Engineering Phage Materials with Desired Peptide Display: Rational Design Sustained through Natural Selection

Anna Merzlyak[†] and Seung-Wuk Lee*

UCSF and UC Berkeley Joint Graduate Group in Bioengineering, Berkeley, California 94720, Physical Biosciences Division, Lawrence Berkeley National Laboratory, Berkeley, California 94720, Department of Bioengineering, University of California, Berkeley, California 94720, and Berkeley Nanoscience and Nanoengineering Institute, Berkeley, California 94720. Received July 8, 2009; Revised Manuscript Received August 24, 2009

Genetic engineering of phage provides novel opportunities to build various nanomaterials by displaying functional peptide motifs on its surface coat protein. However, any genetic modifications of phage coat proteins must be able to accommodate their many biological roles in the phage replication process. To express functional but inherently unfavorable peptide motifs on major coat protein pVIII, we devised a novel genetic conjugation method to circumvent bacterial biological censorship. Constraining the designed peptides among the degenerate flanking residues, we obtained a pVIII library of phage that retained the desired sequences yet could navigate through the phage replication process due to the naturally selected flanking residues. Further, we systematically analyzed the biochemical and size-related compensation mechanisms of the pVIII expressed peptides by constructing four chemically diverse (His, Trp, Glu, Lys) partial library series. Described genetic conjugation methodology can serve to improve the design of engineered phage and allow further exploitation of these particles as functional nanobiomaterials for various applications.

INTRODUCTION

Monodisperse nanofibers with a uniform programmable surface that are able to store information and be reproduced in mass quantities can serve a multitude of biological and material needs. Filamentous bacteriophage (phage) are bacterial viruses 880 nm long and 6.6 nm in diameter (1, 2). Through genetic engineering, phage have been shown to be extremely versatile in peptide presentation on both their major coat protein, pVIII, and minor coat proteins, pIII and pIX (1−7). Additionally, the monodispersity of their shape and chemical character, derived from clonal replication, allows them to self-assemble into directionally organized liquid crystalline structures. Previously, these phage have demonstrated an ability to be formed into one-dimensional ordered fibers, two-dimensional films, and three-dimensional materials (4, 6, 8−10).

The potential to control both structural organization and chemical function of materials at the nanoscale level makes genetically engineered phage very appealing candidates for use as material building blocks. The surface of the filamentous phage is represented by 2700 copies of its major coat protein, pVIII, which packs around the phage DNA to make a tight and cylindrical shell (11). The acidic N-termini of pVIII form a dense periodic display spaced at 2.7 nm (1, 12). The final five residues of pVIII have been determined to be structurally unconstrained and exposed to the solvent (11), therefore presenting an optimal target for engineering, manipulation, and substrate interaction.

A tremendous amount of research has gone into studying the structural characteristics and biological interactions of pVIII through the phage and *Escherichia coli* (*E. coli*) life cycle (13−23). Numerous peptide libraries have been developed for use as information mining tools in a high-throughput evolutionary screening method called phage display (1, 2). Additionally, research efforts have gone into finding out how to improve the diversity of the amino acid libraries presented by the phage, thereby potentially improving the quality of binding information between the peptide and its target (24−26). More recently, functionalized bacteriophage and virus capsids have become an attractive tool for material synthesis (27, 28). The "landscape" peptide presentation on the major coat protein of the filamentous phage has been utilized to template inorganic crystals for energy and memory storage devices (4, 6, 9, 10) and make stimulus-responsive materials (29). The phage has also been exploited for medical applications, such as targeted drug (30), gene (31), and imaging agent (32) delivery, as well as a tissue engineering material (33).

To further exploit and push the limits of the phage particle versatility, it is important to understand the intricate biological processes that occur before pVIII proteins can be successfully packaged into a viable phage. If a very low copy number vector such as fd-tet is utilized (24), the stress on the replication process is reduced, and an engineered octapeptide pVIII library has a capability of forming 10^9 various phage particles (1). The utilization of a phagemid cloning system also allows a much greater versatility of expressed proteins in terms of both size and chemical characteristic, albeit on a much lower percentage of phage proteins (2). Even though the phagemid system offers greater flexibility, limitations still exist. A recent study showed that both the charge of the expressed peptide as well as its size influence the number of fused peptides that could be displayed on the phage particle (34). The potential complexity of the libraries drops when using a higher copy vector or when expressing the desired peptide on every protein of the phage (25). However, as the yield of propagated phage and the density of expressed peptides become much more important for material synthesis applications these attributes should play an important role in choosing both the phage type and the expression system.

Whether the certain insert sequence will be expressed or not depends strongly on its fitness for the processes encountered

* Corresponding author. Seung-Wuk Lee. University of California, Berkeley, CA 94720. Phone: 510-486-4628, Fax: (510) 486-6488, E-mail: leesw@berkeley.edu.
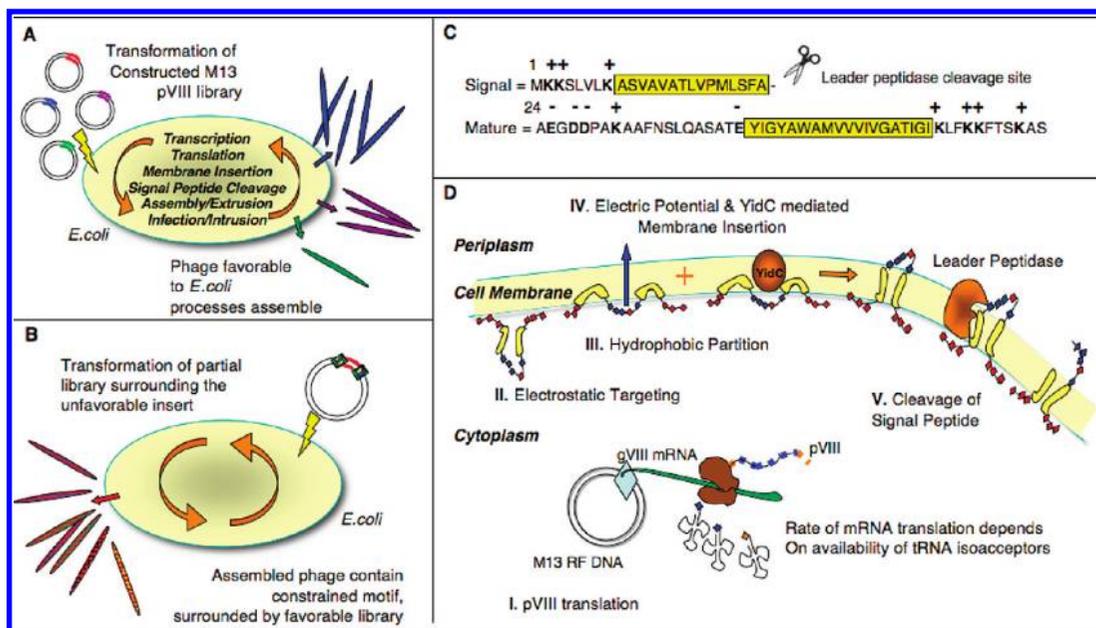† UCSF and UC Berkeley Joint Graduate Group in Bioengineering, Berkeley and Lawrence Berkeley National Laboratory.

Figure 1. pVIII major coat protein, and its biological role in M13 phage replication in *E. coli*. (A) Schematic representation of pVIII library construction. (B) Schematic representation of partially constrained pVIII libraries. (C) Signal and mature peptide sequences of pVIII protein, with hydrophobic portions marked by yellow rectangles, and a marked leader peptidase cleavage site. (D) More detailed diagram of translation, membrane insertion, and procoat cleavage processes most instrumental in biocensoring the pVIII insert expression. (In part adapted from ref *51*).

by pVIII during its lifecycle, including (1) transcription, (2) translation, (3) membrane insertion, (4) signal peptide cleavage, (5) assembly and extrusion, and finally (6) infection and intrusion (*23, 35*); see Figure 1. If the interactions of the engineered pVIII are favorable within this context, its genetic sequence will be more stable, and more phage will be produced during amplification. If the insert sequence makes pVIII interactions less biologically favorable, it will be more likely to mutate and/or produce fewer phage particles. Here, we describe a novel partially constrained library method to add peptide motifs to every copy of pVIII protein. We show how this approach was used to display an Arg-Gly-Asp (RGD) integrin-binding peptide on every copy of pVIII protein to promote cell interaction, which we developed for phage-based novel tissue engineering materials. Additionally, we show how this method can be useful for presenting other functionally designed groups on the major coat protein of the phage. Finally, we quantitatively analyze how the characteristics of the inserts and their constrained sequences affect their expression on a phage particle.

EXPERIMENTAL PROCEDURES

Reagents. Phusion high-fidelity DNA polymerase was purchased from Finnzymes (Espoo, Finland). *Pst*I, *Bam*HI, and *Dpn*I restriction endonucleases, T4 DNA ligase, and dNTP 10 mM solution mix were from New England Biolabs (Ipswich, MA). QuikChange Site-Directed Mutagenesis Kit was from Stratagene (La Jolla, CA). DNA purification was performed with the help of QIAquick PCR Purification and Gel Extraction Kits, and QIAprep Spin Miniprep Kit all from QIAGEN, Inc. (Valencia, CA).

