK(biotin) peptides were obtained from the Swanson Biotechnology Center. For mass-spectrometry, the protein bands of interest were excised, subjected to protease digestion, and analyzed by electrospray ionization tandem mass-spectrometry (MS/MS). Fluorescent gel images were obtained using a variable mode imager (Typhoon 9200; GE Healthcare).

Results

N-terminal labeling of pIII using SrtA<sub>aureus</sub>

pIII has been the most extensively explored of the M13 capsid proteins in phage display because of the flexibility and accessibility of its N-terminus<sup>46</sup>. Thus, we introduced five glycines at the N-terminus of pIII (G<sub>5</sub>-pIII phage) and used SrtA<sub>aureus</sub> to covalently attach a K(biotin)-LPETGG peptide (Fig. 2A). The biotin moiety allowed us to monitor the reaction by immunoblot analysis using streptavidin-HRP. Only when sortase, G<sub>5</sub>-pIII phage, and the peptide are incubated together did we detect a 55kDa streptavidin and anti-pIII reactive protein band (Fig. 2A). The reaction was specific: no other phage proteins were biotinylated. After 3hrs at 37°C, we achieved a yield of 68±9% labeling using 50µM peptide, 50µM SrtA<sub>aureus</sub>, 200nM G<sub>5</sub>-pIII phage, and 10mM CaCl<sub>2</sub>. The efficiency of the reaction was calculated using densitometric analysis of immunoblots where we compared the signal of the biotinylated pIII to biotinylated GFP standards of known concentration. The amount of biotinylated pIII was then divided by the amount of pIII molecules loaded onto the gel, as determined by UV-vis spectrometry. The quantification was repeated for three independent reactions with three samples analyzed for each
reaction. The method of quantification is described in further detail in the Experimental Procedures section.

To determine whether sortase could be exploited to attach pre-folded proteins onto pIII, we used GFP containing an LPETG motif at its C-terminus as a substrate. The reaction was analyzed by immunoblot using an anti-pIII antibody (Fig. 2B). Upon completion of the reaction, a mobility shift of pIII to the ~80kDa region, corresponding to the GFP-pIII fusion product, was detected. The identity of this material was confirmed by mass-spectrometry (Fig. 2B and Fig. S1). After 3hrs at room temperature, we achieved a yield of 56±2% labeling using 20µM GFP-LPETG, 50µM SrtA\textsubscript{aureus}, 200nM G\textsubscript{5}-pIII phage, and 10mM CaCl\textsubscript{2}. The reaction was quantified by densitometry comparing the signal of pIII-GFP to the signal of the intact pIII input loaded into the reaction.

**N-terminal labeling of pIX using SrtA\textsubscript{aureus}**

Because the C-terminus of pIX is buried in the phage structure and therefore unavailable for labeling\textsuperscript{47}, we attempted to label its N-terminus. However, this region of the protein is not as accessible as in pIII and our first attempts at labeling a phage construct displaying five glycines at the N-terminus of pIX using sortase failed (data not shown). To increase accessibility of the five glycines, the N-terminus of pIX was extended with an HA tag, a useful handle for detection, as no pIX-specific antibodies are available. This G\textsubscript{5}HA-pIX phage construct was labeled with the K(biotin)-LPETGG peptide and the reactions were analyzed by immunoblot using streptavidin-HRP and an anti-HA antibody. A 5kDa polypeptide, reactive with both streptavidin and anti-HA, was seen only in the complete reaction (Fig. 3A). We achieved a yield of 73±2% using 50µM peptide, 50µM SrtA\textsubscript{aureus}, 200nM G\textsubscript{5}HA-pIX phage, and 10mM CaCl\textsubscript{2} upon
incubation at 37°C for 3hrs. A similar efficiency was attained when attaching GFP to pIX: 74±1% of pIX was labeled when 20µM GFP-LPETG, 50µM SrtAaureus, 200nM G5HA-pIX phage, and 10mM CaCl₂ were incubated for 3hrs at room temperature. A 35kDa anti-HA reactive polypeptide - consistent with the molecular mass of the GFP-pIX fusion protein - was detected only in the complete reaction and its identity was confirmed by mass-spectrometry (Fig. 3B and Fig. S2).

