ACS chemical biology

Reviews

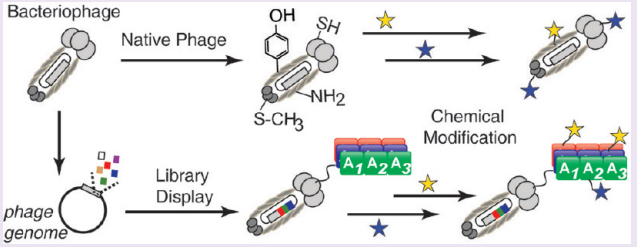
pubs.acs.org/acschemicalbiology

Bacteriophages and Viruses as a Support for Organic Synthesis and Combinatorial Chemistry

Simon Ng, Mohammad R. Jafari, and Ratmir Derda*

Department of Chemistry and Alberta Innovates Centre for Carbohydrate Science, University of Alberta, Edmonton, AB T6G 2G2, Canada

ABSTRACT: Display of polypeptide on the coat proteins of bacteriophages and viruses is a powerful tool for selection and amplification of libraries of great diversity. Chemical diversity of these libraries, however, is limited to libraries made of natural amino acid side chains. Bacteriophages and viruses can be modified chemically; peptide libraries presented on phage thus can be functionalized to yield moieties that cannot be encoded genetically. In this review, we summarize the possibilities for using bacteriophage and viral particles as support for the synthesis of diverse chemically modified peptide libraries. This review critically summarizes the key chemical considerations for on-phage syntheses such as selection of reactions compatible with protein of phage, modification of phage "support" that renders it more suitable for reactions, and characterization of reaction efficiency.



Invention of solid-phase synthesis, in which one reactant is immobilized on a solid support while the other reagents can be easily added and removed, has been recognized by the Nobel Prize in chemistry (Merrifield, 1984). Solid-phase DNA synthesis was critical to most modern advances in molecular biology. Solid-phase synthesis makes it easy to build and handle diverse chemical libraries, and this approach yields a large fraction of lead compounds for drug discovery. Techniques like split-and-pool synthesis can generate one-bead-one-compound (OBOC) libraries that contain large numbers of structurally diverse compounds each attached to a separate bead (Figure 1A).[1] The limitation of OBOC libraries is the ability to identify the library members. The molecules attached to the bead have to be present in sufficient amount for the unambiguous characterization.[2,3]

Molecular biology provides another possibility for building and handling large combinatorial libraries. Cloning a library of DNA molecules into the genome of the host organism, such as bacteriophage, generates a library of organisms. Specifically, in phage display, inserting a library of DNA sequences into the gene of a coat protein physically links the polypeptide gene product to the surface of the bacteriophage (Figure 1B).[4] DNA not only functions as an information tag but also as a "recipe" for host organism to resynthesize the bacteriophages and the phage-displayed compounds. Therefore, even one bacteriophage particle is enough for amplification and genome identification. Genetically encoded libraries can reach diversities of 10^9 in phage or yeast display and of 10^{14} in ribosome and RNA display. Unlike OBOC libraries, genetically encoded libraries are limited to polypeptides or polynucleotides (aptamers).

In this review, we focus on one of the emerging possibilities for combining the advantages of solid-phase synthesis and genetically encoded libraries. From here, we will use "phage" as a general term to represent the bacteriophage particles. Phage can be used as a solid support for multistep organic synthesis. Modification of genetically encoded libraries of peptides displayed on the surface of phage can yield diverse chemical libraries (Figure 1B). Chemical modification on a viral capsid is an established area, and there are several reviews that describe application of modified phage in areas such as sensors,[5] drug delivery,[6] and nanoparticle synthesis.[7,8] This manuscript is aimed to be the first review of the possibilities for multistep synthesis of structurally diverse libraries on phage surfaces. Using the parallel between classical solid-phase synthesis and on-phage synthesis, we overview the key considerations for on-phage syntheses: (i) encoding strategy; (ii) choice of solid support (phage); (iii) choice of reactions and conditions; and (iv) characterization of reaction efficiencies. As part of the review, we describe key examples of multistep organic synthesis on phage particles and synthetic modifications of combinatorial phage-displayed libraries.

WHY PHAGE

Phages are the simplest and the most abundant organisms on our planet. Metagenomic analysis estimates that there are ca. 10^{31} phage on earth (10 times more than bacteria) and 10^6 diverse phage/viral genomes per 1 kg of marine sediments.[9] The particles of phage contain a DNA or RNA genome surrounded by capsid proteins. Owing to the small size of phage, whole genome sequencing[10−12] and even total synthesis of their whole genomes[13,14] were accomplished decades ahead of the simplest bacteria. Phages start to produce their progenies

Received: September 5, 2011
Accepted: October 11, 2011
Published: October 11, 2011

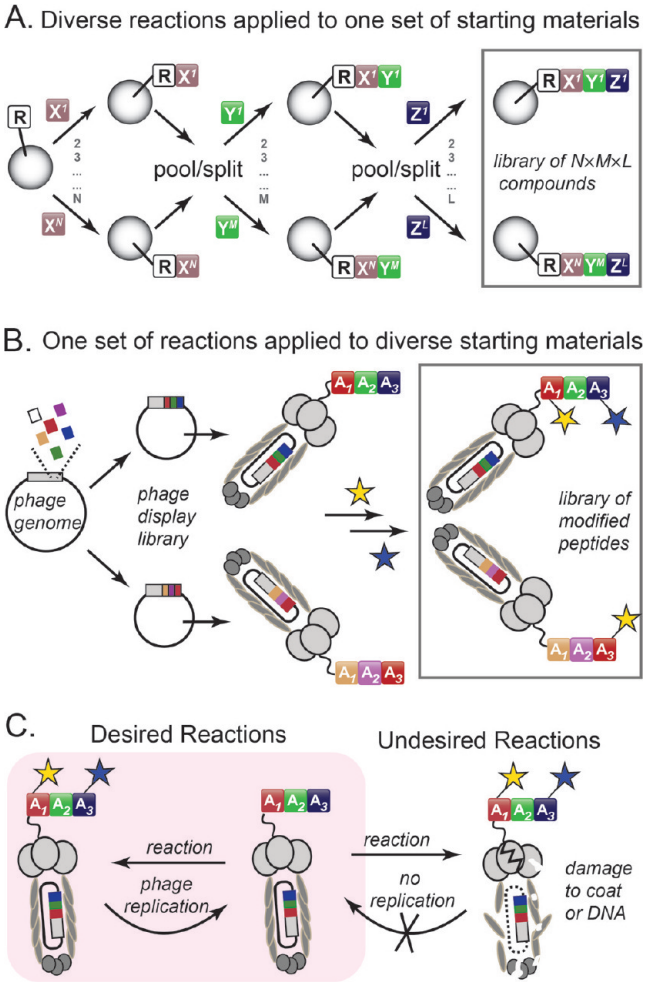


Figure 1. (A) Building diversity using OBOC solid-phase synthesis. R is a starting point for synthesis; X, Y, and Z are derivatives of substrates employed in building diversity. (B) Building diversity by modifying a library of peptides displayed on phage. A_1–A_3 are amino acids encoded in the genome of the phage and displayed on its coat protein. The colored stars are organic molecules that modify specific amino acids. (C) Reactions that modify phage-displayed peptide libraries should not interfere with viability of phage.

as soon as 30 min after the delivery of their DNA to the host organism. Their rapid growth, small size, and ease of manipulation made it simple to study the relationship between DNA sequences (genotype) and functions (phenotype); phage thus was the key tool for establishing the main principles of molecular biology in the second half of the 20th century. The simplicity of phage makes it an attractive biological model system for chemistry/chemical biology communities.

ENCODING STRATEGY

Strategy for Building Library. In OBOC and other solid-phase syntheses, the diversity is built by applying a diverse set of transformations to a limited set of starting materials (Figure 1A). Decoding the structure of the compound requires analysis of the synthesized compounds. The analysis can be facilitated by a chemical or physical information tag present within the support. This tag usually "records" the sequence of the chemical steps. Advanced information tags can route compounds through specific reaction steps[15,16] or bring specific reagents to the reactant.[17,18] In on-phage synthesis, the diversity can be built by applying a constant set of chemical transformations to a diverse set of genetically encoded starting materials (Figure 1B). The compounds, hence, can be decoded from their starting material (polypeptides). In turn, generation of the library of starting materials can be performed using phage-display technology (Figure 1B). Even if the DNA of phage does not directly encode any capsid proteins, differences in phage genome could be used to distinguish phage particles. For example, Woiwode *et al.* used hybridization of phage genome to nitrocellulose support presenting unique complementary DNA regions to generate and track a small library of organic molecules on a phage.[19]

Phage and viral display is a field with over 20 years of history. Display of libraries of polypeptides has been done for M13 phage,[20] T7 phage,[21] lambda phage,[22] baculovirus,[23] adeno-associated virus,[24] RNA phages PP7,[25] MS2,[26] and TBSV.[27] Several M13 and T7-displayed libraries are available as commercial products. Many companies generate custom display platforms on demand.

Requirements for Encoding. There are several considerations for converting a phage into a library display platform: (i) the libraries are made by transformation of DNA plasmids into the host. The diversity of the resulting libraries is equal to the number of hosts that pick up the plasmid. The host, therefore, should be "competent", *i.e.*, undergo highly efficient uptake of DNA plasmids (for an overview of competent bacteria see ref 28). (ii) The peptides should be inserted in locations where they do not interfere with the life cycle of the virus. Preferred location for insertion is the *N*- or *C*-terminus of the coat protein. Insertions into internal loops might lead to undesired conformational constraints. Identification of display position for some viruses can be difficult. For example, in CPMV or FHV viruses, most foreign peptides undergo photolytic cleavage upon viral assembly. (iii) It is important to characterize the diversity of peptide libraries because not every sequence inserted can be displayed effectively.[29,30] In the case of nonlytic filamentous phage, disfavored sequences are those that (i) result from the translation of rare codons,[31] (ii) interfere with electrostatic assembly of particles,[29] or (iii) inhibit export of phage proteins to periplasm.[32,33]

CHOICE OF SOLID SUPPORT

Composition and Size. Support for solid-phase synthesis is usually ca. 100-μm beads made of cross-linked polymer or glass, which present specific reactive groups (*e.g.*, amine, hydroxyl). These reactive groups are used as starting points for synthesis. The beads used for solid-phase synthesis can be easily separated from the reaction by filtration. Submicrometer-size phage and viruses, on the other hand, pass through most filters. Isolation of phage from solution, nevertheless, is relatively simple. Methods such as size exclusion chromatography[34] or precipitation[20] isolate phage in minutes to a few hours. The small size of phage is beneficial because each milliliter of water can contain up to 10^{13} of M13 phage particles. In comparison, one gram of typical 100-μm beads used for solid-phase synthesis contains only 2×10^6 particles. Common viral capsids used for synthesis include simple plant viruses, such as CPMV, and filamentous phage particles, such as M13 phage. CPMV are 30-nm spheres composed of 60 copies each of two different viral proteins, whereas M13 phage are flexible rods about 900 nm long and 6.6 nm wide composed of 5 diverse proteins.

dx.doi.org/10.1021/cb200342h | *ACS Chem. Biol.* 2012, 7, 123–138

Solubility and Stability. Although most phage/viruses are soluble and stable only in water, there are both chemical and genetic ways to surpass this limitation (Figure 2). For example,

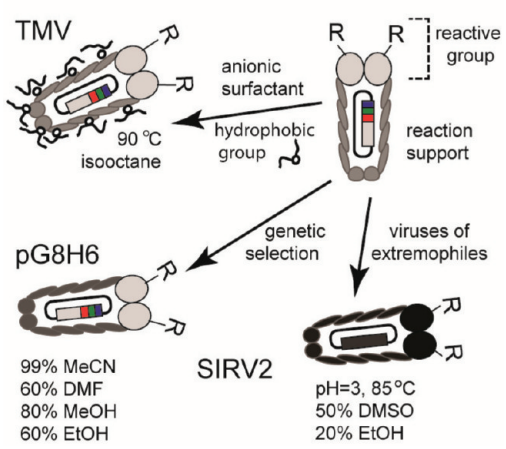


Figure 2. Modifications of phage and viruses that increase their solubility and stability in organic solvents.

chemical modification of viruses by surfactants makes them more soluble and stable in organic solvents.[35] There are examples of coating TMV[36,37] with hydrophobic side chains, although these reports do not identify the viability of virus after such chemical modification. There are several reports of screening for mutants of filamentous phage that have higher stability in organic solvents.[38−40] Identification of these mutants made it possible to perform phage-display screening in a mixture of aqueous and organic solvents.[41]

Perhaps the most interesting viruses that can serve as support for chemical synthesis can come from habitats with high temperature and high concentrations of acid or organic compounds (*e.g.*, oil wells). Steinmetz *et al.* reported the utility of one such virus, *Sulfolobus islandicus* rod-shaped virus 2 (SIRV2), as reaction support.[42] Chemical superiority of thermophile viruses is obvious: they naturally grow at 80 °C and pH 3. They are stable in organic solvents: SIRV2 remains infective after several days in 20% EtOH and 50% DMSO. The coat protein of SIRV2 contains amines and carboxyl groups, along with carbohydrates, which can be oxidized to aldehydes with $NaIO_4$. The virus does not contain cysteine (Cys), and thus the thiol functionality can potentially be inserted in desired locations.