Bacterial and Phage Strains. M13KE single-stranded phage that was used as a base for the library and insert construction was purchased from New England Biolabs. XL10-Gold Ultra-competent Cells (F′ *lacI^qZΔM15*) from Stratagene were used to transform the ligated DNA products. XL1-Blue Competent Cells (F′ *lacI^qZΔM15*) from Stratagene were used for phage amplification.

Inverse PCR Method for M13 Phage Cloning. To present peptide motifs on every copy of M13 major coat protein, an inverse polymerase chain reaction (PCR) cloning method was adapted (*36*). The insert was positioned between the first and fifth amino acids of the wild-type pVIII, replacing wild-type M13KE residues 2−4 (Ala-**Glu-Gly-Asp**-Asp to Ala-(**Insert**)-Asp). All primers were ordered from IDT DNA technologies (Coralville, IA). To allow for recircularization of the vector following the PCR, a *Pst*I restriction site was created upstream of the insert location using a QuikChange Site-Directed Mutagenesis Kit by changing position 1372 of M13KE vector from T to A (CTGC<u>A</u>G), as previously described (*1*). The resulting DNA was verified by picking blue plaques resulting from phage transformation, isolating the DNA using common biological methods (*37*), and sequencing at the University of California, Berkeley DNA sequencing facility (Berkeley, CA).

For the inverse PCR reaction, the forward primer was designed to include a *Pst*I restriction site followed by an insert sequence and a segment complementary to the gVIII 3′−5′ strand. The reverse primer, designed to make the M13 plasmid linear, also included the *Pst*I restriction site and was fully complementary to the gVIII 5′−3′ region. To incorporate the gene sequences, PCR was performed using Phusion high-fidelity DNA polymerase, the two primers, and an M13 vector with an engineered *Pst*I restriction site as the template. The obtained product was purified on an agarose gel, eluted with spin column purification, digested with *Pst*I enzyme, and recircularized with an overnight ligation at 16 °C with T4 DNA ligase (*37*). The ligated DNA vector was transformed into XL10-Gold Ultra-competent bacteria cells, and the resulting phage verified via DNA sequencing.

Partial Library Cloning Method. All the libraries were cloned into the M13 vector by following the above scheme. For the partial libraries, the primers were designed to constrain a region of interest (i.e., RGD) and to allow degeneracy within the flanking codons (i.e., XXXRGDXX). As in previous phage libraries, a 32 codon degeneracy was used (X = NNK, N = A/T/C/G, K = G/T) to reduce the bias among presented amino acids (AAs) and eliminate two of the three potential stop codons (*1, 2*). For a fully unconstrained library, degeneracy was allowed at all of the 8 positions. See Table 1 for primer and corresponding library sequences. To reduce the advantage of

Table 1. Primer Sequences for Phages and Phage Libraries[a]

name	sequence	primer[b]
wild-type	AEGDDP	
unconstrained	A*X^c*(*X*)*₇*DP...	5′ATATAT<u>CTGCAG</u>*NK (NNK)₇*GATCCCGCAAAAGCGG CCTTTAATCCC3′
RGD	A*X^cXX***RGD***XX*DP	5′...<u>CTGCAG</u>*NK (NNK)₂* **CGTGGTGAC**(*NNK)₂*...3′
WHWQ	A*G***WHWQ**GGG*G*DP	5′...<u>CTGCAG</u>*GC TGG CAT TGG CAG GGC GGC GGC*...3′
WHWQ-2X	A*X^c***WHWQ***XG*G*DP	5′...<u>CTGCAG</u> *NK* **TGG CAT TGG CAG** *NNK* **GGC GGC**...3′
WHWQ-3X	A*X^c***WHWQ***XX*G*DP	5′...<u>CTGCAG</u> *NK* **TGG CAT TGG CAG** (*NNK)₂***GGC**...3′
WHWQ-4X	A*X^cX***WHWQ***XX*DP	5′...<u>CTGCAG</u> *NK NNK* **TGG CAT TGG CAG** (*NNK)₂*...3′
	A*X^c***WHWQ***XXX*DP	5′...<u>CTGCAG</u> *NK* **TGG CAT TGG CAG** (*NNK)₃*...3′
1E[d]	A*X^cXX***E***XXXX*DP	5′...<u>CTGCAG</u> *NK (NNK)₂***GAA** (*NNK)₄*...3′
2E	A*X^cXX***EE***XXX*DP	5′...<u>CTGCAG</u> *NK (NNK)₂***GAAGAG** (*NNK)₃*...3′
3E	A*X^cX***EEE***XXX*DP	5′...<u>CTGCAG</u> *NK NNK* **GAAGAGGAA** (*NNK)₃*...3′
4E	A*X^cX***EEEE***XX*DP	5′...<u>CTGCAG</u> *NK NNK* (**GAAGAG**)₂(*NNK)₂*...3′
5E	A*X^c***EEEEE***XX*DP	5′...<u>CTGCAG</u> *NK* (**GAAGAG**)₂ **GAA** (*NNK)₂*...3′
6E	A*X^c***EEEEEE***X*DP	5′...<u>CTGCAG</u> *NK* (**GAAGAG**)₃*NNK* ...3′
p8-rev1376		5′ CCT<u>CTGCAG</u>CGAAAGACAGCATCGG 3′

[a] X's refer to degenerate residues, **Bold** to constrained, and *Italics* to the insert portion. [b] Full primer sequence shown only for the first listed; for the rest, only the *Pst*I restriction site and the insert portion are included. [c] X^c refers to residues coded for by a GNK codon (i.e., A,D,E,G,V). [d] Glu libraries shown as representative; same pattern applied for Lys library (AAA AAG codons); for His (CAT) and W (TGG) libraries, only one type of codon was used.

wild-type or fast-amplifying phage, the cells were allowed to recover for only 30 min after transformation (*38*). To analyze the library sequence space, the cells were plated in agarose top, and phage-formed plaques after an overnight incubation at 37 °C were picked and their DNA sequenced.

Library Quantification and Analysis. *Number of Transformants.* Using 2 μL of each ligation reaction, 50 μL of XL-10 Gold Ultracompetent cells were transformed via heat shock following the manufacturer's recommendations. The cells were recovered in 1 mL LB at 37 °C incubator with shaking for 30 min. Aliquots of 150 and 450 μL of cells were then plated in agarose top on LB plates with IPTG/Xgal (*37*). The number of transformants was quantified for each of the engineered libraries by calculating an average quantity of plaque forming units.

Amino Acid Diversity. Ten sequences were obtained for each library producing plaques (*37*), except for the 3H library where only seven plaques could be obtained with retained 3H sequence. The frequency and number distributions were determined on the basis of the amino acid (AA) side-chain characteristic (i.e., aromatic, negative, and positive), both in the degenerate positions and in the insert as a whole. Observed frequency of occurrence (F) was calculated via the following formula: $F =$ (# of type of AA)/(total # AAs in sequences evaluated) × 100%, with type of AA = Aromatic (Trp, Tyr, Phe), Negative (Asp, Glu), or Positive (Arg, Lys). Expected frequency (E.F.) was calculated on basis of the number of codons corresponding to each AA residue as used in a 32 codon library, and grouped by the type of AA.

RESULTS

Fabrication of Partially Constrained pVIII Library with Desired Peptides. We fabricated a partially constrained pVIII library with desired peptides for construction of engineering materials. The initial purpose for engineering the pVIII protein of M13 phage was to decorate it with cell signaling motifs, so it could serve as a novel material for tissue regeneration scaffolds (*33*). As the target cell type was of neural origin, the signaling motifs chosen were IKVAV and RGD, which are peptide groups taken from laminin and fibronectin and previously shown to have a positive effect on neural progenitor cell attachment and differentiation (*39, 40*). Initial cloning attempts were to display an **ASIKVAV** and a **GGRGD-**SP peptide immediately at the N-terminus of pVIII, thereby maintaining the signaling motifs within their naturally occurring flanking sequences (*41, 42*). After the original attempt failed, the insert sequences were modified to shift the ASIKVAV to

start at position three of the mature pVIII (i.e., AE-*ASIKVAV*DP), and GRGDS was shortened and moved to start after the first Ala residue (i.e., A*GRGDS*DP). Both were redesigned to contain a positive and a negative residue, making the overall charge of pVIII N-terminus −1. The described measures should have theoretically allowed for the expression of the insert (*3*). However, both of the constructed bacteriophage showed a strong tendency for deletions and single nucleotide mutations, which resulted in amino acid substitutions. In nine out of ten sequences obtained, the ASIKVAV insert had a K6N mutation (AE*ASINVAV*), and only one preserved the desired IKVAV sequence with an A3D mutation (AE*DSIKVAV*).

