Phage Display: Simple Evolution in a Petri Dish

Nobel Lecture, December 8, 2018 by
George P. Smith
University of Missouri, Columbia, USA.

INTRODUCTION

In 1997, Valery Petrenko and I published a perspective on phage display in *Chemical Reviews* (Smith and Petrenko, 1997) that adumbrated the over-arching "directed evolution" theme of the 2018 Nobel Prize in Chemistry. "Imagine, then, the applied chemist, not as designer of molecules with a particular purpose, but rather as custodian of a highly diverse population of chemicals evolving *in vitro* as if they were organisms subject to natural selection," we wrote. "A chemical's 'fitness' in this artificial biosphere would be imposed by the custodian for his or her own ends. For instance, the population might be culled periodically of individuals who fail to bind tightly to some biological receptor; the population would then evolve toward specific ligands for that receptor ... Progress toward the custodian's chosen goal would in a sense be 'automatic': once appropriate selection conditions are devised, no plan for how the system is to meet the demands of selection need be specified. And if the chemical population is sufficiently diverse, perhaps this 'blind' process will outperform rational design. The custodian may not comprehend, even in retrospect, how the products of selection work, just as biologists have only the sketchiest understanding of how a fruitfly functions."

Unlike the 1997 perspective, the present essay, like my oral presentation (Smith, 2018), will not attempt to position phage display within

grand themes in the history of scientific ideas. I will try instead to recon-struct the story of the phage-display idea as I personally experienced it. As you will see, it is not a tale of heroic flashes of insight. Rather, it is a case study in how a scientific advance emerges gradually in incremental steps within overlapping global scientific communities. To foster science, to enjoy its many material and cultural benefits, I believe our society must sustain whole scientific communities like the ones from which phage dis-play emerged, not try to identify a small cadre of individuals who are somehow specially destined to make breakthrough discoveries or innova-tions – or to be awarded Nobel prizes.

A PHAGE PRIMER

Phages are viruses that infect bacteria. They are particularly convenient research subjects, both because they can be propagated to enormous numbers inexpensively in microbiological cultures, and because the virus particles, called *virions*, are sturdy structures that can be readily freed of most impurities and manipulated in the laboratory.

The virion has an outer shell, called the *capsid*, composed of a geomet-ric array of phage coat proteins. Inside the capsid is the phage chromo-some, whose genes encode the phage proteins (including the coat pro-teins). The capsid serves not only to protect the chromosome, but also to mediate the infection process, whereby the virion attaches to an unin-fected bacterium and the phage chromosome is transferred into the cell's cytoplasm. Once inside the cell, the phage genes reprogram the cell's machinery to make progeny virions, which are ultimately released from the cell and can go on to infect hundreds of uninfected cells to initiate another round of infection.

Different classes of phages infect different bacterial species and have diverse virion sizes and shapes. Filamentous phages, the particular class chosen for phage display, infect the enteric (gut) bacterium *Escherichia* coli. The virions are long and thin, as seen in the electron micrograph in Figure 1; wildtype virions are about 1 µm long and 6 nm across. The capsid consists of a long tubular array of thousands of major coat pro-tein subunits. One end of the tube is capped by 5 copies each of two minor coat proteins; the opposite end is capped by 5 copies each of two other minor coat proteins.

One of the minor coat proteins, pIII, is the subject of the work described in this essay. Its surface-exposed portion is sometimes visible as irregular knobs at one tip of the virion, as explained in the next sec-tion. Inside the capsid is the phage DNA chromosome. At the right in Figure 1 is a schematic representation of the virion. The phage chromo-some is represented by the black links inside. Coat protein gene III is
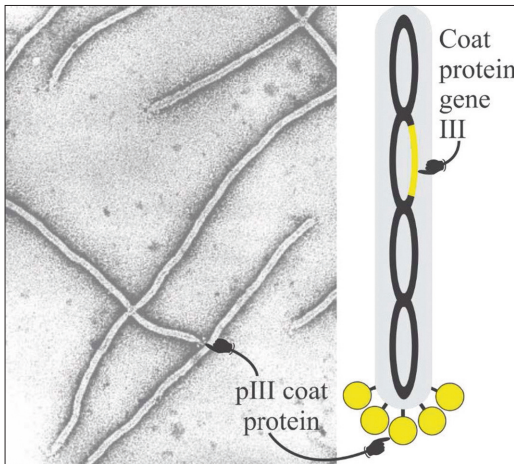
Figure 1. Electron micrograph of filamentous virions (left) and the schematic representation that will be used in this essay (right). The black links represent the phage chromosome; the only phage gene represented, by the yellow segment, is gene III, encoding phage coat protein pIII. Five copies of pIII cap one tip of the virion, with surface-exposed portions that are visible as irregular knobs at the tip of one particle in the electron micrograph (left-hand pointer), and that are represented by yellow circles in the schematic (right-hand pointer).

represented as a yellow segment in the chromosome; the other phage genes aren't represented. Similarly, only the surface-exposed part of pIII is explicitly represented as 5 yellow circles at one tip of the virion.