**N-terminal Labeling of pVIII using SrtApseudomonas**

In the course of phage biogenesis the N-terminus of pVIII is proteolytically cleaved, resulting in the display of an N-terminal alanine. We took advantage of this feature and exploited SrtApseudomonas to label pVIII. Also, the ability of using two orthogonal sortase enzymes (SrtApseudomonas for pVIII and SrtAaureus for pIII and pIX labeling) would further enable dual labeling of the same phage particle.

To be used as a nucleophile in SrtApseudomonas-mediated reactions, pVIII requires display of two N-terminal alanines. Thus, the N-terminus of the mature form of pVIII was modified to AAGGGG (A₂G₄-pVIII phage). The glycines were introduced to extend the N-terminus of pVIII away from the body of the phage, thus improving the accessibility of the Ala-Ala motif for participation in the sortase reaction. Using SrtApseudomonas and a K(biotin)-LPETAA substrate peptide, we showed robust labeling of pVIII based on an immunoblot using streptavidin-HRP (Fig. 4A). Only when A₂G₄-pVIII phage, SrtApseudomonas, and the peptide were mixed together did we detect a biotinylated 10kDa protein, consistent with the size of pVIII. The labeling reaction was site-specific as no other proteins can be detected in the blot. We obtained a yield of 50±3% labeled pVIII when reactions were performed at 37°C for 3hrs with 20µM peptide, 50µM
SrtA<sub>pyogenes</sub>, and 8nM A<sub>2</sub>G<sub>4</sub>-pVIII phage. This translated to 1350±90 biotin molecules on average per phage particle.

Phage assembly limits either the size of the modifications displayed on pVIII to a few residues when using a phage vector, or it limits the number of labels attached to pVIII when using a phagemid vector<sup>20</sup>. In this context, the sortase-labeling strategy is an obvious alternative to overcome such limitations. Using 20μM GFP containing a LPETA motif at its C-terminus, 50μM Srt<sub>pyogenes</sub>, and 8nM A<sub>2</sub>G<sub>4</sub>-pVIII phage, we were able to attach 91±20 GFP molecules on average per phage particle upon incubation at 37°C for 3hrs (Fig. 4B). The identity of the 35kDa anti-GFP reactive protein, consistent with the size of a GFP-pVIII fusion protein, was confirmed by mass-spectrometry (Fig. 4B and Fig. S3). As estimated by nearest neighbor packing (described in Supplementary Information), a single virion can accommodate 385 copies of GFP on its surface. Thus, using the sortase-mediated reaction, we obtained a yield of ~25% of estimated maximum packing.

**Building end-to-body phage structures**

The ability to site-specifically label the M13 capsid proteins provides the opportunity to create novel multi-phage structures, which may provide scaffolds for new materials and devices. One such structure (Fig. 5A) relies on tight binding of the ends of several phage particles (via either pIII or pIX) to the body of another single phage (onto pVIII). However, direct covalent attachment between two phage proteins is not possible using sortase as we were unable to label the C-terminus of pIII, pIX, or pVIII (data not shown). This issue was solved by attaching streptavidin to pIII, biotin to pVIII, and then mixing the two preparations.
Streptavidin, modified to contain a C-terminal LPETG motif in each of its monomers, was attached to the G5-pIII phage using SrtA\textsubscript{aureus}. The samples were boiled, loaded onto an SDS-PAGE gel, and analyzed by immunoblot using an anti-pIII antibody. A 90kDa polypeptide, consistent with the size of pIII fused to a streptavidin monomer, was seen only when all the reaction components were mixed together (Fig. S4). The streptavidin-pIII phage was purified from sortase and free streptavidin by PEG/NaCl precipitation. Dynamic light scattering (DLS) was performed in order to monitor dispersity and aggregation. The normalized autocorrelation function (ACF) of streptavidin-pIII phage showed an exponential decay consistent with monodisperse populations (Fig. 5B). This was confirmed by atomic force microscopy (AFM) that showed individual virions, indicating that only a single phage particle was attached per streptavidin tetramer (Fig. S5). Biotin was conjugated to pVIII using the K(biotin)-LPETAA peptide and SrtA\textsubscript{pyogenes} as described above. The biotinylated phage was purified by PEG/NaCl precipitation to remove free peptide and the sortase-acyl intermediate. The biotinylated phage was observed as individual phage particles by AFM and the ACF showed an exponential decay, again indicating a monodisperse population (Fig. 5B and Fig. S5).