A GRGDS peptide addition on pVIII resulted in a phage that made very small plaques, initially indicating a greatly reduced fitness of the phage (*3, 15, 16, 24*). Furthermore, when the phage was amplified and titrated following standard microbiological procedures (*37*), the resulting phage all had G2D, R3X (i.e., A*DRGDS*DP or A*GXGDS*DP, with X = residue from a single base mutation, i.e., Pro, Leu, Gly, His) or deletion mutations. Even when a DRGDS phage was amplified, the resulting population again showed consistent instability at the Arg position, often mutating it to a neutral residue. From these initial cloning attempts, it was seen that the preference for the phage expression was not to have a positive residue in the N-terminus insert location, unless stabilized by at least two negative charges. However, even with two negative residues the shorter DRGDS peptide continued to remain unstable.

A partial 8mer library was utilized to display the unfavorable RGD peptide, which allowed the unconstrained library residues to be controlled and chosen by the *E. coli* host (see Figure 1). In this manner, only the favorable full inserts would be displayed and contain both the designed constrained motif as well as a flanking amino acid library chosen by the natural selection pressures of the *E. coli* cellular mechanisms. Cloning the RGD sequence with the partial library method as described above resulted in formation of phage particles that upon sequencing were found to retain the RGD sequence (see Table 2). From the resulting RGD library, the most physiologically relevant flanking amino acids GRGDT (for a full sequence of A*DS*-**GRGDT**EDP) were chosen to resemble a fibronectin integrin binding domain GRGDS (*43*). The stability of both the IKVAV- and RGD-modified phages was verified by mass amplification (*33*). Furthermore, they were found to have specific interaction with cells and were used to build tissue engineering scaffolds that could control cell behavior in two and three dimensions (*33*).

Table 2. RGD Constrained Library Sequences

Sequence #	pVIII N-terminus Sequence										Insert charge	Insert mol wt
1[a]	A	G	V	I	R	G	D	G	D	DP	-1	769.82
2	A	G	A	S	R	G	D	D	G	DP	-1	715.69
3	A	G	T	A	R	G	D	G	E	DP	-1	743.75
4	A	E	T	I	R	G	D	G	G	DP	-1	785.82
5	A	G	G	G	R	G	D	E	D	DP	-2	743.69
6[b]	A	D	S	G	R	G	D	T	E	DP	-2	817.78
7	A	D	V	G	R	G	D	I	A	DP	-1	783.86
8	A	A	E	V	R	G	D	I	E	DP	-2	869.96
9	A	D	V	G	R	G	D	D	A	DP	-2	785.79
10	A	D	V	M	R	G	D	S	G	DP	-1	817.88
11	A	E	N	V	R	G	D	A	N	DP	-1	855.9
12	A	A	E	A	R	G	D	H	S	DP	-1	823.85

[a] The insert is shown in *italics,* constrained sequences are marked in **bold** and outlined, positive residues marked in red and the negative in blue.
[b] The sequence of a phage chosen to serve as a part of a tissue-engineering scaffold (*33*).

Table 3. WHWQ Constrained Library Sequences

Sequence #	pVIII WHWQ Sequences										Insert charge	Insert mol wt
1[a]	A	E	E	W	H	W	Q	E	G	DP	-3	1082.11
2	A	D	E	W	H	W	Q	D	G	DP	-3	1054.05
3	A	E	D	W	H	W	Q	D	G	DP	-3	1054.05
4	A	D	D	W	H	W	Q	E	G	DP	-3	1054.05
5	A	E	E	W	H	W	Q	E	G	DP	-3	1082.11
6	A	D	D	W	H	W	Q	D	G	DP	-3	1040.02
7	A	E	W	H	W	Q	E	G	D	DP	-3	1068.08
8	A	D	W	H	W	Q	D	G	D	DP	-3	1040.02
9	A	D	W	H	W	Q	D	S	E	DP	-3	1084.08
10	A	D	W	H	W	Q	E	D	A	DP	-3	1068.09
11	A	D	E	W	H	W	Q	Y	D	DP	-3	1160.18
12	A	E	D	W	H	W	Q	I	E	DP	-3	1124.19
13	A	D	E	W	H	W	Q	E	S	DP	-3	1098.11
14	A	E	E	W	H	W	Q	E	S	DP	-3	1112.14
15	A	D	T	W	H	W	Q	E	E	DP	-3	1112.14
16	A	E	A	W	H	W	Q	E	E	DP	-3	1096.15
17	A	E	D	W	H	W	Q	Q	E	DP	-3	1139.17
18	A	D	D	W	H	W	Q	N	D	DP	-3	1097.08
19	A	E	D	W	H	W	Q	Q	N	DP	-2	1124.16
20	A	D	D	W	H	W	Q	S	T	DP	-2	1056.07

[a] The insert is shown in *italics,* constrained sequences are marked in **bold** and outlined, positive residues marked in red and the negative in blue.

To further verify the utility of the partial library method to express unfavorable short peptide groups on every copy of the major coat protein, we attempted to express a WHWQ sequence found previously in our lab as a trinitrotoluene (TNT) binding peptide (*44*). A fully defined WHWQ insert and four different WHWQ libraries were attempted (see Table 1). The fully constrained WHWQ insert (A*GWHWQGGG*DP) resulted in no plaque forming phage. The transformation of a library with two unconstrained positions (A*XWHWQXGG*DP) resulted in several plaques, but all were either wild-type mutants or original template phage. Finally, libraries constructed with degeneracy at three (A*XXWHWQXG*DP) and four positions (A*XWH-WQXXX*DP or A*XXWHWQXX*DP) resulted in phages that were able to retain the desired WHWQ sequences (see Table 3). Within the resulting sequences, there was a very strong tendency of the unconstrained residues to be negative. From 20 of the phage that were sequenced and found to display the WHWQ motif, 18 had 3 negative residues in the unconstrained space and 2 had 2 negative residues. These 2 phage were from the A*XXWHWQXX*DP library, and both had the negative charges preceding the constrained WHWQ motif.

Limitation and Compensation Trends in Partially Constrained pVIII Libraries. To explore whether there was any compensation between the constrained and unconstrained sequence space of the insert and elucidate any potential trends,

several partially constrained library series were designed. One to six amino acids of the 8mer inserts were defined in the middle of the pVIII libraries as His, Lys, Trp, or Glu, leaving the rest as degenerate (see Table 1). The purpose was to systematically test the constraints of insert expression for amino acids with side groups of diverse chemical characteristics and potential functionalities. Histidine, a very functionally interesting residue, has a nearly physiological pK_a of 6.0, and is therefore very sensitive to its surroundings. However, in a more basic *E. coli* environment (*45*) it is most often found in a neutral confirmation. Lysine (pK_a = 10.53) is consistently positive in *E. coli* and has a modifiable amine group, whereas glutamic acid is negative (pK_a = 4.25) and similarly has a chemically modifiable carboxyl group. Tryptophan is uncharged, but it is the largest (mol wt = 186.21 g/mol) hydrophobic amino acid that is both aromatic and polar by nature, lending itself to a variety of binding interactions.

Effect of pVIII Constrained Sequence on Library Size. The size of the libraries, as measured by the quantity of plaque-forming phage after library transformation, was seen to strongly depend on the sequences constrained within the 8mer library inserts (Figure 2). The most apparent trend was that for all libraries the number of plaque forming phage produced within the library declined as the number of repeating residues increased. For the H-libraries, 1H and 2H constraints produced
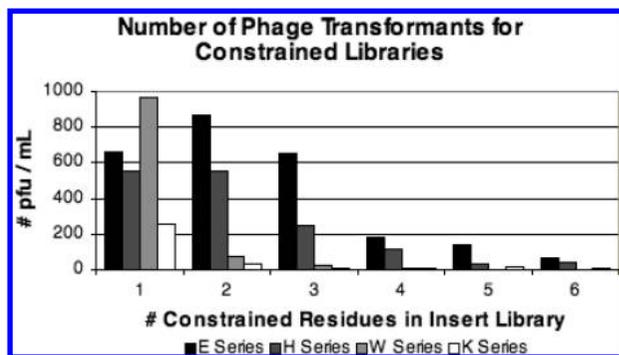
Figure 2. Partial library size dependence on the constrained peptide. Average number of plaque forming units (pfu) is shown in relation to the number of constrained residues in each of the constructed libraries.

an equal population of transformants at 550 pfu/mL and the 3H library population was reduced by half to 240 pfu/mL. The 4H library produced 120 pfu/mL, but did not result in any phage containing four His residues. Instead, the phage sequence always either had a deletion, wild-type revertant mutations, or original template background. Even though the 5H and 6H libraries produced some plaques, they were similar in quality to the 4H library. The W-libraries had the most dramatic change where the 1W seemed a very favorable constraint producing 970 pfu/mL, but the drop to the 2W library population was more than 10 times to 70 pfu/mL, and 3W was again reduced by another half to 30 pfu/mL. The 4W libraries produced less than 10 pfu/mL and had no 4W sequences. The 5W and 6W libraries produced no plaques at all. Within the K-series, only 1K (at ∼260 pfu/mL) contained the constrained Lys residue. The 2K−6K libraries produced a population of plaques, but each was either a deletion or wild-type mutant, or original template DNA background. The E-constrained phage showed high numbers of pfu/mL in the 1−3E libraries (650−870 pfu/mL), with the 2E library being the best producer. The 4E library showed a marked decline (180 pfu/mL), and even though it did contain some deletion and revertant mutants, there was still a population of phage expressing the constrained 4E sequences. The 5E (140 pfu/mL) and 6E (70 pfu/mL) libraries produced a fairly large population of plaque-forming phage compared to other high repeating libraries. However, among the obtained sequences, there were no phage that expressed five or six Glu repeats. Instead, these libraries had a high number of deletion mutants, often displaying one to four Glu residues.