The infection process is extraordinarily efficient; when virions are mixed with a large excess of bacterial cells, the number of successfully infected cells is typically at least half the number of virions (infectivity ≥ 50 percent). It's also possible to artificially introduce "naked" phage DNA (stripped of its protein capsid) into bacterial cells, a procedure called *transfection*. Artificial introduction of the phage chromosome into the cell via transfection initiates the same infection program as does natural infection. Transfection is extremely inefficient (transfectivity many orders of magnitude less than 1 percent), but that inefficiency can be compensated by using huge numbers of input DNA molecules. Transfection is a staple of recombinant DNA technology, which involves manipulating naked DNA molecules in vitro before introducing them into cells.

FILAMENTOUS PHAGE DISPLAY IS BORN

From July 1983 to August 1984, I went on sabbatical to Bob Webster's lab in the Department of Biochemistry at Duke University. He was an accomplished investigator of filamentous phage biology, while I was a newcomer to the field. Much of my phage work before (Crissman and Smith, 1984; Nelson et al., 1981) and during the sabbatical was focused on coat protein pIII. The five copies of this protein were known to include an N-terminal domain that is exposed on the surface at one tip of the virion, and that is sometimes visible in electron micrographs as irregular knobs loosely connected to the bulk of the virion by very thin (usually invisible) threads (Armstrong et al.,

1981; Gray et al., 1981); the pointers in Figure 1 point to such knobs in the micrograph and in the accompanying schematic diagram.

It occurred to me that it might be possible to genetically fuse all or part of a foreign protein to the exposed parts of pIII without greatly impairing pIII's function in the phage infection cycle. If so, the foreign amino acids would be displayed at the tip of the virion, where they'd be accessible to macromolecules such as antibodies and receptors. This was hardly a great leap of imagination. Genetically fusing all or parts of two genes to make a recombinant gene encoding a recombinant "fusion protein" was a commonplace of recombinant DNA technology at the time (and today); choosing a phage coat protein as one of the fusion partners did not seem to stray very far from a well-traveled path.

Paul Modrich's lab down the hall from Webster's lab had three resources I could use to pursue my project: a purified protein called *Eco*RI, which he was studying at the time; the gene encoding the protein; and antibody against the protein. Figure 2 shows schematically how I used the *Eco*RI gene to produce the first filamentous phage-display construct. First, a DNA fragment from the *Eco*RI gene was inserted into the gene-III segment of naked phage DNA *in vitro*, resulting in a modified phage DNA with a recombinant gene III encoding an *Eco*RI-pIII fusion protein; foreign DNA fragments inserted into a chromosome in this way are called *inserts*. Second, the modified DNA was transfected into bacterial cells, which were then cultured in the hope that they would release modified virions bearing the fusion protein at one tip. The cells indeed released virions, which turned out to be infective – the first validation of my plan. I will call the modified virions "test phages."
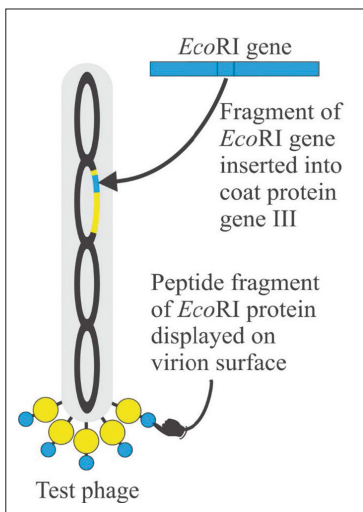

Figure 2. Construction of test phage. See text for details.