The streptavidin-pIII phage and the biotin-pVIII phage were mixed at a 5:1 molar ratio and incubated at room temperature for 15min. Analysis of these samples by DLS showed an increase of the hydrodynamic diameter for the lampbrush phage mixture (700nm) when compared to streptavidin-pIII (516nm) and biotin-pVIII (204nm) phage preparations. When the two types of phage were mixed, the ACF (Fig. 5B) shows a rising shoulder at longer relaxation times, indicating a polydisperse population. The longer relaxation times observed in the shoulder represent structures larger than single phage. These larger structures were observed by AFM (Fig. 5C and Fig. S5). Linkages between the end of one phage and the body of another
phage were observed when streptavidin-pIII and biotin-pVIII are mixed. These linkages were not detected when the individual phages were visualized by AFM (Fig. S5).

**Site-specific labeling of two capsid proteins in the same phage particle**

The two orthogonal sortases used to label different capsid proteins offer the possibility to attach different moieties to the body (using SrtA<sub>pyogenes</sub>) and to the end of phage (using SrtA<sub>aureus</sub>) within the same virion. In such a strategy, either pIII or pIX could be labeled with SrtA<sub>aureus</sub> orthogonally to the pVIII, so as a proof-of-concept, a phage variant that contains a double alanine at the N-terminus of pVIII and the pentaglycine motif at the N-terminus of pIII was generated (this construct is referred to as G<sub>5</sub>-pIII-A<sub>2</sub>-pVIII). Conditions were optimized to label each of these proteins in a site-specific manner. Because such dual-labeled phage could be a useful tool to sort cells by FACS (see below and discussion section), we here provide the proof-of-concept by labeling the body of phage with a fluorophore and the tip of phage with a cell-targeting moiety.

pVIII was labeled with a K(TAMRA)-LPETAA peptide and purified using PEG/NaCl precipitation to remove free peptide and sortase (Fig. 6A). A fluorescent 10kDa protein, corresponding to pVIII, was the only polypeptide detected in the complete reaction. This confirmed successful labeling and site-specificity of SrtA<sub>pyogenes</sub>. The pIII of this fluorescent phage was then incubated with SrtA<sub>aureus</sub> and a 15kDa single domain antibody, VHH7, modified with a C-terminal LPETG motif. VHH7 recognizes murine Class II MHC products (the development and expression of VHH7 will be described elsewhere). Attachment of VHH7 to pIII was monitored by immunoblot using an anti-pIII antibody (Fig. 6B). Comparing the signal intensities of VHH7-pIII 90kDa polypeptide and of pIII, we estimated that on average 2-3 VHH7
molecules are attached per phage particle, a number similar to what can be obtained when screening phagemid libraries of pIII fusions by panning\textsuperscript{48-49}.

**Flow cytometry experiments using fluorescent phage**

Fluorescent phage has been used for targeted staining \textit{in vivo}\textsuperscript{50-51} as well as flow cytometry experiments\textsuperscript{52}. However, these have been performed with short peptide phage display libraries. The ability to label phage with a large number of fluorophores that are site-specifically attached to pVIII is a tool useful for selecting phage of interest from phage display libraries of large moieties (such as antibodies) by fluorescence. With libraries of this type, less specific labeling methods can alter the displayed moiety. To provide proof-of-concept that fluorescent phage can be used for this purpose, we tested the ability of the dual labeled phage - containing TAMRA fluorophore sortagged onto pVIII and VHH7 onto pIII - to stain B cells. As a negative control, we used a fluorescent phage containing an anti-GFP VHH attached to pIII\textsuperscript{53}. An average yield of 2.5 antibodies per phage virion was achieved for both VHH7 and anti-GFP VHH as determined by densitometric analysis.

Mouse lymphocytes obtained from lymph nodes were stained for B cells using a fluorescent Pacific Blue anti-mouse B220 antibody and incubated with phage-VHH7, phage-anti-GFP, or non-targeted phage. All phage preparations were similarly labeled with TAMRA on pVIII. After removal of unbound materials by washing, cells were subjected to flow cytometry (Fig. 6C). When stained with phage-VHH7, we detected an increase in cells double positive for TAMRA and the B cell marker compared to non-specific staining with phage-anti-GFP or non-targeted phage. Staining of cells with phage-VHH7 was vastly superior to VHH7
directly conjugated to TAMRA, as only a few double positive cells were detected when incubated with an equivalent amount of the latter (Fig. 6C).