Dependence of Partial Library Diversity on the Constrained Sequence. In the constructed partial libraries, the unconstrained library amino acids were seen to partially compensate for the constrained amino acids in both chemical character and size, in order to preserve the overall favorability of insert display. The residues were grouped and characterized by their side-chain functionality rather than by individual presence. The positive groups, Lys and Arg, as well as the hydrophobic groups Trp, Tyr, and Phe, were previously described to aide in anchoring the pVIII protein in the cell membrane (46, 47). The negative residues (Asp and Glu) are thought to assist in the electrophoretic dependent insertion of pVIII across the *E. coli* membrane and serve as its hydrophilic anchors (22, 47). This diversity analysis was done on the whole 8mer inserts including the constrained sequences, as well as on only the degenerate parts of the inserts, that exclude the constrained sequences (see Figure 3). These distributions were compared to an 8mer pVIII library population constructed in a similar fashion, as well as expected frequencies of these amino acids as based on the 32 codon representation (NNK codon library). As observed in previously constructed libraries, the frequency of expression of negative Asp and Glu residues were greatly overrepresented in

the fully unconstrained library, whereas both the aromatic and the positive residues were underrepresented in relation to the theoretically expected frequencies (3, 25). However, as more constraints were imposed on the insert space some trends were seen to emerge. The most pronounced effect was the direct relationship between the number of charged residues and the aromatic residues. The constriction of aromatic amino acids showed a dramatic influence on the increase of negatively charged residues within the library portions of these phage pVIII proteins (10% increase for 1W, 41% increase for 2W, and 62% increase for 3W libraries in relation to fully unconstrained). Inversely, while the expressed frequency of aromatic residues did not increase for 1E and 2E libraries, it more than tripled, in relation to the library expression, to 14% for 3E and 15% for 4E library. It is also noteworthy that, whereas the number of aromatic residues is underrepresented in the fully unconstrained and most of the constrained libraries, it exceeds the expected rate of representation (9%) by 50−60% in the 3E and 4E libraries. Additionally, the 1K library had the third largest expression rate of aromatic residues, exactly matching the expected frequency at 9%.

Since the more constrained libraries have a lower quantity of unconstrained residues, a similar number of residues of a certain type will inherently result in a higher frequency of expression. Therefore, to better interpret the behavior of the full pVIII protein, it is useful to also evaluate the average number of residues seen in the whole 8mer insert (see Figure 4) compared to both the fully unconstrained library and the wildtype residues replaced by library insertion (A**EGD**DP). With this approach, even though less dramatic, the same trends as seen with the frequency of expression can be observed. While the number of negative residues is ∼1.7 per insert for the fully unconstrained library, there is a significant increase in negative residues for the 2W and 3W as related to the library (2.5 and 3.1, respectively). Due to the overall low rate of expression of aromatic residues within unconstrained library portions of the phage, there was a large variance within each population. Therefore, besides the constrained aromatic libraries there were no statistically significant changes in the expression of aromatic residues as compared to the library. However, there was an upward trend of aromatic residue expression in the negatively constrained 3E and 4E libraries (0.6 residue/insert) in relation to the library.

When analyzing the expression of positive residues, it was seen that a maximum of one positive residue was expressed in any of the insert libraries, whether it was a constrained Lys residue or positive residues appearing in the unconstrained sections of other libraries. The constrained library had significantly higher expression of positive residues than any other library. Even though all of the negatively constrained libraries had an increased expression of positive residues, the only one that was significantly higher than the fully unconstrained library was the 2E constrained phage library. In fact, the 2E phage library expressed a significantly higher number of positive residues than all the other phage libraries with the exception of the other Glu libraries. Furthermore, in all of the inserts displaying a positive group, at least two negative residues were present.

A direct compensation mechanism seemed to occur for the constrained 1K library. It expressed an average charge of −1.5 (2.5 negative residues and 1 positive residues per insert), the same as the unconstrained library with an average charge of −1.54 (1.74 negative residues and 0.2 positive residues per insert) (Figure 3). The aromatic Trp constrained libraries expressed positive residues but at a lower rate than the unconstrained library. On the other hand, as mentioned above the Lys constrained library showed a relatively high rate of
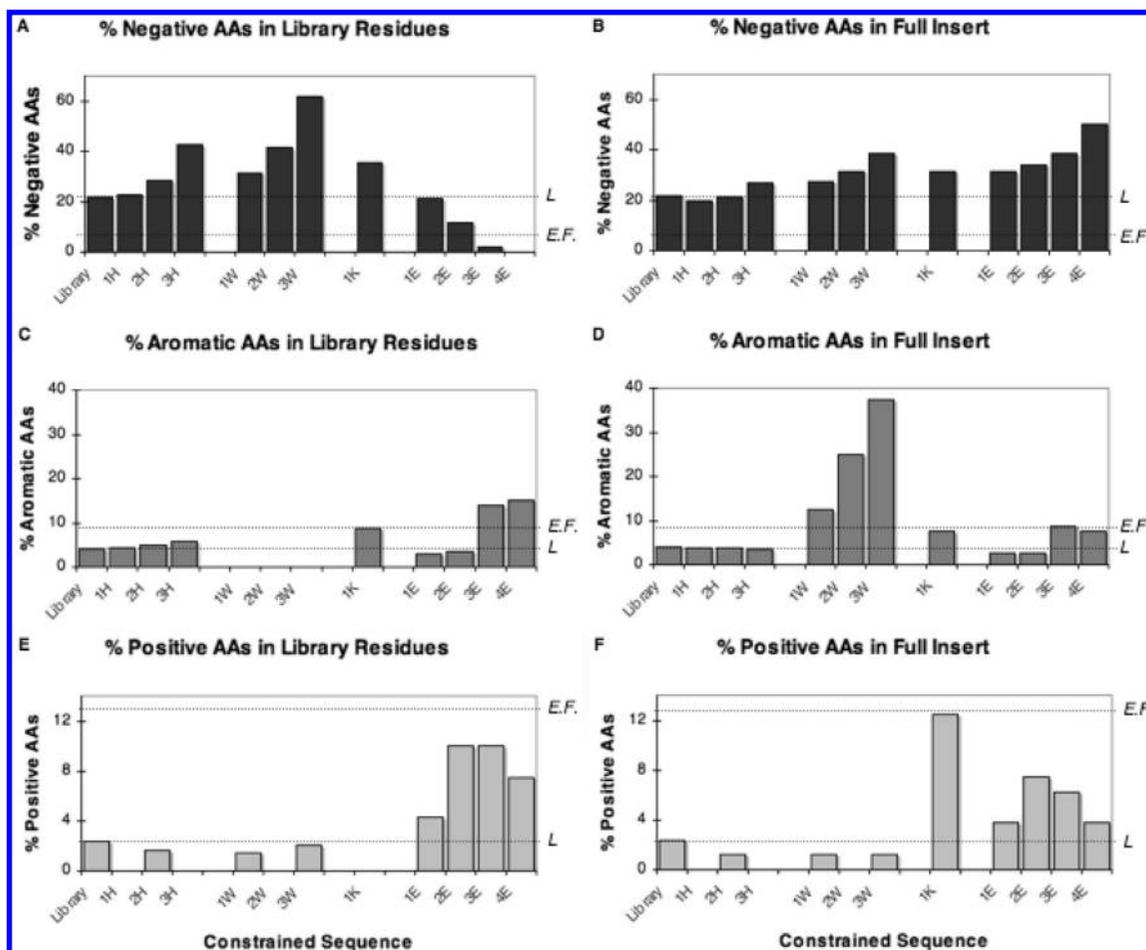
Figure 3. Frequency distribution of functional groups in expressed partial library inserts. The observed frequency of AAs with (A,B) negative, (C,D) aromatic, and (E,F) positive side chain groups is shown for both (A,C,E) the unconstrained library populations only, as well as for (B,D,F) the full 8mer insert containing both the unconstrained and the constrained sequences. Comparison lines are indicated at frequency of differently functional AAs seen in the fully unconstrained 8mer library (marked by L), and the expected frequency (marked by E.F.) as based on number of AA codons in a 32 codon library.

aromatic residue expression. When the whole insert was analyzed, the library had 0.6 aromatic residues/insert, which is similar to 3E and 4E constrained libraries.

Furthermore, with the exception of 1E, 2E, and 3E none of the constrained libraries showed any of the residues of same functionality within their unconstrained regions (i.e., Trp libraries expressed no additional aromatic residues). For the Glu libraries, the unconstrained space for 1E had the same frequency of expression for negative residues as the fully unconstrained library. The 2E library had 50% of that rate, 3E only 10%, whereas the 4E library expressed no extra negative residues within its unconstrained sequences (see Figures 3 and 4).