*Eco*RI gene

Fragment of *Eco*RI gene inserted into coat protein gene III

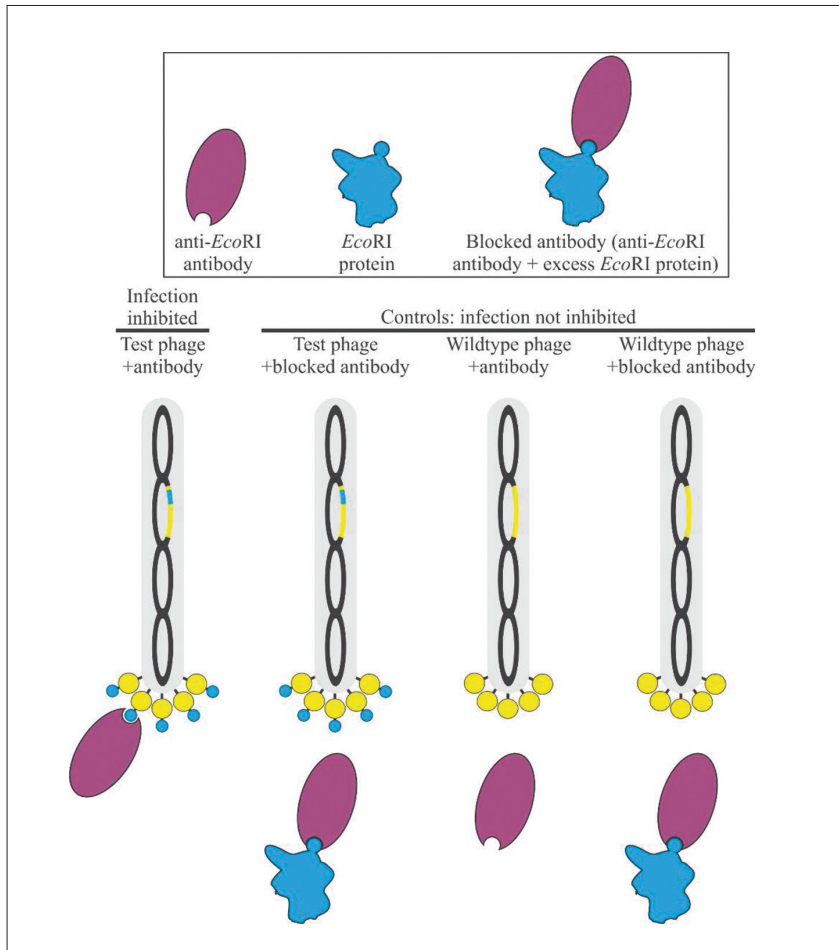Peptide fragment of *Eco*RI protein displayed on virion surface

Test phage

Figure 3. Proposed demonstration that EcoRI peptide was displayed on surface of test phage. See text for details.

According to my scheme, a peptide fragment of the *Eco*RI protein – the amino acids encoded by the insert – should be exposed at one tip of the virion surface; that fragment is symbolized by the blue circles in Figure 2. From now on in this essay, I'll refer to the surface-exposed amino acids encoded by a foreign DNA insert generically as a "peptide," regardless of the number of amino acids, and whether or not they constitute all or part of a natural protein. Figure 3 above shows schematically how I used Modrich's anti-*Eco*RI antibody and *Eco*RI protein to test this supposition. If the test phages were mixed with a large excess of the anti-*Eco*RI antibody (purple ovals), the antibody might well bind to the surface-exposed *Eco*RI peptide (blue circles), and thereby impair test-phage infectivity;

while the antibody should have no such effect on the infectivity of wild-type control phages. However, if the antibody was pre-treated with a large excess of the *Eco*RI protein (irregular blue shape) before being mixed with the phages, the protein would occupy the antibody's binding sites (indentations in the purple ovals), thus blocking its ability to impair test-phage infectivity. Figure 4 graphs the results of this experiment. As anticipated in Figure 3, mixing the antibody with the test phages caused a dramatic time-dependent decline in infectivity, while the antibody had no effect on the infectivity of wildtype control phages. Moreover, if the antibody was pre-reacted with excess EcoRI protein, it no longer affected infectivity of either the test phages or the wildtype control phages. These results showed that the peptide encoded by the insert was indeed exposed on the outside of the virion, where it was accessible to the antibody.

The results summarized in this section were reported in my first phage-display article, in which I called constructs like the test phage "filamentous fusion phage" (Smith, 1985).

REPLACING SCREENING WITH AFFINITY SELECTION

When I submitted my first phage-display publication, I had a specific application in mind: to replace laborious screening of small conventional phage expression libraries with easy affinity selection from enormous phage-display libraries. Explaining the difference between screening and selection will serve to illuminate the essential technological advance that phage display represents.

The signature conventional phage expression vector is λgt11 (Young and Davis, 1983), another phage that infects *E. coli* bacteria. Its chromo-
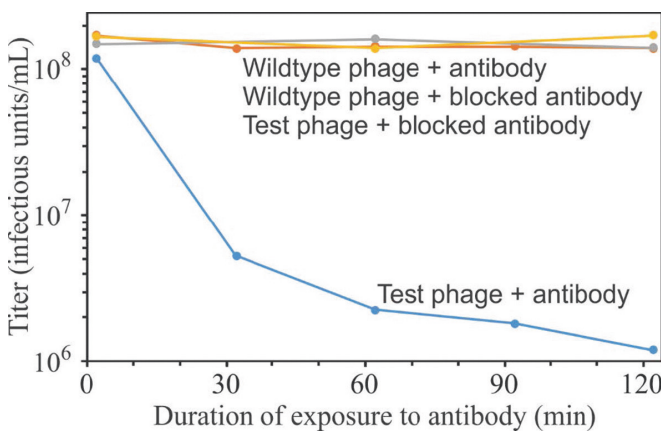

Figure 4. Results of test proposed in Figure 3. See text for details.

some has an extra, non-phage gene into which foreign DNA can be inserted. The resulting recombinant gene is a genetic fusion of the phage's extra gene and the foreign DNA insert. This recombinant gene is *expressed* in λgt11-infected cells: that is, the infected cells synthesize the fusion protein encoded by the recombinant gene. Phage-display constructs like the test phage in Figure 2 are also expression vectors, but there is a crucial difference: the fusion protein expressed in λgt11-infected cells accumulates in those cells but does not become part of the progeny virions they release.