Broad utility of SIRV2, however, hinges on genetic manipulation of its host *S. islandicus*. The culture and genetic manipulation of this organism is one of the simplest among thermophiles: *Sulfolobus* grows aerobically in simple Brock medium. Viruses are detected as plaques on plates prepared with Gelrite gum, a polysaccharide that does not melt at 80 °C. Viral DNA can be engineered in *E. coli* and introduced into the host *via* electroporation with efficiency of 10^2−10^6 transformants per μg of DNA.[43] These transformants can be visualized by reporter gene, such as *lacS* that codes an enzyme with galactosidase activity. Their selection can be done in the presence of antibiotic hygromycin, which does not degrade in hot acidic medium.[44−46]

Despite the availability of methods for genetic manipulation, display of foreign sequences on thermophile viruses is yet to be accomplished. Until recently, a functional plasmid was available only for one *Sulfolobus* virus. Recently, Douglas, Young, and co-workers established a second genetic system for archeal virus *Sulfolobus* turreted icosahedral virus (STIV). The authors attempted to modify a major coat protein, but their manipulations were lethal for viral replications.[47]

REACTIVE HANDLES ON PHAGE

Unlike relatively inert polymeric support, the coat proteins of phage often contain reactive side chains. Nevertheless, techniques for site-specific and bio-orthogonal modification of proteins make it possible to use phage support for the synthesis of molecules with large structural diversity. The following sections highlight manipulation of phage support by removing or introducing specific reactive groups on proteins to make phage more suitable for new types of chemical reactions. Because M13 phage is the most developed system for peptide display, we will focus on this type of phage where possible.

Unnatural Amino Acids. Incorporation of unnatural amino acids (UAA) into the phage capsids can introduce unique reaction handles. Biosynthesis of phage proteins, however, is a temporally orchestrated process. Incorporation of UAA must not only be efficient but also should not alter the kinetics of viral protein synthesis and their assembly. Increase in rate of UAA incorporation, fidelity of incorporation, and coexpression of multiple UAA was the topic of a recent review.[48−51] Here we briefly overview the key examples of UAA incorporation in phage.

Sec-phage. Incorporation of the natural 22nd amino acid selenocysteine (Sec) into M13 phage has been demonstrated by Noren and colleagues at New England Biolabs (Figure 3A).[52] Walker, Noren, and co-workers then coexpressed glycosyltransferase (MurG) and a Sec handle on the same tip of M13 phage.[53] The authors were able to attach biotin selectively on the Sec handle. They then suggested attaching MurG substrate in the same way for directed enzyme evolution. The Sec-containing libraries are simple to produce but difficult to handle because −SeH is rapidly oxidized. To avoid oxidation, handling of libraries should be done in a glovebox (Noren, personal communication).

Met→Aha. One of the least invasive UAA incorporations is replacement of methionine (Met) by isosteric azidohomoalanine (Aha, Figure 3B).[54] Addition of Aha to Met-free media effectively replaces every Met in phage with Aha and places the azide group in a desired position. Three considerations, however, are important: (i) The phage host must be a methionine auxotroph (organism that does not produce its own Met). (ii) Since the start codon (AUG) is translated as *N*-formyl-Met, Met substitution yields *N*-formyl-Aha. Tirrel and co-workers, however, found that in *E. coli N*-terminal Aha can be efficiently deformylated by *E.coli* peptide deformylase and further removed by endogenous methionyl aminopeptidase (MetAP), provided that the second *N*-terminal residue is Ala.[55] (iii) For azide to be a unique functional group in the phage, endogenous Met should be removed (Figure 3B). This removal should be simple because Met-Leu or Met-Ile mutations are well tolerated. Incorporation of Aha into M13 has been reported a few years ago (Connor, R. E., Chapter 5 of Ph.D. Thesis, CalTech, 2008); however, the results have not yet been published in a peer-reviewed publication.

Nonsense Suppression. Another possibility for incorporation of reactive handles is incorporation of UAA by nonsense suppression. Schultz and co-workers generated phage displaying azidophenylalanine (AzPhe) over 7 years ago (Figure 3C).[56,57] The utility of these systems is still limited, potentially

dx.doi.org/10.1021/cb200342h|*ACS Chem. Biol.* 2012, 7, 123−138

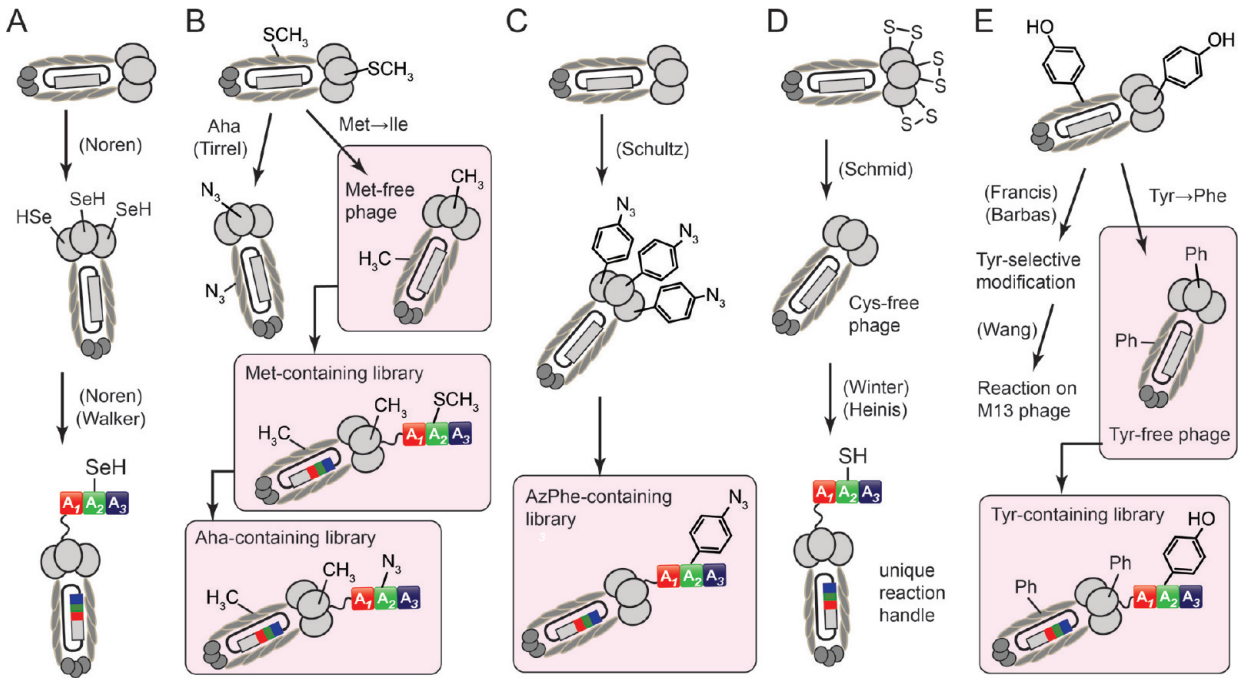


Figure 3. Modification of phage to incorporate unique reaction handles. (A) Incorporation of selenocysteine yields unique reaction handle. (B) Incorporation of azidohomolanine (Aha), which is isosteric to Met. Generation of unique handle using this strategy requires removal of all Met from phage (see Figure 4) and reincorporation of Met only in desired locations (not reported). (C) Incorporation of azidophenylalanine (AzPhe) into M13 by nonsense suppression. (D) Elimination of all Cys and reincorporation of Cys in specific locations. (E) Tyr can be used as reaction handle (see Figure 5). To make it a unique reaction handle, all (solvent-exposed) Tyr have to be mutated. This removal has not been reported, but it could be possible according to mutagenesis studies (see Figure 4).

due to the expression bias against proteins containing UAAs.[58] The AzPhe-containing libraries, for example, have not been used in any published studies. Francis and co-workers reported another example of nonsense UAA incorporation; they produced MS2 phage particles with *p*-amino-L-phenylalanine in their capsids and used this unique reactive handle to perform oxidative coupling.[59]

Side Chains of Natural Amino Acids. Incorporation of unnatural amino acids with orthogonally reactive side chains opens many possibilities, but side chains of natural amino acids could also serve as unique reactive groups (Figure 3D,E). Natural side chains can be rendered "unique" because it is often possible to remove specific amino acids from the viral capsids by mutagenesis (Figure 4). Many of these mutations do not perturb the ability of phage to assemble and replicate. Hence, rather than "expanding the genetic code", it might be possible to "shrink the genetic code"[60] and produce self-replicating phage that contains coat protein without Cys, Lys, Arg, Tyr, Trp, His, or Met.

Cysteines. Cys is one of the least abundant amino acids. *N*-Terminal Cys is usually not found in natural proteins; these residues also undergo reactions that are orthogonal to those of internal Cys (see section on Chemical Modification). Kossiakoff and co-workers used native chemical ligation to modify N-terminal Cys in a M13-displayed library by a peptide thiosters containing unnatural amino acid.[61] Expression of phage-displayed libraries with single N-terminal Cys could be challenging because unpaired Cys disrupts the folding of pIII and reduces phage infectivity.[29] To bypass this problem, *N*-Terminal Cys can be generated after phage production (*in vitro*): (i) The phage coat protein can be expressed as a fusion with intein. Self-splicing generates the phage bearing an *N*-terminal Cys.[62] (ii) Recognition site for proteases, such as tobacco etch virus (TEV) protease can contain Cys at the cleavage site.[63] After expression of the fusion protein and *in vitro* proteolytic cleavage, the *N*-terminal Cys is exposed with high efficiency (95% completion in 24 h) under reducing conditions.[63]

Internal Cys are unique reaction handle in phages that have no natural Cys. Phage with Cys in their proteins can be engineered to be Cys-free even when these residues have important structural roles. Engineering Cys-free pIII protein of M13 phage was difficult due to the unique structural role of disulfides. Folded pIII is necessary for recognition of F-pilus and Toll-like receptor, whereas unfolding in a reducing environment is critical for infection. Nevertheless, Schmidt and co-workers performed rounds of directed evolution and found a stable and functional variant of pIII with N1−N2 domain free of Cys.[64] Subsequently, Heinis, Winter, and co-workers incorporated Cys into this Cys-free phage to introduce a unique reaction handle (Figure 3D).[65]

Using other amino acids, such Tyr, as unique reactive handles requires mutating these residues from the coat proteins (Figure 3E). Sidhu and co-workers used alanine shotgun scan to push minimization of phage to its limits. Their most prominent example is the identification of "mini-pVIII", in which 41 out of 50 amino acids were substituted with Ala or Gly (Figure 4A).[66] The mini-pVIII retains all the structural and functional properties of native pVIII such as α-helix formation and DNA packing. Phage made of mini-pVIII alone, however, are not known. Even phage that form as a hybrid of wild-type pVIII and mini-pVIII are not infective. Still, many reactive residues can be eliminated from pVIII in a way that preserves phage infectivity (Figure 4A, pVIII).[67,68] Similar Ala scan in pIII

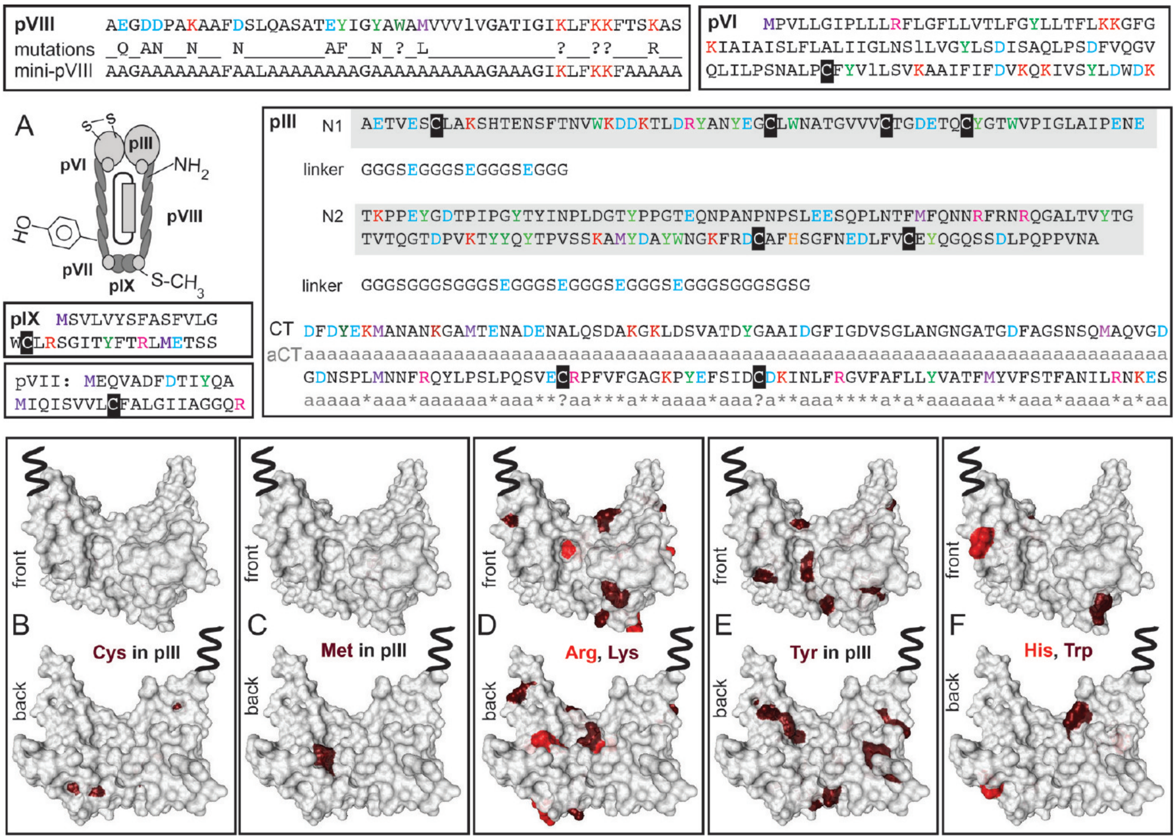


Figure 4. Potential for generation of unique reactive handles on phage by removing reactive amino acids from coat proteins. (A) The basic structure of M13 phage contains ssDNA and coat proteins pVIII, pIII, pVI, pVII, and pIX. Mutagenesis studies in pVIII suggested that many reactive residues can be eliminated. Shotgun alanine studies suggested that most residues could be replaced by Ala to create mini-pVIII. Similar Ala-shotgun study in CT domain of pIII suggested that most residues in this domain could also be mutated. Results are denoted as "aCT". (B−F) X-ray structure of the N1−N2 domain of coat protein pIII (PBD entry 1G3P). Images display location of major reactive residues. Many residues are not solvent-exposed and thus should not be reactive. Black spiral denotes location of N-terminus where peptides are commonly displayed.

protein suggested that the majority of amino acids in the *C*-terminal domain (CT) could be mutated as well (Figure 4A, pIII).[69] "Mini-CT" made of Ala in all nonfunctional positions, however, has not been synthesized yet. Finally, Plückthun and co-workers demonstrated that N1−N2 domains of pIII can be deleted from phage. The phage can still be amplified if the domains are reattached to phage *in vitro* using a noncovalent interaction.[70] Therefore, producing a phage "free of reactive side chains" is not unrealistic. Such phage could be based on functional variants of mini-pVIII, mini-CT, and stripped down versions of other phage proteins.