The His constrained libraries showed no significant variation of the aromatic residues from the fully unconstrained libraries or among each other. There was a slightly lower representation of positive residues in the His libraries, which could be attributed to the His pK_a of 6.0 having the potential to act as a slightly positive residue depending on its surroundings. The number of negative residues in the full 8mer insert did not differ significantly from the library, or each other, matching the fully unconstrained library for 1H and 2H library and then rising slightly by ∼25% for the 3H library (Figure 4).

Correlation between Charge and Size on the pVIII Inserts. The size of the pVIII inserts, quantified by their molecular weights, was found to have a significant inverse relationship with their charges, indicating that larger inserts

tended to be more negatively charged. The size−charge correlation was assessed independently for each of the library series and is depicted here by a linear regression and characterized by the correlation coefficient r, population size n, and the significance value p for a two-tailed t-value, with $t = r\sqrt{[(n - 2)/(1 - r^2)]}$. The correlation strength varied among different libraries. It was moderately weak for the fully unconstrained library ($r = -0.43$, $n = 53$, $p < 0.01$) and the E-series ($r = -0.40$, $n = 40$, $p < 0.02$), moderately strong for both H-series ($r = -0.59$, $n = 27$, $p < 0.01$) and W-series ($r = -0.60$, $n = 30$, $p < 0.001$), and very strong for the K-series ($r = -0.86$, $n = 10$, $p < 0.01$) (see Figure 5C). Regardless of the linear strength of this relationship, it was found to be statistically significant for all libraries considered. Additionally, all of the 8mer libraries expressed on pVIII seemed to have a preferred size and charge range (Figure 5A,B). The majority of inserts expressed and subsequently sequenced from all the libraries had a molecular weight between 700 and 1000, and had a net charge of −1 to −3 charges. Within that space, there was a visible shift toward both greater size and greater charge preference from unconstrained (average: −1.5, 810 Da) to constrained libraries (average: −2.3, 890 Da).

DISCUSSION

M13 bacteriophage can serve as a versatile and multifunctional material building block due to the ease of its genetic manipulation and its monodisperse filamentous shape. The coat proteins of this phage may be modified with rationally designed
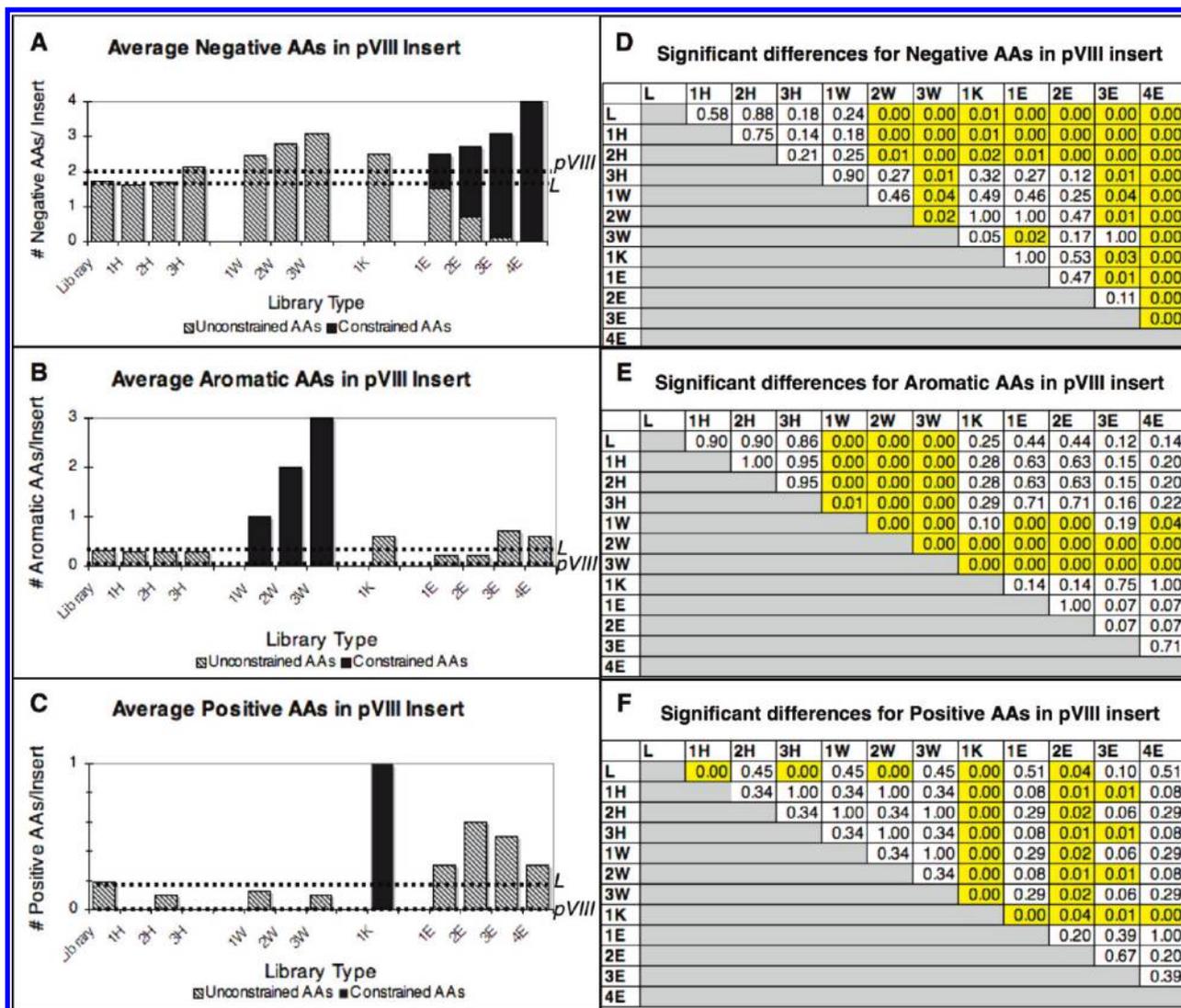
Figure 4. Average number of functional groups in expressed partial library inserts. Average number of AAs with (A) negative, (B) aromatic, and (C) positive side chain groups is shown for the constructed libraries. A contribution to the total number of groups is shown from both the unconstrained and constrained residues within the insert. Comparison lines are indicated for fully unconstrained library (marked L) and for the replaced wild-type pVIII residues at positions 2–4 (A**EGD**DP) (marked pVIII). Significant differences in expression (indicated by yellow shading) among the libraries are indicated in (D–F), with p-values shown as obtained from a Student's t test.

short peptides, or endowed with peptide libraries to find specific binding information via a phage display directed evolution screening method (*2, 48, 49*). We have previously demonstrated the use of such genetically engineered phage for a tissue engineering scaffold application (*33*). As specific phage interaction with the cells was the main goal, a rational design approach was chosen to display well-known RGD and IKVAV cell signaling motifs (*39−41*). Additionally, a peptide expression on every copy of pVIII was preferred over a more sparse phagemid system display (*2, 34*), in order to obtain a high density of the signaling motifs ($\sim 1.5 \times 10^{13}$ epitopes/cm^2) (*33*). Surfaces displaying such a high density of peptide motifs have been shown to promote a stronger cellular response ranging from attachment to differentiation (*39, 40*).

Whereas the sequence of the wild-type pVIII protein has naturally evolved to be a good fit for all of the biological interactions and processes involved in phage replication, any engineered protein additions or modifications have to pass through this imposed censorship prior to expression (see Figure 1). PVIII is a product of M13 phage gene gVIII and begins its existence through transcription and translation by the *E. coli* cell machinery. PVIII is a very high copy number protein, with 0.5 to 5 million units being transcribed and translated every