When a limited number of λgt11 virions are mixed with a vast excess of bacteria and spread evenly over the surface of nutrient agar in a Petri dish, they form *plaques* after the dish has been incubated overnight at 37°C. This procedure is called *plating* the virions (or phages), "plate" here being a synonym for Petri dish. The plaques are visible as small circular zones of clearing in an opaque *lawn* of thick bacterial growth. Each clear zone stems from a single phage-infected cell, which releases hundreds of progeny virions after an hour or so, the cell itself being lysed in the process; the progeny virions immediately go on to infect hundreds of bacterial cells in the immediate neighborhood on the agar surface. A few rounds of infection, each increasing the number of lysed cells by a factor of hundreds, suffice to create a zone of cell lysis large enough to be visible as a clear plaque in the thick bacterial lawn. In its ability to form visible plaques, λgt11 resembles its wildtype ancestor λ as well as countless other classes of phage, including filamentous phages (Figure 5). All the virions in a plaque stem from a single infected cell, thus from a single progenitor virion; they therefore constitute a phage *clone*.
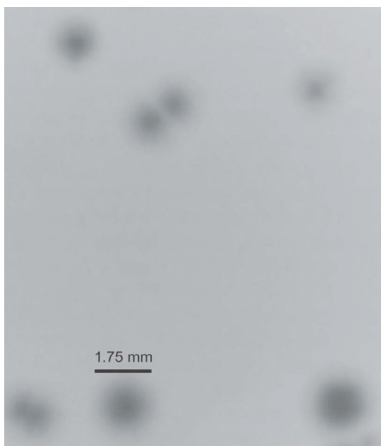

Figure 5. Plaques formed by wildtype filamentous phages. All the virions in one plaque constitute a clone. Unlike λ phages, filamentous phages don't kill the host cell, but they do slow down cell growth; their plaques are due to slowed growth, not cell lysis.

1.75 mm

If a large collection of diverse DNAs – for example, DNAs representing all or most of the protein-encoding genes in an organism – are inserted *en masse* into λgt11 phage DNA, the result is a highly diverse collection of phage chromosomes called an *expression library*. The members of the library are identical apart from the foreign DNA insert portions of their recombinant genes. If the DNA library is transfected *en masse* into bacterial cells, and those cells are then shaken overnight at 37°C in a culture flask, the virions released by the transfected (and subsequently infected) cells are also collectively called a library. At the virion stage, a library's members are clones, not individual DNAs; each clone is represented by hundreds of thousands or millions of identical virions stemming from a single successfully transfected cell. A very large λgt11 library would comprise up to 10 million different clones, representing up to 10 million different recombinant genes with different foreign DNA inserts.

"Library" is in a sense a misnomer for such a collection of 10 million phage "books," each represented by hundreds of thousands of identical phage copies. There is no analog of a card catalog, and all copies of all "books" are jumbled together indiscriminately in a single container. Nevertheless, if a researcher has available an antibody specific for a particular protein of interest, that antibody can be used as a *probe* to *screen* the library for rare phage clones expressing fusion proteins whose foreign fusion partner derives from the protein of interest. Those rare phage clones are the *targets* of the screening.

Screening a λgt11 expression library is demanding of both labor and technical skill (Mierendorf et al., 1987). First, library virions are plated on a large Petri dish at very high density (about 50,000 plaques per 150-mm dish); the resulting plaques are exceedingly crowded, unlike the plaques in Figure 5. Second, once plaques have begun to develop, a special kind of membrane disc is carefully laid over the agar surface and plaques are allowed develop for a few more hours; during this development, fusion protein molecules from each plaque become irreversibly immobilized on the immediately apposing surface of membrane. Third, the membrane and the agar surface are marked so that they can be realigned after being separated, and the membrane is carefully lifted off the agar surface so as not to disturb the plaques; at this stage, the membrane is called a *plaque lift*; the Petri dish is stored temporarily in the refrigerator. Fourth, the plaque lift is incubated with the antibody probe, allowing the antibody to bind specifically to fusion proteins from target plaques; unbound antibody molecules are washed off the lift, which is treated in several steps with reagents that allow the bound antibodies to be detected as small purple spots on the lift surface; each of these spots corresponds to a presumptive target plaque on the stored Petri dish. Fifth, the lift is realigned with the Petri dish and used to localize a small circular area of the agar