Development of mini-pVIII demonstrates that it is difficult to eliminate all residues such as Lys from phage while keeping it intact. The role for three *C*-terminal Lys in pVIII coat protein is DNA binding. Lys would play a similar role in the coat proteins of other phage and viruses. A simple replacement of Lys to Arg is not viable.[67,68] It is possible, however, that "global elimination of reactive side chains" might not be necessary. In phage and viruses, side chains that are not exposed to solvent are often hindered from reactions (Figure 4B−F). Specifically, Lys that binds to DNA cannot be modified by many amine-reactive agents (see section on Chemical Modification for specific examples).

Incorporation of Reactive Handle by PTM. Many post-translational modifications (PTM) can be used as unique reactive handles. Carbohydrates on phage can be oxidized to yield aldehydes.[42] Biotinylation machinery can be tricked to incorporate ketone-containing analogues.[71] Incorporation of sequences that can be recognized by specific enzymes is discussed in depth in the section describing DNA-encoded modification.

■ CHEMICAL MODIFICATION

The unifying practical challenge for all reactions on phage is quantitative transformation. High reaction yields are challenging because most of the reactions on natural amino acid side chains have slow kinetics. Yields could be difficult to characterize (see section on Characterization). The balance between reaction time and amount of reagent used is important for two reasons: (i) long reaction time is impractical, and (ii) the use of large excess of reagent could promote undesired side reactions. Reactivity of the same functional group with different locations on the phage may vary as well. One of the reasons is that different steric and chemical environments change the pK_a of the residues.

The following subsections describe transformations of specific amino acids. The overview focuses on multistep transformations (color boxes in Figure 5) and reactions that could be performed independently of one another. Generally, the concentrations of proteins or viruses range from nano- to high micromolar concentrations, while the concentrations of reagents applied typically fall within the high micro- to millimolar range (Figure 5). The concentrations of the reagents used in the publication should provide the reader an idea of

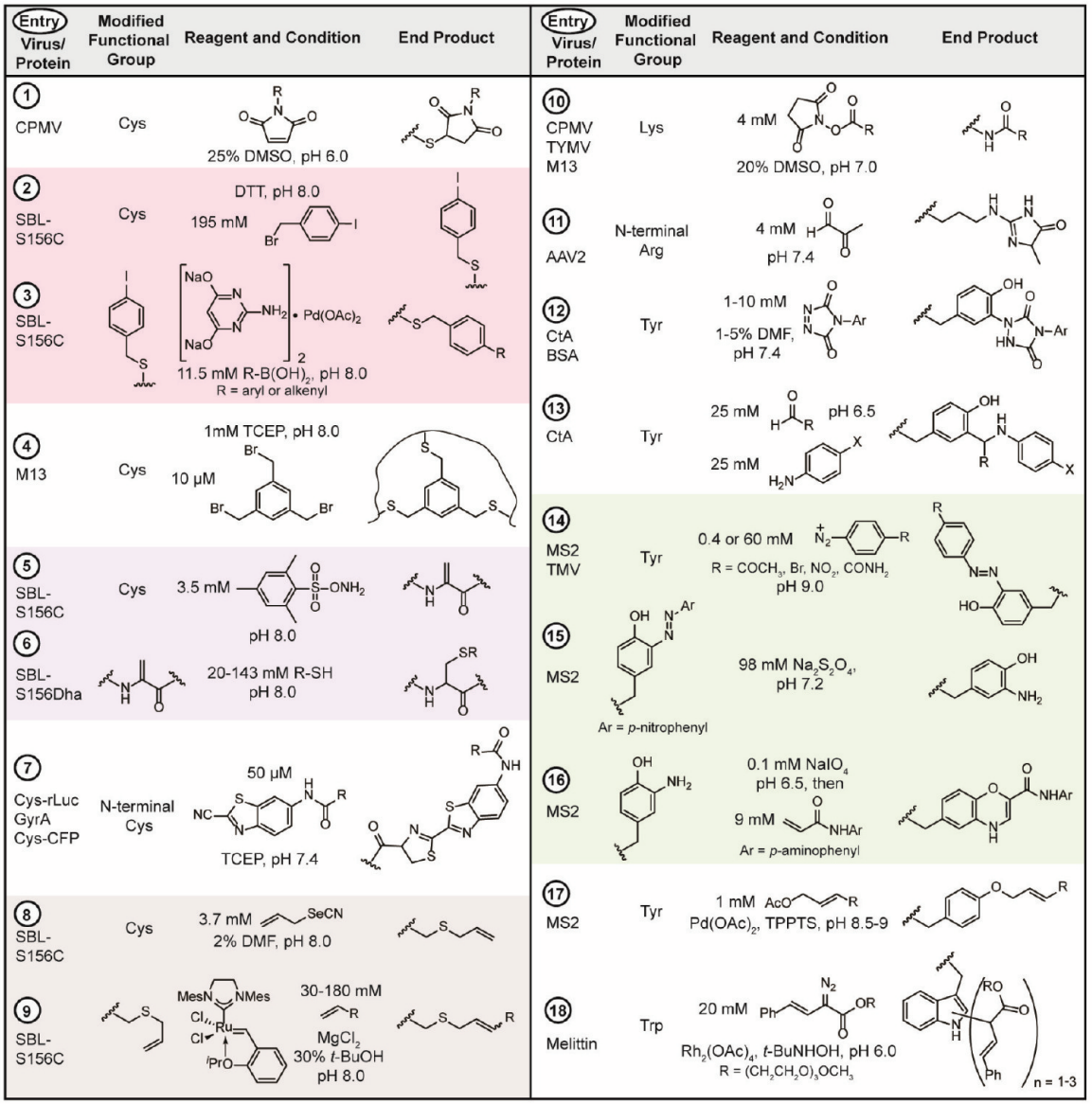


Figure 5. Details of chemical reactions that have been performed on proteins or viruses using side chains of natural amino acids as the reactive handles. Reference for each entry can be found in main text. Colored boxes group multistep transformations.

how efficient these reactions are. Whenever possible, we describe the effect of the local environment and reaction conditions on the efficiency of the reaction on phage.

Cysteine. Michael addition of thiol of Cys to maleimides has been used extensively for protein modification and modification of phage (Figure 5, entry 1).[72] The reaction is selective when pH is maintained below neutrality. Higher pH and excess of reagents tend to promote irreversible alkylation of amino groups.[73] The maleimide derivative is small enough to diffuse across the gap of the CPMV protein assembly and label the Cys located at the interior surface of the capsid proteins. However, introducing Cys on the exterior surface *via* site-directed mutagenesis produced thiols that react significantly faster than the interior residues.[72]

Alkylation of Cys with electrophilic halides is another classical modification. For example, Heinis *et al.* used alkylation to cyclize a library of peptides displayed on M13 phage. The CX_6CX_6C motif displayed on phage can react with tris-(bromomethyl)benzene (TBMB) to constrain the sequences into a bicyclic structure (Figure 5, entry 4).[65] The authors observed that hydrophobic TBMB could penetrate the pIII domain and alkylate the internal Cys (see Figure 4B). This side reaction rendered phage noninfective. Switching to a Cys-free variant of M13 mitigated the problem. Still, side chains of His, Met, Lys, and *N*-terminal amine can compete with the alkylation of Cys.[74] Indeed, the authors observed that TBMB reacts intramolecularly with Lys and *N*-terminal amine in the presence of just one Cys.[65] Also, increasing the concentration of TBMB by 10 times would reduce phage infectivity by ~100-fold, presumably due to cross-linking of the major coat proteins through Lys residues. Unlike Michael addition, this type of chemistry exhibits less chemoselectivity.

Davis and co-workers expanded protein modifications to include Suzuki–Miyaura cross-coupling reaction (Figure 5, entry 3).[75] First, alkylation of Cys installed the orthogonal aryl iodide (Figure 5, entry 2).[75] The authors then performed a cross-coupling in the presence of a palladium complex tolerant to water and oxygen. In the presence of boronic acids, the

dx.doi.org/10.1021/cb200342h | *ACS Chem. Biol.* 2012, 7, 123–138

reaction was very efficient and proceeded to completion within 30 min. Many functional groups were compatible with the reaction condition. Although the thiol of Cys could inhibit the reaction, Cys thiols masked by alkylation or as disulfides were well tolerated. Disulfides were not reduced because the palladium catalyst is phosphine-free. Alcohol of Ser/Thr or amine of Lys could form complexes with boronic acids.[75] These complexes inevitably decrease the effective concentration of boronic acid and might produce an unforeseeable result on the modified residues. Undesired boron complexes, however, could be dissociated if diol or diamine is added after cross-coupling reaction.[75] Due to the size of the catalyst, the accessibility of the aryl iodide and parent Cys residue is critical. The transformation thus could be regioselective, but the authors have not explored this opportunity.

Cys thiol undergoes oxidative elimination to afford dehydroalanine (Figure 5, entry 5) when treated with *O*-mesitylenesulfonylhydroxylamine (MSH).[76] This reaction was effective for SBL-S156C, which contains a sole Cys. ESI-MS analysis demonstrated that no other groups on the protein were modified.[76] However, studies with individual amino acids showed that several functional groups (*e.g.*, Asp, Glu, Met, His, Lys, and *N*-terminal amine) might be aminated by MSH.[77] Although DTT reduction could reverse some of the undesired amination,[77] this treatment might be detrimental to phage that contain disulfide bonds. To circumvent the side reactivity of MSH, Davis and co-workers demonstrated that dibromide derivatives could be superior to MSH in terms of chemoselectivity. The reagent selectively alkylates the thiol of Cys twice and forms the cyclic sulfonium salt that undergoes spontaneous elimination to afford dehydroalanine.[77]

Dehydroalanine, in turn, is a very useful tag and serves as a precursor for Michael addition with a variety of thiol nucleophiles (Figure 5, entry 6).[76] The transformation restores a chiral center to the α-carbon. The diastereoselectivity of the addition depends on the local environment within the peptides. Generation of diastereomers is not a shortcoming but rather an opportunity to increase the diversity of the library. Deciphering the stereochemistry of the "hit" could be easily performed after screening: one could resynthesize two peptides using alkylated *S*- and *R*-Cys and test the isomers separately.

N-Terminal Cys are commonly used in native chemical ligation with peptide thioesters. This opportunity was explored on M13 phage recently.[78] Rao and co-workers observed that *N*-terminal Cys also reacts with 2-cyanobenzothiazole (Figure 5, entry 7).[79] The reaction of internal Cys with 2-cyanobenzothiazole is reversible, whereas *N*-terminal Cys forms a stable condensation product. The reaction has fast kinetics with a second-order rate constant of approximately 9 M^{-1} s^{-1}. It can proceed to completion within a few hours even in the presence of low concentration of reagents.

Davis and co-workers used a Cys handle to perform a two-step modification of proteins; similar steps could be used to modify phage proteins as well. The first step involves a chemoselective installation of an alkene tag on Cys using allyl selenocyanate (Figure 5, entry 8).[80] The allylation is initiated by a S_N2 displacement of the cyanide with thiol, followed by reductive [2,3]-sigmatropic rearrangement to give the allyl sulfide. Unlike the alkylation by allyl halide, this chemistry avoids the undesired modification of *N*-terminus and side chains of Lys, Met, His, or Trp.[74] The authors then employed the installed alkene as a handle for cross-metathesis reaction catalyzed by Hoveyda–Grubbs second generation catalyst (Figure 5, entry 9).[81,82] Diverse alkenes could be introduced at the second step, while allyl sulfide is indispensable for the cross-metathesis.[81] It was proposed that the unique reactivity of allyl sulfide was due to sulfur coordination to the Ru center, thereby enhancing the rate of metathesis by proximity effect.[81] The diversity of alkenes that could be used is limited and is subject to steric hindrance because the ruthenium catalyst is quite bulky. Placing a linker between the allyl sulfides and Cys thiols mitigated the problem and allowed successful cross-metathesis with a variety of allyl glycosides.[82] The drawback of this catalyst is its poor solubility in water. Viruses that are not intrinsically resistant to organic solvents (see Figure 2) might be dissociated under the reaction conditions, in which 30% *tert*-butanol was used. Using purely water-soluble ruthenium catalyst[83] could potentially overcome this problem.