cell division cycle, and 2700 units are needed to package every virus before it is able to leave the cell (*19, 50*). The full protein, called procoat, is 73 amino acids in length, and contains a signal and mature peptide regions (Figure 1C). Positive Lys charges at both its amino and carboxy ends target the procoat molecule to the phospholipid membrane by electrostatic attraction (*51*) (Figure 1D). Once at the membrane, the hydrophobic segments located in the signal and mature portions of the procoat protein partition into the lipid bilayer. YidC transporter protein and electrophoretic potential are responsible for inserting the connecting loop into the periplasm (*22, 23, 52, 53*). Leader peptidase is a protein that is responsible for releasing the signal peptide sequences from many bacterial transmembrane proteins. The cleavage site utilized by the peptidase is dictated by the sequence of the substrate protein, its position in relation to the membrane interface, and the quality of fit to the peptidase (*54−56*). The pVIII procoat is one of the proteins cleaved by the leader peptidase, releasing the N-terminus of the mature peptide (*16, 17, 54*). After a sufficient concentration of the pVIII protein has been processed and collected in the *E. coli* inner membrane, the pVIII units pack around the extruding phage DNA strand (*23*). The phage extrusion process is orchestrated by pI and pXI phage proteins and an *E. coli* protein threoxin.
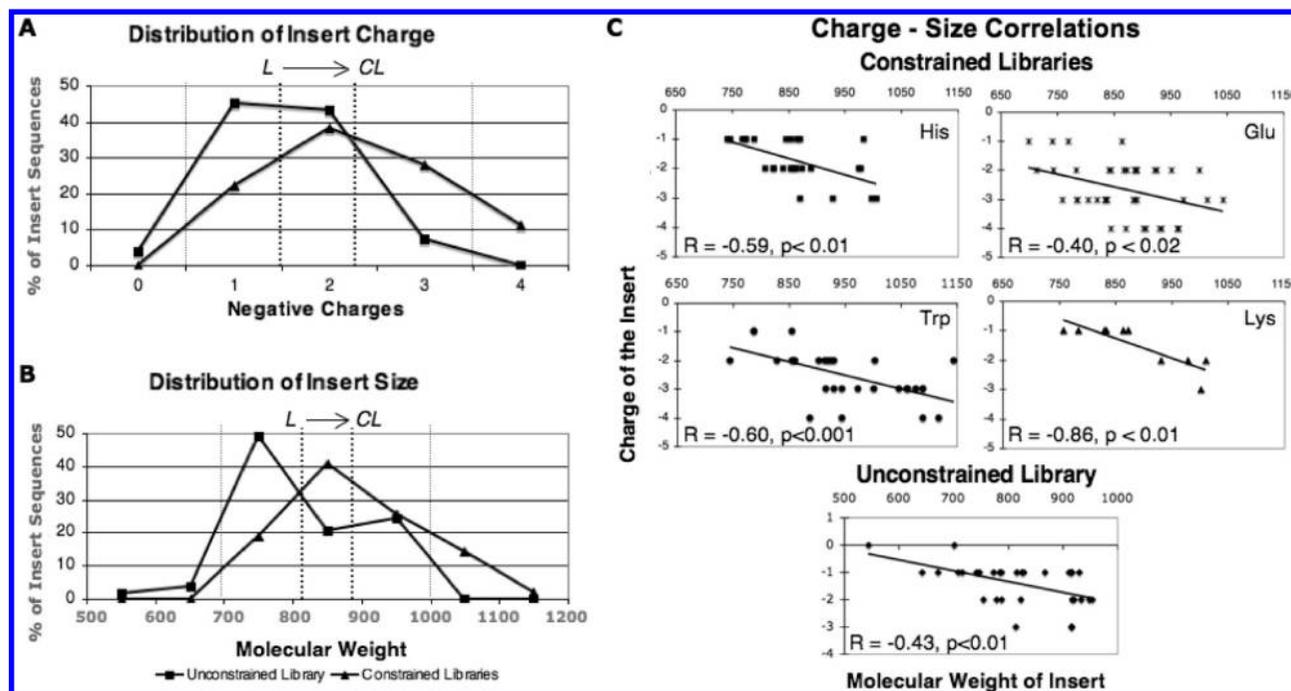
Figure 5. The charge and size of the expressed pVIII library inserts. (A) Charge and (B) molecular weight distributions of the pVIII unconstrained library (■) were compared to the partially constrained insert libraries (▲). The light dashed lines indicate the preferred size and charge for both libraries, whereas the dark dashed lines show a shift of average charge and average size from the unconstrained library (marked by *L*) to the partially constrained libraries (marked by *CL*). (C) An inverse relationship of insert charge to insert size is represented by a linear regression for all of the constructed libraries. The correlation coefficient *R*, and the statistical significance value *p*, is indicated for each of the libraries.

As the phage leaves the cell, it moves through a complex formed by the three above-mentioned proteins from the inner membrane, as well as through a multimeric pore formed by the pIV phage proteins in the outer membrane (*14*). The extrusion is propelled by the packing of the pVIII molecules around the phage DNA (*57*) and by the energy derived from the ATP hydrolysis and proton motive force (*58*). The phage reenters the cell via infection, which depends on the binding event between the pIII minor phage coat protein to the F-pili of the cell and pVIII interaction with both the TolQRA and pIII proteins (*13*). If the infection is successful, the phage DNA is unpackaged and injected into the cell, restarting the process over again.

To find the full insert sequence that would be expressed on pVIII protein, yet contain the desired RGD motif, we used a partially constrained library approach described above (see Figure 1). Although similarly constrained libraries are often used in phage display approaches to improve the binding of the identified short motif sequences (*59, 60*), in our application it was utilized as a means to display the initially unfavorable RGD peptide by allowing the unconstrained library residues to be controlled and chosen by the *E. coli* host. In this manner, the full inserts that were expressed contained both the designed constrained motif and a library chosen through evolutionary selection processes by the *E. coli* cellular mechanisms. This biologically chosen favorability of the insert would then allow for resulting phage particles to preserve their sequence during large-scale amplifications, which could then be used to build structured phage materials.

As seen in previous studies investigating pVIII peptide libraries, when attempting to express IKVAV and RGD cell signaling peptides, or later, Lys constrained libraries on M13 pVIII protein, we succeeded only when the overall charge of the insert was lower than or equal to −1 (*3*). Whereas Lys residues are expressed easily by the bacteria at the signal peptide end as well as the C-terminus of pVIII (Figure 1), their apparent lack in the periplasmic loop, which later becomes the N-terminus

of the mature pVIII is well-documented (*22, 25, 34*). This distribution of positive charges on pVIII can be explained by their function within the phage replication process. As explained above, the presence of positive residues on either end of the procoat protein is needed for the electrostatic binding and anchoring of the protein to the phospohlipid membrane interface of bacteria, and then later to the phage DNA. Their lack in the periplasmic loop can be attributed to the electrophoretic gradient across the bacterial cell membrane that drives the negatively charged portions of the protein across and keeps the positively charged portions in the cytoplasm (*22, 45, 47, 61*).

Further application of the partial library method allowed us to express a TNT binding peptide WHWQ on pVIII (*44*). With expression on all 2700 copies, the WHWQ M13 phage has a very high avidity, likely improving its utilization as a sensor material (*44*). The bulky aromatic peptide motif could only be displayed when surrounded by two to three negative residues, corresponding to an observed increased display of negative residues in the aromatic Trp constrained libraries. The under-representation of aromatic residues at the N-terminus portion of M13 pVIII and other mature peptides of bacterial proteins has previously been observed (*25, 55, 56*). Potentially, this trend can be attributed to the "anti-snorkeling" properties of the aromatic residues located just outside of the membrane interface. Anti-snorkeling refers to the tendency of the aromatic residues to bury their side groups in the membrane interface region, facing the aromatic portion toward the hydrophobic lipid chains and the polar groups toward the negative phospholipid layer (*46, 47, 61−63*).

In the wild-type pVIII, the leader peptidase cleavage site is framed by a transmembrane segment of the signal peptide and the acidic periplasmic loop. A hydrophobic Phe residue at position −2 is likely buried in the membrane, while three negatively charged hydrophilic residues at positions +2, +4, and +5 serve as a hydrophilic anchor of the periplasmic loop, propelling it out of the membrane and away from the negatively charged membrane phospholipid groups (Figure 1C). These

opposing forces expose the leader peptidase cleavage site in the periplasm, allowing the leader peptidase to cleave the signal peptide and release the mature pVIII protein. The increase of negatively charged residues observed in the unconstrained spaces of the library inserts carrying large aromatic segments (WHWQ, 2W, and 3W) could be necessary to counteract the affinity of such constrained sequences toward the interface region. Thus, it is possible that only the proteins where the cleavage site was sufficiently exposed in the periplasm were cleaved and could be further processed into viable phage. The reverse trend was seen for the very negatively constrained libraries (3E and 4E) where the frequency of aromatic residues was 50−60% higher than expected. This could explain the opposite side of this phenomenon, where the aromatic residues served to counteract the tendency of very hydrophilic Glu amino acids to pull the signal peptide out of the membrane, therefore shifting the position of the cleavage site and making it less accessible for processing (*54*).

Protein translation was another process that was seen to impose strong limitations on the allowed insert sequences. In our library engineering efforts, none of the four designed series were able to display five or six long amino acid repeats, and only the 4E library was successful in producing phage with four repeating consecutive residues. This observation can be attributed to a previously shown dependence of protein translation accuracy and library residue expression on tRNA isoacceptor population in the *E. coli* host (*50, 64*). The repetitive use of amino acids combined with the high copy number of the pVIII protein resulted in draining the pool of the available corresponding tRNA isoacceptors. The necessary recharging of the isoacceptors could have slowed the translation process to a rate that was too slow to keep up with the demand of pVIII proteins necessary to package and extrude the accumulated single-stranded phage DNA, thus causing the host cell death (*50, 64−66*). The more repetitive pVIII phage inserts that were incorporated, translated, and packaged had mutations containing a large number of deletions (data not shown), which is common in long stretches of repeating residues (*67*). Even though the 4E phage still occasionally showed deletions and mutations, it was the only kind that could sustain a four amino acid repeat. This was likely due to the large number of Glu isoacceptors in *E. coli* cells (*50*) and to the favorable negative charge at pVIII N-terminus location.