containing the presumptive target plaque (plus a few dozen non-target plaques). This circular area is excised as an agar plug, which is immersed in sterile medium and shaken vigorously to allow virions from the plaques to diffuse into the liquid. Sixth, the virions in the liquid are plated at much lower density – a hundred or so plaques per Petri dish – and the previous four steps are repeated, this time allowing presumptive target virions to be prepared in clonally pure form. The cloned virions are the gateway to further research on the target gene such as DNA sequencing. The scale of a screening project depends on the abundance of the target clones in the original library; a large project might involve a million original plaques on 20 Petri dishes, and even that may be too few for success.

Affinity selection from a phage-display library is in a sense the reverse of screening a conventional phage expression library with an antibody probe. In screening, the antibody probe in solution is used to bind specifically to hundreds of thousands of immobilized fusion proteins on the plaque lift. In affinity selection, which will be detailed in the next section, it's the antibody probe (or other specific binding protein) that's immobilized, and the billions of fusion proteins that are in solution. The exceedingly rare target fusion proteins bind specifically to the immobilized probe, and are thus captured on the immobilizing surface while all the unbound fusion proteins – the unwanted fusion proteins that aren't targets – are washed away. The captured target fusion proteins are then released from the immobilized probe and analyzed. The target fusion proteins are obtained in far too small yields for direct analysis. But that doesn't matter, because each target fusion protein is a coat protein that's attached to the virion whose recombinant coat-protein gene encodes it. In other words, the immobilized probe doesn't capture free-standing target fusion proteins; it captures whole infectious virions displaying the target fusion proteins on the virion surface. Those virions can be propagated to any desired scale simply by infecting fresh bacterial cells.

It's the physical linkage between each fusion protein and the infectious virion that encodes it that is the essential advance of phage display. It's that linkage that permits a vast increase in the number of phage clones that can be effectively surveyed. It's not necessary to prepare clones as separate (if very crowded) plaques on Petri dishes – a requirement that severely limits the number that can be screened. A probe immobilized on a solid support, such as the surface of a small empty plastic Petri dish, can readily be exposed to, say, ten trillion virions representing 100 billion clones. Even if only a handful of those 100 billion starting clones display a target fusion protein, a few rounds of affinity selection should suffice to clone them. Moreover, each round of affinity selection is far easier and less technically demanding than the screening procedure outlined above.

Although the physical linkage explained in the previous paragraph

continues to be understood as the essence of phage display, my vision of the technology's field of application now seems very parochial indeed. Not until 1988 were my horizons greatly expanded, as I'll describe below.

AFFINITY SELECTION

As I returned to Missouri at the end of my sabbatical in August 1984, I resolved to refocus my research program on the phage-display idea. Three practical issues needed to be addressed: development of an efficient phage-display vector – one that would make it easy to insert foreign DNAs into the coat protein gene without unduly impairing phage function; development of an effective procedure for affinity selection; and development of methods for creating huge phage-display libraries, so that the potential of affinity selection could be fully exploited. The first two tasks were the dissertation project taken on by Steve Parmley, a graduate student who came to my lab at that time; his results were published three years later (Parmley and Smith, 1988). The fUSE vector system he introduced continues to be used today, especially vector fUSE5 (GeneBank Accession AF218364). The subject of this section is the affinity selection process introduced in my 1985 article and greatly improved in Parmley's dissertation project.

Affinity selection was closely modeled on "panning" (Wysocki and Sato, 1978), a staple of cellular immunology research of which I was well aware at the time (I didn't think to acknowledge this debt in any of my previous publications). In immunological panning, it is cells rather than virions that are captured by immobilized antibodies, which bind specifically to antigens on the surface of the cells. At first, we ourselves used the term "panning" (or "biopanning" if the super-strong biotin-streptavidin interaction was involved) before settling on the more suitably generic term affinity selection.

Parmley's affinity selection experiments closely followed the cellular immunology model. Thus the capturing molecules were antibodies, which were immobilized by non-specific adsorption to the polystyrene surface of Petri dishes or microplates. Since then, many different types of molecules, which we call generically *selectors*, immobilized on a variety of substrates by various types of bonding, have served to capture target virions.