Lysine. Acylation of Lys is a popular bioconjugation strategy (Figure 5, entry 10).[34,84,85] Although these residues are ubiquitous in proteins, some Lys in viruses can be modified regioselectively. The elegant work from Wang *et al.* demonstrated that only 5 out of at least 15 solvent-exposed Lys in each monomer of CPMV were reactive toward the NHS ester of fluorescein and biotin.[34] In TYMV, only 1 out of 6 Lys residues was modified.[84] On the pVIII coat proteins of M13 phage, only *N*-terminus and Lys8 were modified with NHS ester.[86] Lys40, 43, 44, and 48 bound to DNA were unreactive (see Figure 4A for sequence).

Specific chemical environment can also make the ε-amino group of Lys more susceptible to reactions that do not occur with amino groups in other Lys residues. For example, amino groups in Lys are not typically nucleophilic enough to react with 1,3-diketones at neutral pH. Barbas and co-workers, however, used phage display to find a 21-mer peptide that forms a stable enaminone with 1,3-diketone derivatives.[87] This reaction is highly sequence-specific: Lys residues within peptides of similar sequences do no react. This peptide might not be ideal for generation of chemically modified peptide libraries because libraries would have to contain a constant 21-mer region. On the other hand, the reactive peptide is small compared to the size of an average protein. It is thus a useful tag for the covalent labeling of proteins.

The above references highlight only a few examples of unique chemical environment within phage-displayed peptides that render Lys residues uniquely reactive or unreactive; we anticipate that many more uniquely reactive Lys-containing sequences will be discovered.

Arginine. An Arg residue can be modified with methylglyoxal (Figure 5, entry 11).[88] Glyoxal modifies many nucleophiles in proteins reversibly.[89,90] Arg-specific modification plays on the irreversibility of modification of guanidinum moiety. The yields of this modification are less than 30%, and improving the reaction efficiency might be difficult. Any increase in reactivity could backfire and make the reagent more reactive to other nucleophiles. Arg, however, is a rare residue within phage (Figure 4A,D), and the idea of Arg-specific modification is tempting.

Tyrosine. The electron-rich phenolic side chain of Tyr has activated hydrogen at the *ortho* position. The hydrogen is displaced by an electrophilic substrate during electrophilic aromatic substitution. The reactivity of Tyr is very different from that of Cys and Lys, and modification of Tyr could be done in parallel with Cys- and Lys-specific modification.

Barbas and co-workers demonstrated that an azo-derivative of maleimide can target Tyr chemoselectively (Figure 5, entry

12).[91] Notably, Tyr reacts faster than His, Trp, Ser, Lys, and Cys.[91] The unique reactivity toward Tyr can be explained by the reaction mechanism, which involved a seeming [1,5]-hydride shift from the ene component (Tyr phenyl: C=C−OH) to the enophile (azo-maleimide: RN=NR) *via* a concerted and cyclic transition state. The newly forming C−N bond is exceptionally stable toward acid, base, and strong heat. The yields of the bioconjugation are generally good. The cyclic azo-reagents are relatively stable in aqueous media during the course of reaction. However, this is not always the case, as substrate with increased electrophilicity gave inferior reaction yields.[91]

Francis and co-workers described the first example of three-component Mannich-type reaction between aldehyde, aniline, and the Tyr residue of a protein target (Figure 5, entry 13).[92] The reaction involves the formation of an imine between the aldehyde and aniline, followed by electrophilic aromatic substitution at the Tyr residue. Aliphatic amines have higher pK_a than aniline and undergo condensation with aldehyde at a much slower rate under neutral conditions. As a result, the Mannich reaction with aliphatic amine (*e.g.*, Lys) was not observed. Imine formation is the fastest for formaldehyde; the reaction also works for other aldehydes, such as pyruvaldehyde, crotonaldehyde, or 2-furaldehyde. A protonated aniline Schiff base is the key intermediate for Tyr attack. The Schiff bases formed by electron-rich anilines, which have higher pK_a's,[93] gave better yields than electron-poor anilines.[92] A thorough follow-up study from Francis group[94] showed that Trp, if highly accessible to solvent, reacts as equally efficiently as Tyr.

Small formaldehyde molecules could diffuse through coat proteins, cross-link Lys residues, and render phage noninfective. The Mannich reaction to date was performed only on protein targets, and its effect on phage infectivity is not clear. Multicomponent reactions such as the Mannich reaction, however, are very attractive for library construction.

The first example of three-step synthesis on phage was demonstrated by Francis and co-workers (Figure 5, entry 14).[37,95] The synthesis started from functionalization of Tyr with an electrophilic diazonium salt. The azo-tyrosine conjugate was then reduced to cleave the azo linkage (Figure 5, entry 15).[95] Next, the reduced product, *o*-amino-tyrosine, was oxidized *in situ* to a functional diene, which then reacted with dienophile *via* a Diels−Alder reaction (Figure 5, entry 16).[95] The kinetics of the first step is exceptionally fast for a diazonium salt with a strong electron-withdrawing group such as a nitro substituent attached to the aryl ring. Quantitative conversion can be achieved in 15 min using only 5 equiv of the reagent.[95] The reaction is regioselective: in bacteriophage MS2, only one out of four Tyr is modified. His, however, reacts with diazonium compounds as effectively as Tyr.[96] Selective modification of MS2 and tobacco mosaic virus was successfully achieved because their protein sequences do not contain any His residue. The cross-reactivity with diazonium salts implicates the possibility to modify His in viruses that do not possess solvent-exposed Tyr.

Francis and co-workers also demonstrated the first example of a palladium-catalyzed reaction on a virus target.[97] The phenolic side chain of Tyr undergoes allylic alkylation when exposed to allylic acetates and in the presence of palladium catalyst, using triphenylphosphine-3,3′,3″-trisulfonic acid (TPPTS) as a water-soluble phosphine ligand (Figure 5, entry 17).[97] A competing side reaction is the degradation of the π-allyl palladium complex to an unreactive diene. This complication reduced the effective concentration of allylic acetates for subsequent Tyr attack. Generally, 50−65% conversions were achieved under optimal reaction conditions. Nevertheless, palladium-catalyzed allylic alkylation is attractive for several reasons. Besides allylic acetate, a wide range of leaving groups (*e.g.*, carbonates, carbamates)[97] can also be used. Common residues such as Cys and Lys that are prone to alkylation do not participate in the reaction at any appreciable level. Finally, protein with concentrations as low as 5 μM and the capsids of MS2 can be successfully modified.

Tryptophan. There are not many strategies for functionalizing tryptophan with high chemoselectivity. Francis and co-workers developed a novel Trp modification based on *in situ* generated rhodium carbenoids (Figure 5, entry 18).[98,99] This example is one of the first cases of protein modification catalyzed by transition metals.[98,99] Several improvements have been made since its first introduction.[99] Although the yields of reaction are moderate and multiple modifications were observed on a single Trp residue, the reaction has great potential. Trp is the second least common residue in M13 coat proteins (7 out of 633 amino acids available on all the coat proteins). Only two out of four Trp in N1N2 domain of pIII are solvent-exposed (Figure 4F). Due to the size of the rhodium carbenoid intermediate, it might be possible to modify Trp located at the highly solvent-exposed peptides displayed on *N*-terminus of pIII preferentially over Trp that are not solvent-accessible.

Transition-metal-based bioconjugations offer many opportunities to modify otherwise unreactive residues. A recent follow-up of $Rh_2(OAc)_4$ demonstrated its feasibility to functionalize either Gln, Asn, or Phe selectively.[100,101] The reaction happened upon coiled-coil interaction of two complementary α-helixes, one with the unreactive residue and the other with the Rh catalyst noncovalently bound to two glutamate side chains.

Chemical Modifications: Outlook. Viruses or bacteriophages are an assembly of proteins and DNA. Most methods for protein modification are applicable to viruses as well. Interestingly, surveys of all these methods reveal that Tyr/Trp are generally "orthogonal" to Lys/Cys, potentially allowing dual-site modifications to be carried out with exquisite control over undesired cross-reactivity. The advent of chemoselective His modification will be strongly desired as there are only two His, both on pIII, within all of the M13 coat proteins (Figure 4F).

Part of the newly emerging protein modification methods are based on transition metals. It is possible to introduce desired properties, such as steric congestion of reactive center, solubility in water, and enhanced reactivity, by fine-tuning the properties of ligands coordinated to metal.[102] Generally, side chains of amino acids on phage-displayed peptides possess a higher degree of solvent accessibility and flexibility than other residues on the capsid surface of viruses. With this in mind, we anticipate that it should be possible to modify one residue site-selectively over another through the use of sterically congested metal complexes.

GENETICALLY ENCODED MODIFICATION

Natural Inspiration. In natural communities, phage can equip their host with new functions, such as virulence or photosynthetic ability.[103] In phage display, the phage also equip the host with new function, such as galactosidase activity or resistance to antibiotics. Similarly, phage can carry the

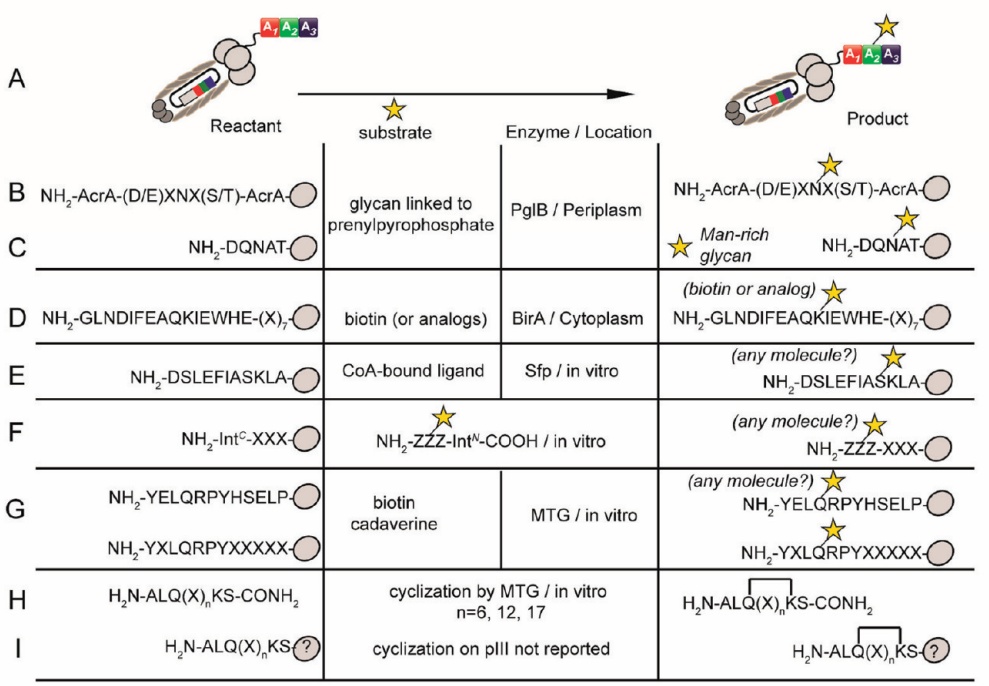


Figure 6. (A) Enzymatic modifications of phage-displayed peptides. (B) Biotinylation of a pIII-displayed library containing AviTag. (C) Glycosylation of library within *C. jejuni* acceptor protein AcrA displayed on pIII.[110] (D) Glycosylation of minimal DQNAT motif displayed in pIII.[113,114] (E) Modification of pIII-displayed peptides by phosphopantenyl transferase Sfp.[108] (F) Modification of pIII-displayed split intein.[78] Int^C and Int^N are *C*- and *N*-terminal domains of split intein. (G) Ligation of biotin-cadaverine to specific pIII-displayed peptides in the presence of MTG. YXLQRPY is a potential consensus sequence for MTG suggested by the authors.[116] (H) Cyclization of peptides by MTG.[117] (I) Cyclization of phage-displayed peptides by MTG could be possible but has not been demonstrated.

instructions for the post-translational modification (PTM) of the peptides displayed on its coat proteins. One such PTM is glycosylation. Although the majority of viruses exploit glycosylation machinery of the host, some viruses carry their own glycosylation instructions in their genome.[104−106] PTMs such as biotinylation, glycosylation, modification by phosphopantenyl transferase Sfp, transglutaminase, and various peptidases have been used to modify M13 phage-displayed libraries of peptides (Figure 6).

Encoding synthetic instructions requires insertions of gene sequences of the modifying enzymes. Filamentous phage are well suited for these purposes because they can tolerate large insertions into their genome. The phage simply become longer in size as the genome increases.