**Discussion**

We show that sortase-mediated reactions overcome many of the limitations of current methods to functionalize M13 capsid proteins. The main body and both ends of the viral capsid can be functionalized with substituents that cannot be encoded genetically (such as biotin and fluorophores), and we can also install properly folded and assembled proteins (such as GFP and streptavidin) in a manner that could easily be extended to oligomeric proteins as well.

One of the major challenges has been the modification of the major capsid protein pVIII. Using sortase, labeling efficiencies were greater than those obtained genetically (Table 1). In the past, biotinylated phage has been produced by display of the biotin acceptor peptide (BAP)\(^{54}\), a 15-amino acid sequence. Peptides similar in size have been displayed at no more than 400-700 copies per phage, with the efficiency being sequence-dependent\(^{20}\). Here we attach 1350 biotin molecules on average per phage particle, a great improvement in the display of a small molecule. Moreover, because the peptide substrate for sortase can be modified with peptides, proteins, fluorophores, etc.\(^{31-35}\), phage can be decorated with a wide range of substituents. As far as display of proteins is concerned, proteins similar in size to GFP have been incorporated at fewer than one copy per phage on pVIII using a phagemid system\(^{18}\). Using sortase, we display ~90 GFP molecules on average per phage particle.

For the pIII and pIX proteins, we show that every phage can be labeled with multiple copies of the desired peptide/protein (Table 1). An advantage of using sortase to covalently attach proteins to phage over genetically engineering pIII directly is that it ensures display of the
correct quaternary structure of the protein. This can be inferred from our experiments using streptavidin. The mixing of two phage particles, one containing streptavidin on pIII and the other containing biotin on pVIII results in a novel and complex phage structure. This shows that the streptavidin structure displayed on phage remains fully active and binds biotin.

Sortase enzymes in combination with the streptavidin-biotin pair or in conjunction with click-chemistry can generate novel structures. The ability of patterning and aligning materials on phage or of increasing its surface area is important for the development of new materials. For example, the lampbrush phage structure generated here (Fig. 5) may find application in light-sensitive processes where phage branching off the stem could be functionalized to act as antennae to capture light.

In addition to N-terminal labeling of single capsid proteins, two capsid proteins were labeled site-specifically on a single phage particle using two orthogonal sortases. This could be explored for panning of antibody libraries displayed on pIII. Due to the exquisite site-specificity of sortase, fluorescent peptides can be added to pVIII without modification of the moiety displayed at pIII. Fluorescent labeling by other chemistries does not easily afford such specificity, especially when displaying a large moiety, such as an antibody fragment. The sensitivity of detection should increase when a phage particle contains many fluorophore groups on pVIII. This is indeed what we observe in our flow cytometry experiments, showing that this strategy greatly enhances the sensitivity of detection. Increased sensitivity would be instrumental in the context of a future panning strategy for detection of rare binding events, whether due to low concentration of the target or low phage concentration.
Modification of pIII and pIX by sortase will be useful first and foremost for material applications, where the physical properties of phage and not its utility as a library vector are of prime concern. Fluorescent modification of pVIII is compatible with the construction and screening of libraries created using pIII genetic fusions. In this case, the site-specificity and yield of the sortase reaction allow the generation of libraries that can be screened directly by fluorescence. Thus, the versatility of the sortase-based labeling strategy described here will enable development of a wide array of tools, expanding the use of phage either for the creation of new materials or for new biological applications.

**Acknowledgments**

We thank Ana Avalos and Lee Kim Swee for help with FACS acquisition and to Debadyuti Ghosh and Nimrod Heldman for helpful discussions. This work was supported by the Institute for Collaborative Biotechnologies through grant W911NF-09-0001 from the U.S. Army Research Office. The content of the information does not necessarily reflect the position or the policy of the Government, and no official endorsement should be inferred. J.J.C. is grateful for a fellowship from The International Human Frontier Science Program Organization.

**Supporting Information**

List of peptides identified by mass spectrometry of GFP-pIII, GFP-pIX, and GFP-pVIII proteins; immunoblot analysis of streptavidin-pIII labeling; atomic force microscopy images of lampbrush phage, streptavidin-pIII phage, and biotin-pVIII phage; table of oligonucleotides used in phage engineering; and method for estimating the number of GFPs that can be packed on the surface of phage. This material is available free of charge via the Internet at http://pubs.acs.org.