Figure 6 below diagrams schematically the steps in a round of affinity selection. *Step 1*: Selector is immobilized by attachment to an immobilizing substrate, such as the surface of a polystyrene Petri dish or microplate well. Stringency can be adjusted by changing the density of immobilized selector on the substrate surface, as explained later in this section. *Step 2*: Input virions are added, so that virions whose displayed peptides bind the
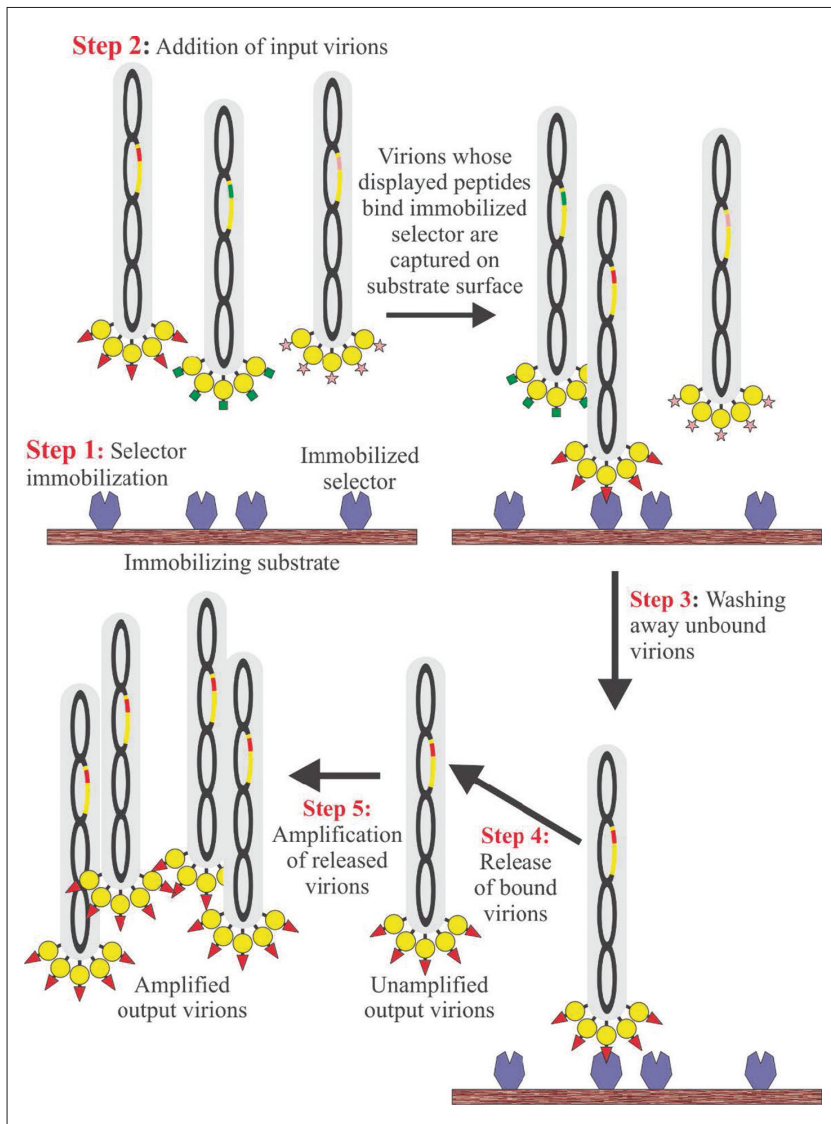
Figure 6. Schematic diagram of affinity selection. Steps will be detailed in the text.

receptor are captured on the substrate surface. The input to the first round of affinity selection is the starting unselected library; this input might consist of up to 10–100 trillion virions representing up to 100 billion clones. The inputs to each subsequent round of selection are the amplified output virions (Step 5 below) from the previous round of selection. *Step 3*: The substrate surface is thoroughly washed to remove virions that are not captured by immobilized selector. *Step 4*: Captured virions

that remain after washing are released by weakening the bond between the selector and the displayed peptide; alternatively (not shown; see next paragraph), they may be released by cleaving the segment of amino acids linking the displayed peptide to the rest of the virion. The released virions are the unamplified output of selection. *Step 5*: The unamplified output virions are "amplified" by infecting fresh bacterial cells and shaking the cells overnight in culture medium. Each virion in an unamplified output gives rise to millions or billions of identical progeny virions in the amplified output. Amplified output virions are the input virions for the next round of affinity selection. Virions from either unamplified or amplified outputs can be cloned so that individual displayed peptides can be sequenced (by sequencing the DNA inserts that encode them) and analyzed in countless other ways.

Releasing captured virions from the immobilized selector is a key step in affinity selection (step 4 in Figure 6). At first this was accomplished by elution in acid, which filamentous phages tolerate without losing infectivity. That was another methodological choice arising from my immunological background: acidic elution buffer had long been used in immunoaffinity chromatography to release antibodies from immobilized antigens or antigens from immobilized antibodies. Acid greatly weakens the interaction between antibodies and most antigens, generally without irreversibly inactivating either binding partner. However, many practitioners have come to appreciate the advantages of releasing phages using the intestinal protease trypsin to cleave a bond between the displayed insert-encoded peptide and the remainder of the virion, as summarized by my graduate student Will Thomas and me in a critical review (Thomas and Smith, 2010). Filamentous virions have been known since the 1960s to be extremely resistant to trypsin and other intestinal proteases (Salivar et al., 1964), as expected for a phage that infects enteric bacteria like *E. coli*. Thomas and I added the trypsin-release phage-display vector f3TR1 (GeneBank Accession HM355479) to the fUSE family.