Kinetic and Spatial Considerations for PTM. Co-localization of substrate and enzyme is important. To ensure that enzyme and substrate are in proximity, the modifying enzyme can be co-displayed on the surface of the phage. Although such display is useful for directed evolution of enzymes,[53,107,108] it can be problematic: fusing viral capsid with large protein generally disfavors phage secretion. Many produced particles do not contain any displayed protein.[33]

Alternatively, the enzyme that performs PTM can be present in the cytoplasm or periplasm of the bacteria and modify the phage particles as they are produced. In the latter strategy, timely expression and location of the modification enzymes is important. For example, in classical pIII-display on M13 phage, peptides spend most of their time in the periplasm. For modification, it might be better to use enzymes that are naturally residing in the periplasm. If phage encodes its own instructions for PTM, it is important to think about kinetics of this PTM: (i) Expression and folding of PTM enzyme must be faster than expression of phage coat proteins. (ii) The half-life of the modification must be shorter than the residency time of the phage in the specific location (*e.g.*, cytoplasm or periplasm). In general these kinetic requirements might be difficult to achieve. The problems could be potentially solved by coencapsulation of clonal population of phage and bacteria in an artificial compartment, such as test tube, plaque in agar, or microdroplet.[109,110] Several examples below demonstrate the use of PTM on phage. In majority of the cases, the gene for PTM was already present in the host, or PTM enzyme was used *in vitro*.

M13 Glycophage. One of the PTMs that take place in periplasm is *N*-linked glycosylation. Initially discovered by Szymanski and co-workers in *Campylobacter jejuni*,[111] protein glycosylation locus (*Pgl*) was subsequently cloned in *E. coli*.[112] Glycosylation of phage was performed simply by infecting this *E. coli* with M13 phage displaying (D/E)XNX(S/T) acceptor motif on pIII protein (Figure 6B,C).[113,114] The efficiency of such glycosylation, however, was low: less than 0.1% of the (D/E)XNX(S/T)-containing phage were glycosylated.[114]

Many viruses produced by archaea, including aforementioned STIV and SIRV2 are naturally glycosylated.[42,47] STIV carries its own glycosyltransferase gene, which is important for viability of virus.[47] Archaea platforms might be very interesting for engineering of this PTM because structural diversity of glycosylation in archaea is far greater than in eukaryotes.[115]

Biotinylation. Biotinylation of M13 phage particles that carry biotinylation sequence (AviTag) by BirA is common, but the majority of the reports use this modification after the production of phage ("*in vitro* biotinylation"). Kay and co-workers published the only report, to our knowledge, describing biotinylation of M13 *in vivo* using a SS320 pBirA-

dx.doi.org/10.1021/cb200342h | *ACS Chem. Biol.* 2012, 7, 123−138

Table 1. Methods Used for Characterization of Modification on Phage

entry	method	details of modification	amt/conc	virus/ref
I	inductively coupled plasma atomic emission spectroscopy (ICP-OES)	measures presence of new elements (Gd or Se)[a]	0.1 mg mL^{-1}	Qβ-Virus[135]
			2 mg mL^{-1}	Qβ-Virus[134]
	UV−vis spectroscopy	C_{60}[a]	225 mg	Qβ-Virus[139]
		rhodamine[a]	1.0 mg mL^{-1}	M13[85]
		coumarine[a]	1.0 mg mL^{-1}	TYMV[140]
		fluorescein[a]	1.0 mg mL^{-1}	TYMV[84]
	NMR[b]	purified proteins only	10.8 mg	fd[131]
			5 mg	Pf1[132]
	densitometry	His_6 tag	10 μg	Qβ-Virus[141]
II	transmission electron microscopy (TEM)	detecting structural integrity	2 μg	TYMV[84]
		detecting structural integrity after modification	0.2 mg mL^{-1}	TMV [37]
	scanning transmission electron microscopy	C_{60}	10 ng	Qβ-Virus[139]
III	Western blot	intact phage	10 μg	TYMV[84]
		addition of peptidyl carrier protein	0.1 nmol	M13[123]
		AP-streptavidin	10−20 μg	SIRV2[42]
		anti C_{60}	10−20 μg	Qβ-Virus[139]
	agarose gel	mutated coat protein	1 μg	MS2[142]
		intact phage	10 μg	Qβ-Virus[139]
		polymer-coated capsid	10 μg	Qβ-Virus[143]
	SDS−PAGE	conjugation of fluorescent molecules	1.0 mg mL^{-1}	TYMV[140]
			1 μg	TYMV[84]
			20−50 μg	SIRV2[42]
			10−20 μg	Qβ-Virus[141]
			6−10 μg	PVX[144]
	size-exclusion chromatography	His_6 tag[a]	10 μg	Qβ-Virus[141]
		C_{60}	1 mg mL^{-1}	Qβ-Virus[139]
		oligo(ethylene glycol)	1 mg mL^{-1}	TYMV[140]
		fluorescein	80 μg	CPMV[34]
	ultracentrifuge	fluorescein	500 μg	CPMV[34]
		Intact phage	0.1 mg mL^{-1}	CPMV[145]
		Small molecules	5 mg	TMV[146]
	dynamic light scattering(dls)	C_{60}	0.1 mg mL^{-1}	Qβ-Virus[139]
		polymer-coated capsid	0.1 mg mL^{-1}	Qβ-Virus[133]
		His_6 tag	30 μg	Qβ-Virus[141]
IV	mass analysis (MALDI)	polymer-coated capsid	0.1 mg mL^{-1}	Qβ-Virus[133]
		His_6 tag	30 μg	Qβ-Virus[141]
		small molecules	1−5 mg mL^{-1}	TMV[37]
	MS/MS analysis (ESI)	identifying reactive site of protein	1−5 mg mL^{-1}	TMV[37]
			0.1 mg mL^{-1}	M13[85]
V	zeta-potential analysis	His_6 tag	0.01 mg mL^{-1}	Qβ-Virus[141]
		polycationic peptide	0.01 mg mL^{-1}	Qβ-Virus[147]

[a]The method was used to quantify the amount of modification. [b]Phage was not modified.

expressing bacterial strain.[118] The authors produced a biotinylated peptide library by encoding an Avitag followed by a random 7-mer library at the pIII *N*-terminus. (Figure 6D). The authors observed biotinylation of 50% of the phage particles in optimal conditions (low expression of BirA). Interestingly, overexpression of BirA did not improve efficiency but, instead, decreased phage production. The authors hypothesized that biotinylation impedes the export of pIII to the periplasm.[118]

Biotinylation *in vivo* was performed efficiently in lytic phage that assemble in the cytoplasm of bacteria; for example, phage lambda can be biotinylated by *E. coli* biotin transferase.[119] Eukaryotic viruses were also biotinylated by BirA expressed in the endoplasmic reticulum (ER) of the host. This process has been used to produce biotinylated HIV,[120] baculovirus,[121] and lentivirus.[122]

Sfp. The first demonstration of the proof-of-principle use of Sfp for modification of phage was performed by Walsh and coworkers. The authors displayed the 80-amino-acid long peptidyl carrier protein (PcP) on the surface of the phage. Incubation of PcP-phage with Sfp and small molecules attached to coenzyme A (CoA) yielded a display of molecule on PcP-phage.[123] Subsequent identification of shorter 11- and 12-mer peptides that can be modified by Sfp enzyme made this method more practical.[124,125] These short tags could also be displayed on phage and modified by Sfp *in vitro* (Figure 6E).[108] The yield of Sfp-catalyzed decoration of phage is unclear. The authors visualized the modification by Western blot but did not comment as to what fraction of the phage-displayed substrates had been modified.

Split Inteins. Another interesting modification is the ligation of two peptide or protein segments catalyzed by split

dx.doi.org/10.1021/cb200342h|*ACS Chem. Biol.* 2012, 7, 123−138

intein.[126] Split intein is composed of two fragments, *N*-terminal domain (Int^N) and *C*-terminal domain (Int^C), each fused to two different segments that serve as the precursors for ligation. Split intein can tolerate most functionality on their ligated substrates (extein), provided that the modification does not impose significant constraint on the sequence that immediately flanks the intein. Binding of Int^N and Int^C promotes the ligation; the intein is excised from the ligated product. Recently, Int^C was successfully expressed on a truncated pIII of M13 phage.[78] The truncated pIII has been ligated specifically to a pentapeptide-Int^N bearing biotin or desthiobiotin (Figure 6F).

TGases. Transglutaminases (TGases), enzymes that catalyze amide bond formation between side chains of Gln and Lys, have been used to modify phage-displayed libraries. Specifically, phage display has been used to identify the sequences that form amide bond with biotin cadaverine in the presence of TGase 2, Factor XIIIa,[127] or microbial TGase (MTG, Figure 6G).[116] Interestingly, recent work of Heinis and co-workers suggests that MTG can cyclize peptides of general structure H_2N-ALQ(X)$_n$KS-$CONH_2$. (Figure 6H).[117] The authors, however, tested only a few synthetic peptides with n = 6, 12, and 17 amino acids and only a few sequences.

Outlook. Many PTM are useful for modification of proteins but are not very interesting for library production because the modification site is embedded within a large constant recognition sequences (*e.g.*, GLNDIFEAQKIEWHE for BirA, DSLEFIASKLA for Sfp; for a recent review of other PTMs and their sequences see ref 128). In other PTMs, such as glycosylation, random sequences can be immediately next to a glycosylation site. For example, Aebi and co-workers randomized sequences around (D/E)XNX(S/T) motif to construct a phage-displayed glycopeptide library (Figure 6B).[114] Subsequently, it should be possible to use carbohydrates within this library as unique bioorthogonal handles. The promiscuity of cyclization reactions catalyzed by MTG is promising as well. MTG could be used to produce diverse phage-displayed libraries of cyclic peptides (Figure 6I).[117] Another enzyme, sortase (SrtA), can form covalent bonds to very short sequence LPXTG on any protein.[129] To date, however, there are no reports of the use of sortase for modifying of phage-displayed libraries.

The best PTM enzymes for library production are those that have recognition and modification sites separated from one another. Such recognition property, for example, is known for restriction enzymes: enzyme SfiI cleaves any XXXXX sequence flanked by GGCC sequences. Distal site recognition is known for some enzymes, such as Lacticin 481 Synthetase, which can dehydrate Ser/Thr residues as far as 42 residues downstream of the recognition motif.[130] Engineering of other enzymes to have separate modification and recognition site might be an exciting challenge for protein engineering.

CHARACTERIZATION OF REACTION EFFICIENCY

Characterization of reaction efficiency can be divided into three classes: quantitative, qualitative, and structural. The last can provide not only amount but also the structure of the modification. In organic chemistry, NMR and high-resolution mass spectrometry are the cornerstone techniques used for structural characterization. To the best of our knowledge, there are no reports of characterization modification of phage by NMR. NMR has been used only to determine the structures of wild-type phage proteins.[131,132] Utility of mass spectrometry is limited to detection of modification on proteins present at high concentrations (see below). Therefore, the researchers often have to resort to indirect, qualitative characterization methods.

Commonly used qualitative methods measure changes in global physical properties of phage, such as its density, hydrodynamic radius, surface charge (zeta-potential), or electrophoretic mobility (Table 1). It could be hard to correlate these changes with the exact amount of modifications present on phage. The readout, however, can be very sensitive and can be used to prove that the modification is indeed present on phage. Some techniques can give only partial quantification. For example, if modified and unmodified phage particles have different density, the exact ratio of modified to unmodified phage can be measured by centrifugation or size-exclusion chromatography. The method, however, would not provide any information about the exact number of modifying groups in the modified phage.

It is alarming that in many publications researchers often use techniques that *could* be quantitative without quantifications. Examples are Western blots, modification by fluorescent probes, TEM of immuno-gold labeled samples, and mass-spectrometry. Quantification in these methods is simply a matter of appropriate controls and statistical analysis. Emission of fluorescently labeled phage can be compared to that of the standard. The absorption of different bands on an electrophoresis gel can be quantified by densitometry. Multiple TEM images can be quantified and averaged. Intensity of peaks in mass spectrometry can be compared to that of a calibration curve based on a purified sample.

In every quantification, sensitivity is a major issue. As the instruments and methods are developed to measure smaller amounts of analyte, the research is always one step forward: the need for measuring smaller quantity persists. Producing viruses or phage in large scale is time-consuming and costly, but large amounts of coat proteins are necessary for most characterization techniques. It is therefore not surprising that examples of chemical modifications are most common for phage/viruses that can be purified in milligram to gram quantities (*e.g.*, TYMV,[84] TMV,[37] and Qβ[133]). In Table 1, we list the amount and concentration of proteins of phage or viral particles used for characterization.

Many viruses and phage contain multiple copies of a single structural protein. The presence of multiple reaction sites on the phage can dramatically simplify quantification.[85] For example, M13 capsid contains 2700 copies of the pVIII protein and only 5 copies of pIII. A saturated solution of M13 phage contains 10^{13} pfu mL^{-1} (ca. 17 nM). Thus the concentration of pVIII is about 46 μM, whereas the concentration of pIII is only about 85 nM. In MS analysis of disrupted M13 phage, only pVIII protein can be detected, as the *concentration* of pIII is beyond the limit of detection of most MS instruments.