**Abbreviations**
ACF=autocorrelation function; AFM=atomic force microscopy; BAP=biotin acceptor peptide; DLS=dynamic light scattering; FACS=fluorescence-activated cell sorting; GFP=green fluorescent protein; HA=hemagglutinin, HRP=horseradish peroxidase, IPTG=isopropyl β-D-1-thiogalactopyranoside; LB=Luria broth; MHC=major histocompatibility complex; PBS=phosphate buffer saline, PEG=polyethylene glycol; pfu=plaque forming units; SDS-PAGE=sodium dodecyl sulfate polyacrylamide gel electrophoresis; SrtA\textsubscript{aureus}=sortase A from \textit{Staphylococcus aureus}; SrtA\textsubscript{pyogenes}=sortase A from \textit{Streptococcus pyogenes}; TBS=Tris buffered saline; TAMRA=tetramethylrhodamine, VHH=camelid heavy-chain antibody; X-gal=5-bromo-4-chloro-indolyl-β-D-galactopyranoside.
References


Table 1 – Labeling efficiency for each of the phage coat proteins using sortase.

<table>
<thead>
<tr>
<th>Minor Capsid Proteins</th>
<th>Capsid Protein</th>
<th>Probe</th>
<th>Efficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>pIII</td>
<td>Biotin</td>
<td>68±9%</td>
</tr>
<tr>
<td></td>
<td>pIII</td>
<td>GFP</td>
<td>56±2%</td>
</tr>
<tr>
<td></td>
<td>pIX</td>
<td>Biotin</td>
<td>73±2%</td>
</tr>
<tr>
<td></td>
<td>pIX</td>
<td>GFP</td>
<td>74±1%</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Major Capsid Protein</th>
<th>Capsid Protein</th>
<th>Probe</th>
<th>Optimal Packing</th>
<th>Copy Number/Phage Using Sortase</th>
<th>Literature</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>pVIII</td>
<td>Biotin</td>
<td>2700</td>
<td>1350±90</td>
<td>400-700</td>
</tr>
<tr>
<td></td>
<td>pVIII</td>
<td>GFP</td>
<td>385</td>
<td>91±20</td>
<td>&lt;1</td>
</tr>
</tbody>
</table>
Figure 1 - M13 bacteriophage structure and sortase-based reaction schemes. M13 bacteriophage is composed of five capsid proteins. pVIII is the major capsid protein with ~2700 copies in each phage particle. The pVII (light blue) and pIX (blue) are located at one end and start the assembly process, while pIII (green) and pVI (red) are at the other end and cap the phage. Note: the image is not to scale (a). Schematic representation of the mechanism of chemo-enzymatic labeling mediated by Staphylococcus aureus (SrtA<sub>aureus</sub>-left) or Streptococcus pyogenes (SrtA<sub>pyogenes</sub>-right) (b).
Figure 2 – pIII labeling. G5-pIII modified phage (there are five copies of pIII/ phage particle) was incubated with SrtA_Aureus and K(biotin)-LPETGG peptide (a), or GFP-LPETG (b), for 3hrs at 37˚C or room temperature, respectively. The reactions were monitored by SDS-PAGE under reducing conditions followed by immunoblotting using streptavidin-HRP (a-top panel) or an anti-pIII antibody (a-bottom panel and b). The molecular weight markers are shown on the left. The unidentified anti-pIII reactive protein (*) is most probably a proteolytic fragment of pIII. The identity of the GFP-pIII fusion product was determined by mass-spectrometry. The amino acid sequences of pIII and GFP are shown in blue and green, respectively. The peptides identified are highlighted in bold. The tryptic peptide comprising the GFP C-terminus, followed by the SrtA_Aureus cleavage site, fused to the N-terminal glycines of pIII is shown in red.
Figure 3 - pIX labeling. G5:HA-pIX modified phage (there are five copies of pIX/phage particle) was incubated with SrtA_aureus and K(biotin)-LPETGG peptide (a), or GFP-LPETG (b), at 37°C and room temperature, respectively, for the times indicated. The reactions were monitored by SDS-PAGE under reducing conditions followed by immunoblotting using streptavidin-HRP (a-top panel) or an anti-HA antibody (a-bottom panel and b). The molecular weight markers are shown on the left. The identity of the GFP-pIX fusion product was determined by mass-spectrometry. The amino acid sequences of pIX and GFP are shown in blue and green, respectively. The peptides identified are highlighted in bold. The AspN digestion-resultant peptide comprising the GFP C-terminus, followed by the SrtA_aureus cleavage site, fused to the N-terminal glycines of pIX is shown in red.