The inverse charge−size relationship seen within the libraries might be attributed to the increased number of negatively charged groups expressed in the large aromatic group containing inserts. However, this relationship is seen even when the inserts are not aromatic and is most pronounced when the constrained insert is carrying a positively charged residue. The size of the insert presented at the N-terminus of pVIII protein is thought to be partially limited by both the size and the properties of the phage extrusion channel formed by pIV proteins (*12, 35*), measuring 6-9 nm in diameter (*18*). The exact structure of pIV protein has not yet been solved, and the exact mechanism of phage extrusion is unknown. However, a strong homology has been identified between the phage extrusion mechanism and type II secretion proteins, which control macromolecule export from bacterial outer membranes (*19, 68*). Many of the type II secretion proteins, including phage pIV protein, have highly conserved C-terminal domains, composed of strongly amphipathic β-sheet structures. These structures anchor the pore in the membrane with the hydrophobic residues and face the charged residues toward the center of the pore to make a large aqueous channel (*20, 69*). Additionally, it has been shown that transport through these channels depends on both ATP hydrolysis and the proton motive force (*58*). Thus, in our system, the greater hydrophilicity of the phage surface is favorable to reduce the hydration necessary for hydrophobic molecules when traveling through a hydrophilic pore (*70*), and the negative charge is favorable to help the phage molecule move down the proton motive force gradient.

In summary, we have described a novel partially constrained library cloning method that allows for the expression of short functional peptide motifs on all of the phage major coat proteins. Our cloning method works through adaptation of the flanking library residues by host induced natural selection, to make the full protein favorable for phage replication. Furthermore, we analyzed the ability of such 8mer insert partial library systems to display positive, negative, and hydrophobic groups, as well as the relationship between the insert charge and its size. Such host−pathogen interaction analysis could be combined with a recently developed virus protein association assay to further our knowledge of general phage biology, and potentially identify novel targets for antibacterial drug development (*71*).

The utilization of such a library engineering approach would be useful in any application where a short functional peptide motif is known, and the current phage display methodology might not provide superior information (*72, 73*). This approach allows not only to express a desired functional group on every copy of the major coat protein of the phage, but also to identify the phage particles that can preserve their sequence during large-scale amplification. Thus, it would be beneficial in a variety of materials applications where density and uniformity of peptide display are favored, including crystal templation (*10, 27, 28*), cell differentiation (*39, 40*), and sensor platforms (*44*). Overall, the utilization of the novel partial library method to express desired peptide groups on all of the major coat proteins of the M13 bacteriophage can serve to further exploit it as a functional nanobiomaterial.

ACKNOWLEDGMENT

LITERATURE CITED

(1) Petrenko, V. A., Smith, G. P., Gong, X., and Quinn, T. (1996) A library of organic landscapes on filamentous phage. *Protein Eng. 9*, 797–801.

(2) Smith, G. P., and Petrenko, V. A. (1997) Phage display. *Chem. Rev. 97*, 391–410.

(3) Iannolo, G., Minenkova, O., Petruzzelli, R., and Cesareni, G. (1995) Modifying filamentous phage capsid: limits in the size of the major capsid protein. *J. Mol. Biol. 248*, 835–44.

(4) Lee, S. W., Mao, C. B., Flynn, C. E., and Belcher, A. M. (2002) Ordering of quantum dots using genetically engineered viruses. *Science 296*, 892–895.

(5) Huang, Y., Chiang, C. Y., Lee, S. K., Gao, Y., Hu, E. L., De Yoreo, J., and Belcher, A. M. (2005) Programmable assembly of nanoarchitectures using genetically engineered viruses. *Nano Lett. 5*, 1429–1434.

(6) Merzlyak, A., and Lee, S. W. (2006) Phage as templates for hybrid materials and mediators for nanomaterial synthesis. *Curr. Opin. Chem. Biol. 10*, 246–252.

(7) Petrenko, V. A., Smith, G. P., Mazooji, M. M., and Quinn, T. (2002) alpha-helically constrained phage display library. *Protein Eng. 15*, 943–950.

(8) Dogic, Z., and Fraden, S. (2006) Ordered phases of filamentous viruses. *Curr. Opin. Colloid Interface Sci. 11*, 47–55.

(9) Lee, S. W., Wood, B. M., and Belcher, A. M. (2003) Chiral smectic C structures of virus-based films. *Langmuir 19*, 1592–1598.

(10) Nam, K. T., Kim, D. W., Yoo, P. J., Chiang, C. Y., Meethong, N., Hammond, P. T., Chiang, Y. M., and Belcher, A. M. (2006) Virus-enabled synthesis and assembly of nanowires for lithium ion battery electrodes. *Science 312*, 885–888.

(11) Wang, Y. A., Yu, X., Overman, S., Tsuboi, M., Thomas, G. J., and Egelman, E. H. (2006) The structure of a filamentous bacteriophage. *J. Mol. Biol. 361*, 209–215.

(12) Makowski, L. (1993) Structural constraints on the display of foreign peptides on filamentous bacteriophages. *Gene 128*, 5–11.

(13) Click, E. M., and Webster, R. E. (1998) The TolQRA proteins are required for membrane insertion of the major capsid protein of the filamentous phage f1 during infection. *J. Bacteriol. 180*, 1723–8.

(14) Feng, J. N., Model, P., and Russel, M. (1999) A trans-envelope protein complex needed for filamentous phage assembly and export. *Mol. Microbiol. 34*, 745–755.

(15) Hohn, B., Lechner, H., and Marvin, D. A. (1971) Filamentous bacterial viruses 0.1. DNA synthesis during early stages of infection with Fd. *J. Mol. Biol. 56*, 143.

(16) Kuhn, A., and Wickner, W. (1985) Isolation of mutants in M13 coat protein that affect its synthesis, processing, and assembly into phage. *J. Biol. Chem. 260*, 15907–13.

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

(18) Opalka, N., Beckmann, R., Boisset, N., Simon, M. N., Russel, M., and Darst, S. A. (2003) Structure of the filamentous phage pIV multimer by cryo-electron microscopy. *J. Mol. Biol. 325*, 461–470.

(19) Russel, M., Linderoth, N. A., and Sali, A. (1997) Filamentous phage assembly: Variation on a protein export theme. *Gene 192*, 23–32.

(20) Marvin, D. A., Welsh, L. C., Symmons, M. F., Scott, W. R., and Straus, S. K. (2006) Molecular structure of fd (f1, M13) filamentous bacteriophage refined with respect to X-ray fibre diffraction and solid-state NMR data supports specific models of phage assembly at the bacterial membrane. *J. Mol. Biol. 355*, 294–309.

(21) Samuelson, J. C., Jiang, F. L., Yi, L., Chen, M. Y., de Gier, J. W., Kuhn, A., and Dalbey, R. E. (2001) Function of YidC for the insertion of M13 Procoat protein in Escherichia coli - Translocation of mutants that show differences in their membrane potential dependence and Sec requirement. *J. Biol. Chem. 276*, 34847–34852.

(22) Schuenemann, T. A., Delgado-Nixon, V. M., and Dalbey, R. E. (1999) Direct evidence that the proton motive force inhibits membrane translocation of positively charged residues within membrane proteins. *J. Biol. Chem. 274*, 6855–6864.

(23) Stopar, D., Spruijt, R. B., Wolfs, C. J., and Hemminga, M. A. (2003) Protein-lipid interactions of bacteriophage M13 major coat protein. *Biochim. Biophys. Acta 1611*, 5–15.

(24) Smith, G. P. (1988) Filamentous phage assembly - morphogenetically defective-mutants that do not kill the host. *Virology 167*, 156–165.

(25) Kuzmicheva, G. A., Jayanna, P. K., Sorokulova, I. B., and Petrenko, V. A. (2009) Diversity and censoring of landscape phage libraries. *Protein Eng., Des. Sel. 22*, 9–18.

(26) Rodi, D. J., Soares, A. S., and Makowski, L. (2002) Quantitative assessment of peptide sequence diversity in M13 combinatorial peptide phage display libraries. *J. Mol. Biol. 322*, 1039–52.

(27) Yamashita, I. (2008) Biosupramolecules for nano-devices: biomineralization of nanoparticles and their applications. *J. Mater. Chem. 18*, 3813–3820.

(28) Uchida, M., Klem, M. T., Allen, M., Suci, P., Flenniken, M., Gillitzer, E., Varpness, Z., Liepold, L. O., Young, M., and Douglas, T. (2007) Biological containers: Protein cages as multifunctional nanoplatforms. *Adv. Mater. 19*, 1025–1042.

(29) Bermudez, H., and Hathorne, A. P. (2008) Incorporating stimulus-responsive character into filamentous virus assemblies. *Faraday Discuss. 139*, 327–35, discussion 399−417, 419−20.

(30) Krag, D. N., Shukla, G. S., Shen, G. P., Pero, S., Ashikaga, T., Fuller, S., Weaver, D. L., Burdette-Radoux, S., and Thomas, C. (2006) Selection of tumor-binding ligands in cancer patients with phage display libraries. *Cancer Res. 66*, 8925–8925.

(31) Hajitou, A., Trepel, M., Lilley, C. E., Soghomonyan, S., Alauddin, M. M., Marini, F. C., Restel, B. H., Ozawa, M. G., Moya, C. A., Rangel, R., Sun, Y., Zaoui, K., Schmidt, M., von Kalle, C., Weitzman, M. D., Gelovani, J. G., Pasqualini, R., and Arap, W. (2006) A hybrid vector for ligand-directed tumor targeting and molecular imaging. *Cell 125*, 385–398.