The hurdles of synthesis on phage are reminiscent of problems in characterization of surfaces, which is also hampered by low concentrations of analytes and a high amount of noise from the non-surface species. In surface science, XPS and other techniques that quantify incorporation of new elements into the surface are the workhorses of the field. Similar methods could quantify introduction of new elements into the phage or virus. For example, incorporation of metals (*e.g.*, Gd)[134] or incorporation of amino acids with "unnatural elements" (*e.g.*, selenomethionine)[135] has been detected by Finn and co-workers using ICP-OES. Halogens are not naturally present in phage, yet they are common in bioactive organic molecules and therapeutic compounds. Synthesis and

dx.doi.org/10.1021/cb200342h | *ACS Chem. Biol.* 2012, 7, 123–138

characterization of halogen-containing molecules on phage should be simple with an elemental analysis technique like ICP-OES.

Other methods commonly used by surface scientists are sum-frequency (SFG) or second harmonic generation (SSG) spectroscopies, which detect surface-immobilized molecules only. These techniques can visualize crystalline proteins in heterogeneous media, such as living cells.[136] To our knowledge, there are only a few reports of SFG or SHG detection of viruses.[137] Detecting viruses by SFG exploited the same symmetry breaking rules that enable detection of surface-immobilized molecules in a sea of analogous molecules in the bulk. Theoretically, this technique should also work for chiral cylindrical structures,[138] such as filamentous phage.

Quantitative structural characterization of the chemically modified phage and virus remains a challenge. Many chemistry journals, such as *The Journal of Organic Chemistry* (JOC), have a standard for characterization: NMR and high-resolution MS must be presented, where applicable, for publication. In on-phage synthesis, no single method is sufficient to prove the incidence and/or amount of modification. Still, devising a general JOC-like set of standards for on-phage synthesis would be desirable. Table 1 could suggest a set of methods that could be used in the future as part of the "standard characterization set".

In the absence of an experimental standard, researchers often pick methods according to their availability (*e.g.*, MS, UV−vis spectroscopy) and ease of use (*e.g.*, SDS-PAGE, UV−vis spectroscopy). The researchers interested in using phage as synthetic support could also learn from surface science and other areas in which structural characterization of low abundance analytes is a common problem.

CONCLUSION

To date there are only three examples of library synthesis that unite the power of phage display and chemical modification of phage.[61,65,148] A few more reports demonstrated chemical modification of genetically encoded libraries displayed on RNA. These examples were discussed in the recent minireview by Angelini and Heinis.[149] Although examples of synthesis of diverse libraries by modification of genetically encoded libraries are still rare, this review demonstrates that there are many unexplored approaches for generation of these libraries. Phage are tolerant to drastic modification to coat proteins and large insertions in the genome. The phage genome could carry the information about starting materials (peptides); it can also carry the gene of enzymes that modify these peptides post-translationally. Since phage are made of proteins, a growing body of chemistry for modifying protein will be the key tool for generating diverse libraries on phage. The major hurdle in synthesis on phage will be characterization of the exact reaction yield and structural details of products displayed on individual phage particles. Most techniques require high concentration of phage proteins. Characterizations of "one-phage-one-compound" libraries would be very challenging but might not be necessary for the discovery process. Modifications could be characterized in clonal populations containing a large number of genetically identical phage particles.

Minimizing the phage particle to make it suitable for broad range of chemical modifications is attractive but, still, a challenging task. Mutagenesis studies suggest, however, that generating chemically inert phage is not unrealistic. Exploring Nature's solutions to this problem might be even more fruitful: many interesting phage resistant to organic solvents and high temperature came out of extreme habitats. Metagenomic studies suggest that millions more phage await discovery.

AUTHOR INFORMATION

Corresponding Author

*E-mail: ratmir.derda@ualberta.ca.

ACKNOWLEDGMENTS

We thank Dr. Mickey Richards and Gary Freschauf for critical comments on the manuscript. This work was supported by funds from University of Alberta, Alberta Innovates Centre for Carbohydrate Science, NSERC Discovery Grant 402511 and Grand Challenges Canada "Rising Star in Global Health" award (to R.D.).

ABBREVIATIONS

AAV2: adeno-associated virus serotype 2; BSA: bovine serum albumin; CPMV: cowpea mosaic virus; CtA: Chymotrypsinogen A; Cys-CFP: N-terminal cysteine cyan fluorescent protein; Cys-rLuc: N-terminal cysteine luciferase; Dha: dehydroalanine; DMF: dimethylformamide; DMSO: dimethyl sulfoxide; DTT: dithiothreitol; EDC: 1-ethyl-3-(3-dimethylaminopropyl)carbodiimide; GyrA: gyrase A; ICP-OES: inductively coupled plasma-atomic emission spectroscopy; Mes: mesityl; MTG: microbial transglutaminase; PTM: post-translational modifications; Sulfo-NHS: *N*-hydroxysulfosuccinimide; SBL: subtilisin from *Bacillus lentus*; SFG: sum frequency generation; SHG: second harmonic generation; TCEP: tris(2-carboxyethyl)phosphine; TMV: tobacco mosaic virus; TPPTS: triphenylphosphine-3,3′,3″-trisulfonic acid; TYMV: turnip yellow mosaic virus

REFERENCES

(1) Lam, K. S., Salmon, S. E., Hersh, E. M., Hruby, V. J., Kazmierski, W. M., and Knapp, R. J. (1991) A new type of synthetic peptide library for identifying ligand-binding activity. *Nature 354*, 82−84.

(2) Meldal, M. (2004) "One bead two compound libraries" for detecting chemical and biochemical conversions. *Curr. Opin. Chem. Biol. 8*, 238−244.

(3) Wilson, R., Cossins, A. R., and Spiller, D. G. (2006) Encoded microcarriers for high-throughput multiplexed detection. *Angew. Chem., Int. Ed. 45*, 6104−6117.

(4) Kehoe, J. W., and Kay, B. K. (2005) Filamentous phage display in the new millennium. *Chem. Rev. 105*, 4056−4072.

(5) Mao, C. B., Liu, A. H., and Cao, B. R. (2009) Virus-based chemical and biological sensing. *Angew. Chem., Int. Ed. 48*, 6790−6810.

(6) Pokorski, J. K., and Steinmetz, N. F. (2011) The art of engineering viral nanoparticles. *Mol. Pharm. 8*, 29−43.

(7) Lee, L. A., Niu, Z. W., and Wang, Q. (2009) Viruses and virus-like protein assemblies−chemically programmable nanoscale building blocks. *Nano Res. 2*, 349−364.

(8) de la Escosura, A., Nolte, R. J. M., and Cornelissen, J. (2009) Viruses and protein cages as nanocontainers and nanoreactors. *J. Mater. Chem. 19*, 2274−2278.

(9) Breitbart, M., and Rohwer, F. (2005) Here a virus, there a virus, everywhere the same virus? *Trends Microbiol. 13*, 278−284.

(10) Sanger, F., Air, G. M., Barrell, B. G., Brown, N. L., Coulson, A. R., Fiddes, J. C., Hutchison, C. A., Slocombe, P. M., and Smith, M. (1977) Nucleotide-sequence of bacteriophage Phichi174 DNA. *Nature 265*, 687−695.

(11) Beck, E., Sommer, R., Auerswald, E. A., Kurz, C., Zink, B., Osterburg, G., Schaller, H., Sugimoto, K., Sugisaki, H., Okamoto, T., and Takanami, M. (1978) Nucleotide-sequence of bacteriophage-Fd DNA. *Nucleic Acids Res. 5*, 4495−4503.

(12) Fiers, W., Contreras, R., Haegeman, G., Rogiers, R., Vandevoorde, A., Vanheuverswyn, H., Vanherreweghe, J., Volckaert, G., and Ysebaert, M. (1978) Complete nucleotide-sequence of Sv40 DNA. *Nature 273*, 113−120.
(13) Smith, H. O., Hutchison, C. A., Pfannkoch, C., and Venter, J. C. (2003) Generating a synthetic genome by whole genome assembly: phi X174 bacteriophage from synthetic oligonucleotides. *Proc. Natl. Acad. Sci. U.S.A. 100*, 15440−15445.
(14) Mandecki, W., Hayden, M. A., Shallcross, M. A., and Stotland, E. (1990) A totally synthetic plasmid for general cloning, gene-expression and mutagenesis in *Escherichia coli*. *Gene. 94*, 103−107.
(15) Nicolaou, K. C., Xiao, X. Y., Parandoosh, Z., Senyei, A., and Nova, M. P. (1995) Radiofrequency encoded combinatorial chemistry. *Angew. Chem., Int. Ed. Engl. 34*, 2289−2291.
(16) Halpin, D. R., and Harbury, P. B. (2004) DNA display I. Sequence-encoded routing of DNA populations. *PLoS Biol. 2*, 1015−1021.
(17) Silverman, S. K. (2010) DNA as a versatile chemical component for catalysis, encoding, and stereocontrol. *Angew. Chem., Int. Ed. 49*, 7180−7201.
(18) Li, X. Y., and Liu, D. R. (2004) DNA-Templated organic synthesis: Nature's strategy for controlling chemical reactivity applied to synthetic molecules. *Angew. Chem., Int. Ed. 43*, 4848−4870.
(19) Woiwode, T. F., Haggerty, J. E., Katz, R., Gallop, M. A., Barrett, R. W., Dower, W. J., and Cwirla, S. E. (2003) Synthetic compound libraries displayed on the surface of encoded bacteriophage. *Chem. Biol. 10*, 847−858.
(20) Smith, G. P., and Scott, J. K. (1993) Libraries of peptides and proteins displayed on filamentous phage. *Methods Enzymol. 217*, 228−257.
(21) Krumpe, L. R. H., Atkinson, A. J., Smythers, G. W., Kandel, A., Schumacher, K. M., McMahon, J. B., Makowski, L., and Mori, T. (2006) T7 lytic phage-displayed peptide libraries exhibit less sequence bias than M13 filamentous phage-displayed peptide libraries. *Proteomics 6*, 4210−4222.
(22) Beghetto, E., and Gargano, N. (2011) Lambda-display: A powerful tool for antigen discovery. *Molecules 16*, 3089−3105.
(23) Makela, A. R., and Oker-Blom, C. (2008) The baculovirus display technology−An evolving instrument for molecular screening and drug delivery. *Combi. Chem. High Throughput Screening 11*, 86−98.
(24) Muller, O. J., Kaul, F., Weitzman, M. D., Pasqualini, R., Arap, W., Kleinschmidt, J. A., and Trepel, M. (2003) Random peptide libraries displayed on adeno-associated virus to select for targeted gene therapy vectors. *Nat. Biotechnol. 21*, 1040−1046.
(25) Caldeira, J. D., Medford, A., Kines, R. C., Lino, C. A., Schiller, J. T., Chackerian, B., and Peabody, D. S. (2010) Immunogenic display of diverse peptides, including a broadly cross-type neutralizing human papillomavirus L2 epitope, on virus-like particles of the RNA bacteriophage PP7. *Vaccine 28*, 4384−4393.
(26) Peabody, D. S., Manifold-Wheeler, B., Medford, A., Jordan, S. K., Caldeira, J. D., and Chackerian, B. (2008) Immunogenic display of diverse peptides on virus-like particles of RNA phage MS2. *J. Mol. Biol. 380*, 252−263.
(27) Kumar, S., Ochoa, W., Singh, P., Hsu, C., Schneemann, A., Manchester, M., Olson, M., and Reddy, V. (2009) Tomato bushy stunt virus (TBSV), a versatile platform for polyvalent display of antigenic epitopes and vaccine design. *Virology 388*, 185−190.
(28) Aune, T. E. V., and Aachmann, F. L. (2010) Methodologies to increase the transformation efficiencies and the range of bacteria that can be transformed. *Appl. Microbiol. Biotechnol. 85*, 1301−1313.
(29) 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*, 1039−1052.
(30) Derda, R., Tang, S. K. Y., Li, S. C., Ng, S., Matochko, W., and Jafari, M. R. (2011) Diversity of phage-displayed libraries of peptides during panning and amplification. *Molecules 16*, 1776−1803.
(31) Rodi, D. J., Makowski, L. (1997) Transfer RNA isoacceptor availability contributes to sequence censorship in a library of phage displayed peptides, in *Structure and Function of Macromolecular Assembly* (Namba, K., Ed.), pp 155−164, Matsushita Electric Industrial Co. Ltd, Matsushita, Japan.
(32) Bothmann, H., and Pluckthun, A. (1998) Selection for a periplasmic factor improving phage display and functional periplasmic expression. *Nat. Biotechnol. 16*, 376−380.
(33) Jestin, J. L., Volioti, G., and Winter, G. (2001) Improving the display of proteins on filamentous phage. *Res. Microbiol. 152*, 187−191.
(34) Wang, Q., Kaltgrad, E., Lin, T., Johnson, J. E., and Finn, M. G. (2002) Natural supramolecular building blocks: Wild-type Cowpea mosaic virus. *Chem. Biol. 9*, 805−811.
(35) Johnson, H. R., Hooker, J. M., Francis, M. B., and Clark, D. S. (2007) Solubilization and stabilization of bacteriophage MS2 in organic solvents. *Biotechnol. Bioeng. 97*, 224−234.
(36) Holder, P. G., Finley, D. T., Stephanopoulos, N., Walton, R., Clark, D. S., and Francis, M. B. (2010) Dramatic thermal stability of virus-polymer conjugates in hydrophobic solvents. *Langmuir 26*, 17383−17388.
(37) Schlick, T. L., Ding, Z., Kovacs, E. W., and Francis, M. B. (2005) Dual-surface modification of the tobacco mosaic virus. *J. Am. Chem. Soc. 127*, 3718−3723.
(38) Olofsson, L., Ankarloo, J., Andersson, P. O., and Nicholls, I. A. (2001) Filamentous bacteriophage stability in non-aqueous media. *Chem. Biol. 8*, 661−671.
(39) Olofsson, L., Ankarloo, J., and Nicholls, I. A. (1998) Phage viability in organic media: insights into phage stability. *J. Mol. Recognit. 11*, 91−93.
(40) Oh, J. S., Davies, D. R., Lawson, J. D., Arnold, G. E., and Dunker, A. K. (1999) Isolation of chloroform-resistant mutants of filamentous phage: localization in models of phage structure. *J. Mol. Biol. 287*, 449−457.
(41) Olofsson, L., Söderberg, P., Ankarloo, J., and Nicholls, I. A. (2008) Phage display screening in low dielectric media. *J. Mol. Recognit. 21*, 330−337.
(42) Steinmetz, N. F., Bize, A., Findlay, K. C., Lomonossoff, G. P., Manchester, M., Evans, D. J., and Prangishvili, D. (2008) Site-specific and spatially controlled addressability of a new viral nanobuilding block: Sulfolobus islandicus rod-shaped virus 2. *Adv. Funct. Mater. 18*, 3478−3486.
(43) Schleper, C., Kubo, K., and Zillig, W. (1992) The particle Ssv1 from the extremely thermophilic Archaeon Sulfolobus is a virus−Demonstration of infectivity and of transfection with viral-DNA. *Proc. Natl. Acad. Sci. U.S.A. 89*, 7645−7649.
(44) Berkner, S., and Lipps, G. (2008) Genetic tools for Sulfolobus spp.: vectors and first applications. *Arch. Microbiol. 190*, 217−230.
(45) Lipps, G. (2006) Plasmids and viruses of the thermoacidophilic crenarchaeote Sulfolobus. *Extremophiles 10*, 17−28.
(46) Leigh, J. A., Albers, S. V., Atomi, H., and Allers, T. (2011) Model organisms for genetics in the domain Archaea: methanogens, halophiles, Thermococcales and Sulfolobales. *FEMS Microbiol. Rev. 35*, 577−608.
(47) Wirth, J. F., Snyder, J. C., Hochstein, R. A., Ortmann, A. C., Willits, D. A., Douglas, T., and Young, M. J. (2011) Development of a genetic system for the archaeal virus Sulfolobus turreted icosahedral virus (STIV). *Virology. 415*, 6−11.
(48) Hoesl, M. G. H. M. G., and Budisa, N. (2011) In vivo incorporation of multiple noncanonical amino acids into proteins. *Angew. Chem., Int. Ed. 50*, 2896−2902.
(49) Wang, L., Xie, J., and Schultz, P. G. (2006) Expanding the genetic code. *Annu. Rev. Biophys. Biomol. Struct. 35*, 225−249.
(50) Sletten, E. M., and Bertozzi, C. R. (2009) Bioorthogonal chemistry: Fishing for selectivity in a sea of functionality. *Angew. Chem., Int. Ed. 48*, 6974−6998.
(51) Algar, W. R., Prasuhn, D. E., Stewart, M. H., Jennings, T. L., Blanco-Canosa, J. B., Dawson, P. E., and Medintz, I. L. (2011) The controlled display of biomolecules on nanoparticles: A challenge suited to bioorthogonal chemistry. *Bioconjugate Chem. 22*, 825−858.
(52) Sandman, K. E., Benner, J. S., and Noren, C. J. (2000) Phage display of selenopeptides. *J. Am. Chem. Soc. 122*, 960−961.