Affinity selection is not perfectly discriminatory; the yield of a binding clone might be only 10,000 times higher than the yield of each non-binding clone. If binding clones comprise fewer than one in a million clones in the input virions, therefore, they may well not be represented in the small sample of a hundred or so cloned output virions that are typically subject to further analysis. Consequently, one or two additional rounds of affinity selection are normally carried out, the input virions for each round being amplified output virions from the previous round. Only then are individual cloned output virions analyzed. More recently, next-gen sequencing (NGS) makes it possible to sequence the DNA inserts of many millions of phages, allowing identification of particularly abundant clones in ear-

ly-round outputs, even when they comprise far fewer than 1 percent of
the total ('t Hoen et al., 2012).

Quite apart from imperfect discrimination, any practical affinity-selec-
tion strategy must confront the possibility of "target-unrelated phages"
(TUPs) – that is, phage clones that are favored for reasons other than
affinity of their displayed peptide for the immobilized selector (unlike
elsewhere in this essay, the "target" referred to in the definition of TUP is
a synonym for the selector rather than displayed peptides that bind the
selector). Thomas and I, along with our colleague Miriam Golomb, have
published a critical review of the TUP problem (Thomas et al., 2010). The
ability to distinguish TUPs from genuine selector-binding clones is an
additional attraction of NGS as the read-out for affinity selections ('t
Hoen et al., 2012).

Stringency is a key parameter of affinity-selection strategy. It refers to
the degree to which selection favors virions displaying high-affinity pep-
tides over virions displaying low-affinity peptides. Stringency can be
increased in several ways, such as decreasing the density of selectors on
the immobilizing substrate (Step 1 in Figure 6). High stringency must
almost always be purchased at the cost of reduced *yield* – the percent of
binding virions in the input that are successfully recovered in the output.
In the first round of affinity selection, yield is all-important: in a very
large input library, each clone (thus each displayed peptide) is repre-
sented by a limited number of virions in the input, and any clone that is
lost in this round cannot be recovered in subsequent rounds. In the out-
put of the first round of selection, thus in the input to the second round,
the number of clones represented has been reduced by many orders of
magnitude compared to the unselected library, and each clone is therefore
represented by orders of magnitude more virions. Only in the second and
subsequent rounds, therefore, does it make sense to increase stringency
in order to favor virions displaying high-affinity peptides or proteins.

1988 – A YEAR OF DISCOVERY

The title of this section is from a preface I wrote to a 1996 phage-display
laboratory manual (Smith, 1996). There I described three encounters –
two personal and one with a published article – that greatly expanded my
vision for the phage-display idea in 1988.

The first encounter occurred in January, when I visited Vidal de la
Cruz and Tom McCutchan in the Laboratory of Parasitic Diseases at the
National Institutes of Health in Bethesda, Maryland. Following my 1985
article, they had constructed phages displaying antigens from the malaria
parasite *Plasmodium falciparum* (de la Cruz et al., 1988). In the course our
conversation, McCutchan casually remarked "you know, you could use

phage to out-Geysen Geysen." Mario Geysen, I learned, had developed
methods for chemically synthesizing hundreds of short peptides simulta-
neously, and a clever plan for using this technology to delineate the pep-
tide *epitope* recognized by an antibody, without the need for advance
knowledge of the antibody's specificity (Geysen et al., 1986a, b); an
epitope is the part of an antigen that makes close contact with an anti-
body's binding site. Geysen couldn't accomplish this task by testing all
possible short peptides individually: even examining all possible hexa-
peptides would have required him to synthesize and test 64 million com-
pounds. Instead, he tested 400 peptide mixtures, each mixture having
specified amino acids at two positions and an equal mixture of all the
amino acids at the other positions. A position with a mixture of amino
acids is said to be *degenerate*, and a peptide mixture with degenerate posi-
tions is called a degenerate peptide. The degenerate peptide showing the
best binding then serves as the starting-point for synthesis of a second
series of 400 new degenerate peptides, this time with the optimal amino
acids at the first two non-degenerate positions and 400 combinations of
specified amino acids at two additional positions (thus reducing the num-
ber of degenerate positions by two). In this way, he could in theory pro-
gressively narrow the antibody's binding specificity to a single peptide
sequence.