Figure 4 – pVIII labeling. A$_2$G$_4$-pVIII modified phage (there are 2700 copies of pVIII/phage particle) was incubated with SrtA$_{pyogenes}$ and K(biotin)-LPETAA peptide (a), or GFP-LPETAA (b), at 37°C for the times indicated in the figure. The reactions were monitored by SDS-PAGE under reducing conditions followed by immunoblotting using streptavidin-HRP (a) or an anti-GFP antibody (b). The molecular weight markers are shown on the left. The unidentified anti-GFP reactive protein (*) is attributed to proteolyzed GFP forming an intermediate with SrtA$_{pyogenes}$. The identity of the GFP-pVIII fusion product was determined by mass-spectrometry. The amino acid sequences of pVIII and GFP are shown in blue and green, respectively. The peptides identified are highlighted in bold. The tryptic peptide comprising the GFP C-terminus, followed by the SrtA$_{pyogenes}$ cleavage site, fused to the N-terminal alanines of pVIII is shown in red.
**Figure 5 – Creation of a multi-phage structure.** Schematic representation of the strategy used to build a lampbrush structure (a). Upon labeling of the N-terminus of pIII with streptavidin and of the N-terminus of pVIII with biotin using sortase-mediated reactions, the phage were mixed (see Experimental Procedures section for details). The resulting product was visualized by dynamic light scattering (b) and by atomic force microscopy (c).
Figure 6 – Dual labeling of phage using orthogonal SrtA<sub>pyogenes</sub> and SrtA<sub>aureus</sub>. Schematic representation of the strategy used to couple two different moieties to two different capsid proteins (a). Labeling of pVIII with a K(TAMRA)-LPETAA peptide mediated by SrtA<sub>pyogenes</sub> was followed by labeling of pIII with a single domain antibody directed to Class II MHC as a cell targeting moiety and SrtA<sub>aureus</sub> (see Experimental Procedures section for details). The final product was analyzed by fluorescent scanning imaging to visualize labeling of pVIII, followed by immunoblotting using an anti-pIII antibody to monitor the efficiency of labeling (b). There are five copies of pIII/ phage particle. The asterisks indicate unidentified anti-pIII reactive proteins, which probably correspond to proteolytic fragments of pIII. Binding of the dual labeled phage to lymphocytic Class II MHC+ cells was observed by flow cytometry (c). The Class II MHC+ enriched cell fraction of the lymph nodes of a C57BL/6 mouse was stained for B220 together with the dual labeled phage (phage-TAMRA-VHH7), TAMRA labeled phage (no cell targeting motif, phage-TAMRA), or anti-Class II MHC directly conjugated to TAMRA (TAMRA-VHH7).
### Table of Content Graphic

<table>
<thead>
<tr>
<th>Capital Form</th>
<th>Private</th>
<th>Efficiency</th>
</tr>
</thead>
<tbody>
<tr>
<td>2AP</td>
<td>Halo</td>
<td>0.03 10%</td>
</tr>
<tr>
<td>3AP</td>
<td>GFP</td>
<td>0.03 10%</td>
</tr>
<tr>
<td>2AP</td>
<td>Halo</td>
<td>0.04 10%</td>
</tr>
<tr>
<td>2AP</td>
<td>GFP</td>
<td>0.04 10%</td>
</tr>
</tbody>
</table>

### Table: Capital Formation

<table>
<thead>
<tr>
<th>Capital Form</th>
<th>Name</th>
<th>Transcript Chance</th>
</tr>
</thead>
<tbody>
<tr>
<td>2AP</td>
<td>Bone</td>
<td>2.56</td>
</tr>
<tr>
<td>2AP</td>
<td>GFP</td>
<td>0.07</td>
</tr>
</tbody>
</table>

400-700
Supplementary Information

An M13 bacteriophage display framework that allows sortase-mediated modification of surface-accessible phage proteins