(32) Frenkel, D., and Solomon, B. (2002) Filamentous phage as vector-mediated antibody delivery to the brain. *Proc. Natl. Acad. Sci. U.S.A. 99*, 5675–9.

(33) Merzlyak, A., Indrakanti, S., and Lee, S. W. (2009) Genetically engineered nanofiber-like viruses for tissue regenerating materials. *Nano Lett. 9*, 846–52.

(34) Imai, S., Mukai, Y., Takeda, T., Abe, Y., Nagano, K., Kamada, H., Nakagawa, S., Tsunoda, S., and Tsutsumi, Y. (2008) Effect of protein properties on display efficiency using the M13 phage display system. *Pharmazie 63*, 760–764.

(35) Rodi, D. J., Mandova, S., Makowski, L. (2005) Filamentous Bacteriophage Structure and Biology, in *Phage Display in Biotechnology and Drug Discovery* (Sidhu, S. S., Ed.) p 748, CRC Press Taylor & Francis Group, Boca Raton, FL.

(36) Chen, G., and Courey, A. J. (1999) Generation of epitope-tagged proteins by inverse PCR mutagenesis. *BioTechniques 26*, 814–816.

(37) Sambrook, J., and Russell, D. W. (2001) *Molecular Cloning: A Laboratory Manual*, 3rd ed., CSHL Press.

(38) Noren, K. A., and Noren, C. J. (2001) Construction of high-complexity combinatorial phage display peptide libraries. *Methods 23*, 169–78.

(39) Schense, J. C., Bloch, J., Aebischer, P., and Hubbell, J. A. (2000) Enzymatic incorporation of bioactive peptides into fibrin matrices enhances neurite extension. *Nat. Biotechnol. 18*, 415–9.

(40) Silva, G. A., Czeisler, C., Niece, K. L., Beniash, E., Harrington, D. A., Kessler, J. A., and Stupp, S. I. (2004) Selective differentiation of neural progenitor cells by high-epitope density nanofibers. *Science 303*, 1352–5.

(41) Pierschbacher, M. D., and Ruoslahti, E. (1984) Cell attachment activity of fibronectin can be duplicated by small synthetic fragments of the molecule. *Nature 309*, 30–3.

(42) Sephel, G. C., Tashiro, K. I., Sasaki, M., Greatorex, D., Martin, G. R., Yamada, Y., and Kleinman, H. K. (1989) Laminin A chain synthetic peptide which supports neurite outgrowth. *Biochem. Biophys. Res. Commun. 162*, 821–9.

(43) Ruoslahti, E. (1996) RGD and other recognition sequences for integrins. *Annu. Rev. Cell Dev. Biol. 12*, 697–715.

(44) Jaworski, J. W., Raorane, D., Huh, J. H., Majumdar, A., and Lee, S. W. (2008) Evolutionary screening of biomimetic coatings for selective detection of explosives. *Langmuir 24*, 4938–43.

(45) Zilberstein, D., Agmon, V., Schuldiner, S., and Padan, E. (1984) Escherichia coli intracellular pH, membrane potential, and cell growth. *J. Bacteriol. 158*, 246–52.

(46) Adamian, L., Nanda, V., DeGrado, W. F., and Liang, J. (2005) Empirical lipid propensities of amino acid residues in multispan alpha helical membrane proteins. *Proteins 59*, 496–509.

(47) Stopar, D., Spruijt, R. B., and Hemminga, M. A. (2006) Anchoring mechanisms of membrane-associated M13 major coat protein. *Chem. Phys. Lipids 141*, 83–93.

(48) Whaley, S. R., English, D. S., Hu, E. L., Barbara, P. F., and Belcher, A. M. (2000) Selection of peptides with semiconductor

binding specificity for directed nanocrystal assembly. *Nature 405*, 665–668.

(49) Lee, S. K., Yun, D. S., and Belcher, A. M. (2006) Cobalt ion mediated self-assembly of genetically engineered bacteriophage for biomimetic Co-Pt hybrid material. *Biomacromolecules 7*, 14–7.

(50) Rodi, D. J., and Makowski, L. (1997) In *Proceedings of the 22nd Tanaguchi International Symposium*, p 155.

(51) Kuhn, A., Wickner, W., and Kreil, G. (1986) The cytoplasmic carboxy terminus of M13 procoat is required for the membrane insertion of its central domain. *Nature 322*, 335–9.

(52) Russel, M. (1993) Protein-protein interactions during filamentous phage assembly. *J. Mol. Biol. 231*, 689–697.

(53) Samuelson, J. C., Chen, M. Y., Jiang, F. L., Moller, I., Wiedmann, M., Kuhn, A., Phillips, G. J., and Dalbey, R. E. (2000) YidC mediates membrane protein insertion in bacteria. *Nature 406*, 637–641.

(54) Jain, R. G., Rusch, S. L., and Kendall, D. A. (1994) Signal peptide cleavage regions. Functional limits on length and topological implications. *J. Biol. Chem. 269*, 16305–10.

(55) Choo, K. H., Tong, J. C., and Ranganathan, S. (2008) Modeling Escherichia coli signal peptidase complex with bound substrate: determinants in the mature peptide influencing signal peptide cleavage. *BMC Bioinf. 9* (Suppl. 1), S15.

(56) Choo, K. H., and Ranganathan, S. (2008) Flanking signal and mature peptide residues influence signal peptide cleavage. *BMC Bioinf. 9* (Suppl. 12), S15.

(57) Papavoine, C. H. M., Christiaans, B. E. C., Folmer, R. H. A., Konings, R. N. H., and Hilbers, C. W. (1998) Solution structure of the M13 major coat protein in detergent micelles: A basis for a model of phage assembly involving specific residues. *J. Mol. Biol. 282*, 401–419.

(58) Feng, J. N., Russel, M., and Model, P. (1997) A permeabilized cell system that assembles filamentous bacteriophage. *Proc. Natl. Acad. Sci. U.S.A. 94*, 4068–73.

(59) Brissette, R., and Goldstein, N. I. (2007) The use of phage display peptide libraries for basic and translational research, In *Methods in Molecular Biology* (Fisher, P., Ed.) pp 203−213, Humana Press, Inc., Totowa, NJ.

(60) Li, R., Hoess, R. H., Bennett, J. S., and DeGrado, W. F. (2003) Use of phage display to probe the evolution of binding specificity and affinity in integrins. *Protein Eng. 16*, 65–72.

(61) von Heijne, G. (1994) Membrane proteins: from sequence to structure. *Annu. Rev. Biophys. Biomol. Struct. 23*, 167–92.

(62) Liang, J., Adamian, L., and Jackups, R., Jr. (2005) The membrane-water interface region of membrane proteins: structural bias and the anti-snorkeling effect. *Trends Biochem. Sci. 30*, 355–7.

(63) Braun, P., and von Heijne, G. (1999) The aromatic residues Trp and Phe have different effects on the positioning of a transmembrane helix in the microsomal membrane. *Biochemistry 38*, 9778–82.

(64) Elf, J., Nilsson, D., Tenson, T., and Ehrenberg, M. (2003) Selective charging of tRNA isoacceptors explains patterns of codon usage. *Science 300*, 1718–22.

(65) Pratt, D., Tzagoloff, H., and Erdahl, W. S. (1966) Conditional lethal mutants of the small filamentous coliphage M13. I. Isolation, complementation, cell killing, time of cistron action. *Virology 30*, 397–410.

(66) Hohn, B., Schutz, H. V., and Marvin, D. A. (1971) Filamentous bacterial viruses 0.2. killing of bacteria by abortive infection with Fd. *J. Mol. Biol. 56*, 155.

(67) Vilette, D., Ehrlich, S. D., and Michel, B. (1995) Transcription-induced deletions in Escherichia coli plasmids. *Mol. Microbiol. 17*, 493–504.

(68) Genin, S., and Boucher, C. A. (1994) A superfamily of proteins involved in different secretion pathways in gram-negative bacteria: modular structure and specificity of the N-terminal domain. *Mol. Gen. Genet. 243*, 112–8.

(69) Nikaido, H. (2003) Molecular basis of bacterial outer membrane permeability revisited. *Microbiol. Mol. Biol. Rev. 67*, 593–656.

(70) Nikaido, H. (1992) Porins and specific channels of bacterial outer membranes. *Mol. Microbiol. 6*, 435–42.

(71) Zlotnick, A., Lee, A., Bourne, C. R., Johnson, J. M., Domanico, P. L., and Stray, S. J. (2007) In vitro screening for molecules that affect virus capsid assembly (and other protein association reactions). *Nat. Protocols 2*, 490–498.

(72) Brammer, L. A., Bolduc, B., Kass, J. L., Felice, K. M., Noren, C. J., and Hall, M. F. (2008) A target-unrelated peptide in an M13 phage display library traced to an advantageous mutation in the gene II ribosome-binding site. *Anal. Biochem. 373*, 88–98.

(73) Kolb, G., and Boiziau, C. (2005) Selection by phage display of peptides targeting the HIV-1 TAR element. *RNA Biol. 2*, 28–33.

BC900303F