(53) Love, K. R., Swoboda, J. G., Noren, C. J., and Walker, S. (2006) Enabling glycosyltransferase evolution: A facile substrate-attachment strategy for phage-display enzyme evolution. *ChemBioChem 7*, 753–756.
(54) Kiick, K. L., Saxon, E., Tirrell, D. A., and Bertozzi, C. R. (2002) Incorporation of azides into recombinant proteins for chemoselective modification by the Staudinger ligation. *Proc. Natl. Acad. Sci. U.S.A. 99*, 19–24.
(55) Wang, A., Nairn, N. W., Johnson, R. S., Tirrell, D. A., and Grabstein, K. (2008) Processing of N-terminal unnatural amino acids in recombinant human interferon-beta in *Escherichia coli*. *ChemBioChem. 9*, 324–330.
(56) Tian, F., Tsao, M.-L., and Schultz, P. G. (2004) A phage display system with unnatural amino acids. *J. Am. Chem. Soc. 126*, 15962–15963.
(57) Tsao, M. L., Tian, F., and Schultz, P. G. (2005) Selective Staudinger modification of proteins containing *p*-azidophenylalanine. *ChemBioChem. 6*, 2147–2149.
(58) Liu, C. C., Schultz, P. G. (2010) Adding New Chemistries to the Genetic Code, in *Annual Review of Biochemistry*, Vol. *79*, pp 413–444, Annual Reviews, Palo Alto.
(59) Carrico, Z. M., Romanini, D. W., Mehl, R. A., and Francis, M. B. (2008) Oxidative coupling of peptides to a virus capsid containing unnatural amino acids. *Chem. Commun.*, 1205–1207.
(60) Sidhu, S. S., and Kossiakoff, A. A. (2007) Exploring and designing protein function with restricted diversity. *Curr. Opin. Chem. Biol. 11*, 347–354.
(61) Dwyer, M. A., Lu, W. Y., Dwyer, J. J., and Kossiakoff, A. (2000) Biosynthetic phage display: a novel protein engineering tool combining chemical and genetic diversity. *Chem. Biol. 7*, 263–274.
(62) Evans, T. C., Benner, J., and Xu, M.-Q. (1999) The in vitro ligation of bacterially expressed proteins using an intein from *Methanobacterium thermoautotrophicum*. *J. Biol. Chem. 274*, 3923–3926.
(63) Tolbert, T. J., and Wong, C.-H. (2002) New methods for proteomic research: Preparation of proteins with N-terminal cysteines for labeling and conjugation. *Angew. Chem., Int. Ed. 41*, 2171–2174.
(64) Kather, I., Bippes, C. A., and Schmid, F. X. (2005) A stable disulfide-free gene-3-protein of phage fd generated by in vitro evolution. *J. Mol. Biol. 354*, 666–678.
(65) Heinis, C., Rutherford, T., Freund, S., and Winter, G. (2009) Phage-encoded combinatorial chemical libraries based on bicyclic peptides. *Nat. Chem. Biol. 5*, 502–507.
(66) Roth, T. A., Weiss, G. A., Eigenbrot, C., and Sidhu, S. S. (2002) A minimized M13 coat protein defines the requirements for assembly into the bacteriophage particle. *J. Mol. Biol. 322*, 357–367.
(67) Williams, K. A., Glibowicka, M., Li, Z., Li, H., Khan, A. R., Chen, Y. M. Y., Wang, J., Marvin, D. A., and Deber, C. M. (1995) Packing of coat protein amphipathic and transmembrane helices in filamentous bacteriophage M13: Role of small residues in protein oligomerization. *J. Mol. Biol. 252*, 6–14.
(68) Held, H. A., and Sidhu, S. S. (2004) Comprehensive mutational analysis of the M13 major coat protein: Improved scaffolds for C-terminal phage display. *J. Mol. Biol. 340*, 587–597.
(69) Weiss, G. A., Roth, T. A., Baldi, P. F., and Sidhu, S. S. (2003) Comprehensive mutagenesis of the C-terminal domain of the M13 gene-3 minor coat protein: The requirements for assembly into the bacteriophage particle. *J. Mol. Biol. 332*, 777–782.
(70) Spada, S., and Plückthun, A. (1997) Selectively infective phage (SIP) technology: A novel method for in vivo selection of interacting protein-ligand pairs. *Nat. Med. 3*, 694–696.
(71) Chen, I., Howarth, M., Lin, W. Y., and Ting, A. Y. (2005) Site-specific labeling of cell surface proteins with biophysical probes using biotin ligase. *Nat. Methods 2*, 99–104.
(72) Wang, Q., Lin, T., Johnson, J. E., and Finn, M. G. (2002) Natural supramolecular building blocks: Cysteine-added mutants of Cowpea mosaic virus. *Chem. Biol. 9*, 813–819.
(73) Smyth, D. G., Blumenfeld, O. O., and Konigsberg, W. (1964) Reactions of *N*-ethylmaleimide with peptides and amino acids. *Biochem. J. 91*, 589–595.
(74) Hermanson, G. T. (2008) *Bioconjugate Techniques*, 2nd ed., Academic Press, New York.
(75) Chalker, J. M., Wood, C. S. C., and Davis, B. G. (2009) A convenient catalyst for aqueous and protein Suzuki–Miyaura cross-coupling. *J. Am. Chem. Soc. 131*, 16346–16347.
(76) Bernardes, G. J. L., Chalker, J. M., Errey, J. C., and Davis, B. G. (2008) Facile conversion of cysteine and alkyl cysteines to dehydroalanine on protein surfaces: Versatile and switchable access to functionalized proteins. *J. Am. Chem. Soc. 130*, 5052–5053.
(77) Chalker, J. M., Gunnoo, S. B., Boutureira, O., Gerstberger, S. C., Fernandez-Gonzalez, M., Bernardes, G. J. L., Griffin, L., Hailu, H., Schofield, C. J., and Davis, B. G. (2011) Methods for converting cysteine to dehydroalanine on peptides and proteins. *Chem. Sci. 2*, 1666–1676.
(78) Garbe, D., Thiel, I. V., and Mootz, H. D. (2010) Protein trans-splicing on an M13 bacteriophage: towards directed evolution of a semisynthetic split intein by phage display. *J. Pept. Sci. 16*, 575–581.
(79) Ren, H., Xiao, F., Zhan, K., Kim, Y.-P., Xie, H., Xia, Z., and Rao, J. (2009) A biocompatible condensation reaction for the labeling of terminal cysteine residues on proteins. *Angew. Chem., Int. Ed. 48*, 9658–9662.
(80) Chalker, J. M., Lin, Y. A., Boutureira, O., and Davis, B. G. (2009) Enabling olefin metathesis on proteins: chemical methods for installation of S-allyl cysteine. *Chem. Commun.*, 3714–3716.
(81) Lin, Y. A., Chalker, J. M., Floyd, N., Bernardes, G. a. J. L., and Davis, B. G. (2008) Allyl sulfides are privileged substrates in aqueous cross-metathesis: Application to site-selective protein modification. *J. Am. Chem. Soc. 130*, 9642–9643.
(82) Lin, Y. A., Chalker, J. M., and Davis, B. G. (2010) Olefin cross-metathesis on proteins: Investigation of allylic chalcogen effects and guiding principles in metathesis partner selection. *J. Am. Chem. Soc. 132*, 16805–16811.
(83) Burtscher, D., and Grela, K. (2009) Aqueous olefin metathesis. *Angew. Chem., Int. Ed. 48*, 442–454.
(84) Barnhill, H. N., Reuther, R., Ferguson, P. L., Dreher, T., and Wang, Q. (2007) Turnip yellow mosaic virus as a chemoaddressable bionanoparticle. *Bioconjugate Chem. 18*, 852–859.
(85) Li, K., Chen, Y., Li, S., Nguyen, H. G., Niu, Z., You, S., Mello, C. M., Lu, X., and Wang, Q. (2010) Chemical modification of M13 bacteriophage and its application in cancer cell imaging. *Bioconjugate Chem. 21*, 1369–1377.
(86) Lim, R. K. V., and Lin, Q. (2010) Bioorthogonal chemistry: Recent progress and future directions. *Chem. Commun. 46*, 1589–1600.
(87) Tanaka, F., Fuller, R., Asawapornmongkol, L., Warsinke, A., Gobuty, S., and Barbas Iii, C. F. (2007) Development of a small peptide tag for covalent labeling of proteins. *Bioconjugate Chem. 18*, 1318–1324.
(88) Horowitz, E. D., Weinberg, M. S., and Asokan, A. (2011) Glycated AAV vectors: Chemical redirection of viral tissue tropism. *Bioconjugate Chem. 22*, 529–532.
(89) Takahashi, K. (1977) The reactions of phenylglyoxal and related reagents with amino acids. *J. Biochem. 81*, 395–402.
(90) Takahashi, K. (1977) Further studies on the reactions of phenylglyoxal and related reagents with proteins. *J. Biochem. 81*, 403–414.
(91) Ban, H., Gavrilyuk, J., and Barbas, C. F. (2010) Tyrosine bioconjugation through aqueous ene-type reactions: A click-like reaction for tyrosine. *J. Am. Chem. Soc. 132*, 1523–1525.
(92) Joshi, N. S., Whitaker, L. R., and Francis, M. B. (2004) A three-component Mannich-type reaction for selective tyrosine bioconjugation. *J. Am. Chem. Soc. 126*, 15942–15943.
(93) Dirksen, A., Hackeng, T. M., and Dawson, P. E. (2006) Nucleophilic catalysis of oxime ligation. *Angew. Chem., Int. Ed. 118*, 7743–7746.