†The David H. Koch Institute for Integrative Cancer Research, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
‡Biophysics Program, Harvard University, Cambridge, MA 02138, USA
§Whitehead Institute for Biomedical Research, Cambridge, MA 02142
‖Department of Materials Science and Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA
⊥Department of Biology, Massachusetts Institute of Technology, Cambridge, MA 02139
#Department of Biological Engineering, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

*Corresponding authors:

Hidde L. Ploegh / Carla P. Guimaraes
Whitehead Institute for Biomedical Research
9 Cambridge Center
Cambridge, MA 02142
Phone: (617) 324-2031
Fax: (617) 452-3566
e-mail: ploegh@wi.mit.edu / guimarae@wi.mit.edu

Angela M. Belcher
Massachusetts Institute of Technology
77 Massachusetts Avenue, Room 76-561
Cambridge, MA 02139
(617) 252 1163
e-mail: belcher@mit.edu

Figure S1  2
Figure S2  3
Figure S3  4
Figure S4  5
Figure S5  6
Table S1  7
Supplemental Methods  8
Figure S1 – Characterization of the GFP-pIII conjugate by mass spectrometry. The polypeptide corresponding to GFP-pIII was excised from the SDS-PAGE gel and digested with trypsin. The resulting peptides were analyzed by liquid chromatography MS/MS. Peptides positively identified by sequence are highlighted in yellow and bold.
Figure S2 – Characterization of the GFP-pIX conjugate by mass spectrometry. The polypeptide corresponding to GFP-pIII was excised from the SDS-PAGE gel and digested with AspN. The resulting peptides were analyzed by liquid chromatography MS/MS. Peptides positively identified by sequence are highlighted in yellow and bold.
Figure S3 – Characterization of the GFP-pVIII conjugate by mass spectrometry. The polypeptide corresponding to GFP-pVIII was excised from the SDS-PAGE gel and digested with trypsin. The resulting peptides were analyzed by liquid chromatography MS/MS. Peptides positively identified by sequence are highlighted in yellow and bold.
### Figure S4 – pIII labeling with streptavidin.

Gs-pIII phage (there are five copies of pIII/phage particle) was incubated with SrtA\textsubscript{aureus} and streptavidin containing a C-terminal LPETG motif in each monomer. The reactions were monitored by SDS-PAGE under reducing conditions followed by immunoblotting using an anti-pIII antibody. The molecular weight markers are shown on the left. The unidentified anti-pIII reactive protein (*) is attributed to proteolyzed pIII. The identity of the streptavidin-pIII fusion product was determined by mass-spectrometry. The amino acid sequences of pIII and streptavidin monomer are shown in blue and green, respectively. The peptides identified are highlighted in bold. The tryptic peptide comprising the streptavidin C-terminus, followed by the SrtA\textsubscript{aureus} cleavage site, fused to the N-terminal glycines of pIII is shown in red.

<table>
<thead>
<tr>
<th>Phage (200 nM)</th>
<th>-</th>
<th>-</th>
<th>+</th>
<th>+</th>
<th>+</th>
</tr>
</thead>
<tbody>
<tr>
<td>Streptavidin (µM)</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>50</td>
</tr>
<tr>
<td>Sortase (50 µM)</td>
<td>-</td>
<td>+</td>
<td>-</td>
<td>+</td>
<td>+</td>
</tr>
</tbody>
</table>

**Streptavidin-pIII**

| MAEAGITGTV YNQLGSTFIV TAGADGALTG TYESAVGNAE |
| SRYVLTYRD SAPATDGSST ALGMTVAWKH NVRNAHSATT |
| WSGQYVGCAE ARINTQWLII SGTEANAWK STLVDHDTPT |