(94) McFarland, J. M., Joshi, N. S., and Francis, M. B. (2008) Characterization of a three-component coupling reaction on proteins by isotopic labeling and nuclear magnetic resonance spectroscopy. *J. Am. Chem. Soc. 130*, 7639−7644.
(95) Hooker, J. M., Kovacs, E. W., and Francis, M. B. (2004) Interior surface modification of bacteriophage MS2. *J. Am. Chem. Soc. 126*, 3718−3719.
(96) Sokolovsky, M., and Vallee, B. L. (1966) The reaction of diazonium-1H-tetrazole with proteins. Determination of tyrosine and histidine content. *Biochemistry (Moscow) 5*, 3574−3581.
(97) Tilley, S. D., and Francis, M. B. (2006) Tyrosine-selective protein alkylation using π-allylpalladium complexes. *J. Am. Chem. Soc. 128*, 1080−1081.
(98) Antos, J. M., and Francis, M. B. (2004) Selective tryptophan modification with rhodium carbenoids in aqueous solution. *J. Am. Chem. Soc. 126*, 10256−10257.
(99) Antos, J. M., McFarland, J. M., Iavarone, A. T., and Francis, M. B. (2009) Chemoselective tryptophan labeling with rhodium carbenoids at mild pH. *J. Am. Chem. Soc. 131*, 6301−6308.
(100) Popp, B. V., and Ball, Z. T. (2010) Structure-selective modification of aromatic side chains with dirhodium metallopeptide catalysts. *J. Am. Chem. Soc. 132*, 6660−6662.
(101) Popp, B. V., and Ball, Z. T. (2011) Proximity-driven metallopeptide catalysis: Remarkable side-chain scope enables modification of the Fos bZip domain. *Chem. Sci. 2*, 690−695.
(102) Antos, J. M., and Francis, M. B. (2006) Transition metal catalyzed methods for site-selective protein modification. *Curr. Opin. Chem. Biol. 10*, 253−262.
(103) Rohwer, F., and Thurber, R. V. (2009) Viruses manipulate the marine environment. *Nature 459*, 207−212.
(104) Dong, S. L., McPherson, S. A., Wang, Y., Li, M., Wang, P. F., Turnbough, C. L., and Pritchard, D. G. (2010) Characterization of the enzymes encoded by the anthrose biosynthetic operon of *Bacillus anthracis*. *J. Bacteriol. 192*, 5053−5062.
(105) Zhang, Y., Xiang, Y., Van Etten, J. L., and Rossmann, M. G. (2007) Structure and function of a chlorella virus-encoded glycosyltransferase. *Structure 15*, 1031−1039.
(106) Xiang, Y., Baxa, U., Zhang, Y., Steven, A. C., Lewis, G. L., Van Etten, J. L., and Rossmann, M. G. (2010) Crystal structure of a virus-encoded putative glycosyltransferase. *J. Virol. 84*, 12265−12273.
(107) Fernandez-Gacio, A., Uguen, M., and Fastrez, J. (2003) Phage display as a tool for the directed evolution of enzymes. *Trends Biotechnol. 21*, 408−414.
(108) Sunbul, M., Emerson, N., and Yin, J. (2011) Enzyme-catalyzed substrate attachment to phage surfaces for the selection of catalytic activities. *ChemBioChem 12*, 380−386.
(109) Derda, R., Tang, S. K. Y., and Whitesides, G. M. (2010) Uniform amplification of phage with different growth characteristics in individual compartments consisting of monodisperse droplets. *Angew. Chem., Int. Ed. 49*, 5301−5304.
(110) Agresti, J. J., Antipov, E., Abate, A. R., Ahn, K., Rowat, A. C., Baret, J. C., Marquez, M., Klibanov, A. M., Griffiths, A. D., and Weitz, D. A. (2010) Ultrahigh-throughput screening in drop-based microfluidics for directed evolution. *Proc. Natl. Acad. Sci. U.S.A. 107*, 4004−4009.
(111) Szymanski, C. M., Yao, R. J., Ewing, C. P., Trust, T. J., and Guerry, P. (1999) Evidence for a system of general protein glycosylation in *Campylobacter jejuni*. *Mol. Microbiol. 32*, 1022−1030.
(112) Wacker, M., Linton, D., Hitchen, P. G., Nita-Lazar, M., Haslam, S. M., North, S. J., Panico, M., Morris, H. R., Dell, A., Wren, B. W., and Aebi, M. (2002) N-linked glycosylation in Campylobacter jejuni and its functional transfer into *E. coli*. *Science 298*, 1790−1793.
(113) Celik, E., Fisher, A. C., Guarino, C., Mansell, T. J., and DeLisa, M. P. (2010) A filamentous phage display system for N-linked glycoproteins. *Protein Sci. 19*, 2006−2013.
(114) Durr, C., Nothaft, H., Lizak, C., Glockshuber, R., and Aebi, M. (2010) The *Escherichia coli* glycophage display system. *Glycobiology 20*, 1366−1372.
(115) Magidovich, H., and Eichler, J. (2009) Glycosyltransferases and oligosaccharyltransferases in Archaea: Putative components of the N-glycosylation pathway in the third domain of life. *FEMS Microbiol. Lett. 300*, 122−130.
(116) Sugimura, Y., Yokoyama, K., Nio, N., Maki, M., and Hitomi, K. (2008) Identification of preferred substrate sequences of microbial transglutaminase from Streptomyces mobaraensis using a phage-displayed peptide library. *Arch. Biochem. Biophys. 477*, 379−383.
(117) Touati, J., Angelini, A., Hinner, M. J., and Heinis, C. (2011) Enzymatic cyclisation of peptides with a transglutaminase. *ChemBioChem 12*, 38−42.
(118) Scholle, M. D., Kriplani, U., Pabon, A., Sishtla, K., Glucksman, M. J., and Kay, B. K. (2006) Mapping protease substrates by using a biotinylated phage substrate library. *ChemBioChem 7*, 834−838.
(119) Stolz, J., Ludwig, A., and Sauer, N. (1998) Bacteriophage lambda surface display of a bacterial biotin acceptor domain reveals the minimal peptide size required for biotinylation. *FEBS Lett. 440*, 213−217.
(120) Belshan, M., Matthews, J. M., and Madson, C. J. (2011) Replication of biotinylated human immunodeficiency viruses. *J. Virol. Methods 171*, 299−302.
(121) Kaikkonen, M. U., Viholainen, J. I., Narvanen, A., Yla-Herttuala, S., and Airenne, K. J. (2008) Targeting and purification of metabolically biotinylated baculovirus. *Hum. Gene Ther. 19*, 589−600.
(122) Nesbeth, D., Williams, S. L., Chan, L., Brain, T., Slater, N. K. H., Farzaneh, F., and Darling, D. (2006) Metabolic biotinylation of lentiviral pseudotypes for scalable paramagnetic microparticle-dependent manipulation. *Mol. Ther. 13*, 814−822.
(123) Yin, J., Liu, F., Schinke, M., Daly, C., and Walsh, C. T. (2004) Phagemid encoded small molecules for high throughput screening of chemical libraries. *J. Am. Chem. Soc. 126*, 13570−13571.
(124) Yin, J., Straight, P. D., McLoughlin, S. M., Zhou, Z., Lin, A. J., Golan, D. E., Kelleher, N. L., Kolter, R., and Walsh, C. T. (2005) Genetically encoded short peptide tag for versatile protein labeling by Sfp phosphopantetheinyl transferase. *Proc. Natl. Acad. Sci. U.S.A. 102*, 15815−15820.
(125) Zhou, Z., Cironi, P., Lin, A. J., Xu, Y. Q., Hrvatin, S., Golan, D. E., Silver, P. A., Walsh, C. T., and Yin, J. (2007) Genetically encoded short peptide tags for orthogonal protein labeling by sfp and AcpS phosphopantetheinyl transferases. *ACS Chem. Biol. 2*, 337−346.
(126) Mootz, H. D. (2009) Split Inteins as Versatile Tools for Protein Semisynthesis. *ChemBioChem 10*, 2579−2589.
(127) Keresztessy, Z., Csosz, E., Harsfalvi, J., Csomos, K., Gray, J., Lightowlers, R. N., Lakey, J. H., Balajthy, Z., and Fesus, L. (2006) Phage display selection of efficient glutamine-donor substrate peptides for transglutaminase 2. *Protein Sci. 15*, 2466−2480.
(128) Rabuka, D. (2010) Chemoenzymatic methods for site-specific protein modification. *Curr. Opin. Chem. Biol. 14*, 790−796.
(129) Proft, T. (2009) Sortase-mediated protein ligation: An emerging biotechnology tool for protein modification and immobilisation. *Biotechnol. Lett. 32*, 1−10.
(130) You, Y. O., Levengood, M. R., Ihnken, L. A. F., Knowlton, A. K., and van der Donk, W. A. (2009) Lacticin 481 synthetase as a general serine/threonine kinase. *ACS Chem. Biol. 4*, 379−385.
(131) Zeri, A. C., Mesleh, M. F., Nevzorov, A. A., and Opella, S. J. (2003) Structure of the coat protein in fd filamentous bacteriophage particles determined by solid-state NMR spectroscopy. *Proc. Natl. Acad. Sci. U.S.A. 100*, 6458−6463.
(132) Thiriot, D. S., Nevzorov, A. A., Zagyanskiy, L., Wu, C. H., and Opella, S. J. (2004) Structure of the coat protein in Pf1 bacteriophage determined by solid-state NMR spectroscopy. *J. Mol. Biol. 341*, 869−879.
(133) Manzenrieder, F., Luxenhofer, R., Retzlaff, M., Jordan, R., and Finn, M. G. (2011) Stabilization of virus-like particles with poly(2-oxazoline)s. *Angew. Chem., Int. Ed. 50*, 2601−2605.
(134) Prasuhn, D. E. Jr, Singh, P., Strable, E., Brown, S., Manchester, M., and Finn, M. G. (2008) Plasma clearance of bacteriophage Qβ particles as a function of surface charge. *J. Am. Chem. Soc. 130*, 1328−1334.

(135) Hong, V., Udit, A. K., Evans, R. A., and Finn, M. G. (2008) Electrochemically protected copper(I)-catalyzed azide-alkyne cycloaddition. *ChemBioChem 9*, 1481−1486.
(136) Campagnola, P. J., and Loew, L. M. (2003) Second-harmonic imaging microscopy for visualizing biomolecular arrays in cells, tissues and organisms. *Nat. Biotechnol. 21*, 1356−1360.
(137) Liu, T. M., Lee, Y. W., Chang, C. F., Yeh, S. C., Wang, C. H., Chu, S. W., and Sun, C. K. (2008) Imaging polyhedral inclusion bodies of nuclear polyhedrosis viruses with second harmonic generation microscopy. *Opt. Express 16*, 5602−5608.
(138) Iwamoto, M., Liu, F., and Zhong-Can, O. Y. (2011) Polarization-dependence of optical second harmonic generation for chiral cylindrical structure and explanation for nonlinear optical imaging of cholesteric liquid crystals. *Chem. Phys. Lett. 511*, 455−460.
(139) Steinmetz, N. F., Hong, V., Spoerke, E. D., Lu, P., Breitenkamp, K., Finn, M. G., and Manchester, M. (2009) Buckyballs meet viral nanoparticles: Candidates for biomedicine. *J. Am. Chem. Soc. 131*, 17093−17095.
(140) Zeng, Q., Saha, S., Lee, L. A., Barnhill, H., Oxsher, J., Dreher, T., and Wang, Q. (2011) Chemoselective modification of turnip yellow mosaic virus by Cu(I) catalyzed azide-alkyne 1,3-dipolar cycloaddition reaction and its application in cell binding. *Bioconjugate Chem. 22*, 58−66.
(141) Udit, A. K., Brown, S., Baksh, M. M., and Finn, M. G. (2008) Immobilization of bacteriophage Qβ on metal-derivatized surfaces via polyvalent display of hexahistidine tags. *J. Inorg. Biochem. 102*, 2142−2146.
(142) Peabody, D. S. (2003) A viral platform for chemical modification and multivalent display. *J. Nanobiotechnol. 1*, 5.
(143) Pokorski, J. K., Breitenkamp, K., Liepold, L. O., Qazi, S., and Finn, M. G. (2011) Functional virus-based polymer-protein nanoparticles by atom transfer radical polymerization. *J. Am. Chem. Soc. 133*, 9242−9245.
(144) Steinmetz, N. F., Miertens, M. E., Taurog, R. E., Johnson, J. E., Commandeur, U., Fischer, R., and Manchester, M. (2010) Potato virus X as a novel platform for potential biomedical applications. *Nano Lett. 10*, 305−312.
(145) Chatterji, A., Ochoa, W. F., Paine, M., Ratna, B. R., Johnson, J. E., and Lin, T. (2004) New addresses on an addressable virus nanoblock: Uniquely reactive Lys residues on cowpea mosaic virus. *Chem. Biol. 11*, 855−863.
(146) Bruckman, M. A., Kaur, G., Lee, A. L., Xie, F., Sepulveda, J., Breitenkamp, R., Zhang, X., Joralemon, M., Russell, T. P., Emrick, T., and Wang, Q. (2008) Surface modification of tobacco mosaic virus with "click" chemistry. *ChemBioChem 9*, 519−523.
(147) Udit, A. K., Everett, C., Gale, A. J., Kyle, J. R., Ozkan, M., and Finn, M. G. (2009) Heparin antagonism by polyvalent display of cationic motifs on virus-like particles. *ChemBioChem 10*, 503−510.
(148) Jespers, L., Bonnert, T. P., and Winter, G. (2004) Selection of optical biosensors from chemisynthetic antibody libraries. *Protein Eng. Des. Sel. 17*, 709−713.
(149) Angelini, A., and Heinis, C. (2011) Post-translational modification of genetically encoded polypeptide libraries. *Curr. Opin. Chem. Biol. 15*, 355−361.