<table>
<thead>
<tr>
<th>pIII phage</th>
<th>(there are five copies of pIII/phage particle)</th>
</tr>
</thead>
<tbody>
<tr>
<td>KVKPSSAESL LPETGSSGGGS AETVESCLAK SHTENSFTNV</td>
<td></td>
</tr>
<tr>
<td>WKDDKTDLRY ANYEGCLWNA TGVVCTGDE TQCYGTVWPI</td>
<td></td>
</tr>
<tr>
<td>GLAIIFENEG GSEGSSGSEG GSEGSGFKPP EYGDTP1PGY</td>
<td></td>
</tr>
<tr>
<td>TINPLGQFTY POTEQNFPN PNPSLEEQQP LNTMFQUNR</td>
<td></td>
</tr>
<tr>
<td>FRNRQGALTV YTTVTQGTD PVKTYTQFTP VSSKAMYDAY</td>
<td></td>
</tr>
<tr>
<td>WNGKFRDCAP HGFGSDLVY CEYQQSSLDL P0FFVNAVGG</td>
<td></td>
</tr>
<tr>
<td>SGGGSNSGGE GGGSEGSGSG GGGSEGSGDY</td>
<td></td>
</tr>
<tr>
<td>EKMANNKGA MTENADENAL QSDAKGLDLS VATDGAAJID</td>
<td></td>
</tr>
<tr>
<td>GFIGDVSSGLA NNGATGDFGA GSNQMAQVG DGDNSPLMN</td>
<td></td>
</tr>
<tr>
<td>FRQYLPSLFQ SVECRPFFVG AGKPYEFSID CDKINLFRGV</td>
<td></td>
</tr>
<tr>
<td>FAVLLYVATF MYVFSTFANI LRNKES</td>
<td></td>
</tr>
</tbody>
</table>

**α-pIII**

<table>
<thead>
<tr>
<th>15</th>
<th>25</th>
<th>37</th>
<th>50</th>
<th>75</th>
<th>100</th>
</tr>
</thead>
</table>

---

5/9
Figure S5 – AFM characterization of lampbrush phage structure. Phage with the N-terminus of pIII labeled with streptavidin and phage with the N-terminus of pVIII conjugated to biotin were created using sortase-mediated reactions. The phage preparations were visualized by AFM before (top right and top left panels) and after mixing (bottom panels; see Experimental Procedures section for details).
**Table S1 - Oligonucleotides for phage engineering**

<table>
<thead>
<tr>
<th>Name</th>
<th>Sequence (5’-3’)</th>
</tr>
</thead>
<tbody>
<tr>
<td>G5pIIIC</td>
<td>GTACCTTTTCTATTCTCACTCTGTGGAGGCGGTGGATC</td>
</tr>
<tr>
<td>G5pIIIINC</td>
<td>GGCCGATCCACCCTCCACCAGAGTGGAGAATAGAAAG</td>
</tr>
<tr>
<td>A2G4pVIIIC</td>
<td>GCTGCGCGGGGAGGG</td>
</tr>
<tr>
<td>A2G4pVIIINC</td>
<td>GATCCCTCCCCCGCCAGCTGCA</td>
</tr>
<tr>
<td>G5HApIXC</td>
<td>CGGCCATGGCGGCGGCGGAGGTGGAGGCTACCATACGATGT</td>
</tr>
<tr>
<td>G5HApIXNC</td>
<td>TCCAGATTACGCTCAGGG</td>
</tr>
<tr>
<td></td>
<td>TGAGCGTATATCGGAAACATCGTAGGCTATGGGTAGCCTCCACCTC</td>
</tr>
<tr>
<td></td>
<td>CGCCCCGATTGGCCGGCT</td>
</tr>
<tr>
<td>AADSPH-pVIII-Top</td>
<td>GTCAGTGCTGCTTTCGCTGCTGCAGATTCGGCCGCATAC</td>
</tr>
<tr>
<td>AADSPH-pVIII-Bottom</td>
<td>CTCAGTBGCAGCGGAATCTGCAGCGACGAGAAAGACAGCATC</td>
</tr>
<tr>
<td></td>
<td>GGAAC</td>
</tr>
</tbody>
</table>
Supplemental Methods

Estimating nearest neighbor packing of GFP on phage surface

Using the crystal structure of the pVIII capsid protein (1IFJ)\textsuperscript{1}, a model viral capsid was constructed with five fold symmetry serving as a model of the phage surface. A crystal structure of GFP (1GFL)\textsuperscript{2} was oriented such that its C-terminus was adjacent to the N-terminus of pVIII. By analyzing this image, it was determined that one GFP molecule blocked the N-termini of the six pVIII proteins surrounding the GFP-pVIII fusion meaning at most one out of seven pVIII proteins can be labeled with a GFP. From this, it was calculated that a single virion with 2700 pVIII proteins would have at most 385 GFP molecules. The visualizations were performed using WinCoot\textsuperscript{3}.
References