It was obvious to all three of us – de la Cruz, McCutchan, and me –
that phage display might well allow 64 million (or many more) peptides to
be surveyed individually for affinity for an antibody – not by testing them
one by one, but by using immobilized antibody to select virions displaying
binding peptides from a phage-display library displaying tens of millions
(or many more) random peptides altogether. Furthermore, synthetic DNA
inserts encoding tens of millions (or many more) random peptides could
be purchased cheaply from chemical DNA synthesis companies. The
companies wouldn't have to synthesize them one by one. Instead, they'd
synthesize them all at once in the form of a degenerate oligonucleotide.
Just as the synthetic degenerate peptides in the previous paragraph had
mixtures of amino acids at various positions in the peptides, so the
degenerate oligonucleotide would have a mixture of nucleotides at each
position in the peptide coding sequence. Such degenerate oligonucleo-
tides were already in widespread use in molecular biology.

Construction of a random peptide library would be easy. We'd start
with a synthetic DNA insert in which each three-position codon would
have the degenerate sequence NNK, where N stands for an equal mixture
of all four nucleotides and K stands for an equal mixture of G and T. An
NNK degenerate codon is thus an equal mixture of 32 codons, including
the TAG stop (nonsense) codon and 31 sense codons that together include
codons for all 20 amino acids. An insert with six such codons would com-

prise 32^6 = 1.074 billion sequences altogether, including 31^6 = 888 million sequences encoding all possible 64 million six-amino-acid peptides without stop codons. Random peptide libraries, which we called "epitope libraries" at the time, seemed a much more promising application of phage display than libraries of natural proteins or protein fragments. The epitope library concept was included at the last minute in the 1988 article (Parmley and Smith, 1988); and soon became the main focus of my research program, as I'll described in the next section.

The second encounter, in October 1988, was with an article reporting the construction of functional single-chain antibodies (Bird et al., 1988). I was electrified by this report. If 240-amino acid single-chain antibodies could be displayed on the virion surface (and the prospects seemed reasonably good), immobilized antigens could be used to affinity-select specific antibodies from synthetic phage-antibody libraries, completely bypassing the natural immune system. I soon learned that the same vision occurred independently to a number of other investigators. Indeed, Robert Ladner of Genex Corporation had fully articulated the phage-antibody concept in a patent application a year and a half earlier, though he chose phage λ rather than filamentous phage as the prospective display vector (Ladner et al., 1987). In any case, it is others, including my co-laureate Greg Winter and his colleagues, who have succeeded in bringing this vision to a markedly successful reality.

The third encounter, in December 1988, occurred in the office of Jim Larrick at Genelabs, a biotech company in Redwood City, California. "Drugs," he said cryptically after brief greetings. Larrick's message was that using cellular receptors or other medically significant biomolecules as immobilized selectors might open up a new gateway to drug discovery. This was a dramatic expansion of my horizons, which had hitherto been largely limited to immunological concerns. That's when we came to use the generic term "selector," and when selectors other than antibodies and antigens came to figure prominently in our plans and publications.

AFFINITY SELECTION FROM RANDOM PEPTIDE LIBRARIES

Our first attempts at constructing a random peptide library were undertaken in the 1988–1989 academic year by Shannon Flynn as part of his undergraduate research project. He ran up against a severe bottleneck: inability to transfect bacterial cells with naked DNA in sufficient numbers to put the phage-display concept to the test. Transfection was a necessary step in library construction, since inserting foreign DNA into the phage coat protein gene requires manipulation of naked phage DNA. At the end of 1988, Bill Dower and his colleagues reported a practical method for very large-scale transfection of bacterial cells (Dower et al., 1988), and

work on large phage-display libraries began in earnest. That's when two key colleagues came to my lab: Robert Davis and Jamie Scott.

Davis served as my lab manager and chief technician for over two decades, reducing to practice several of the large-scale routines in the lab. One of the most important routines was a radioactive DNA sequencing strategy tailored specifically to the short inserts encoding random peptides (Haas and Smith, 1993). Davis was able to sequence almost 800 phage clones per week in this way, with time left over to take on other tasks as well. We estimate that he generated well over 1 million bases of sequence altogether before we finally started outsourcing sequencing to the University's DNA core facility – at considerable increase in cost.

Meanwhile, Scott was a postdoc, who, with Davis's expert technical help, constructed vector fUSE5 and used it to create our first large random peptide library: f3-6mer (GeneBank Accession AF246446), which has 200 million clones and displays random 6-amino acid peptides (6mers). With the aid of our colleague Hannah Alexander, who had come to Missouri from Richard Lerner's lab at the Scripps Research Institute, we obtained from that lab samples of two monoclonal antibodies specific for a known peptide epitope (Fieser et al., 1987). The sequences of the 6mer peptides Scott affinity-selected from the library closely matched the antibodies' known epitope, even though knowledge of that epitope had in no way influenced construction of the library or conduct of the experiment. This was a dramatic validation of the phage-display idea, and the essential breakthrough recognized by my Nobel Prize award. Scott's results were published in *Science* in July 1990 (Scott and Smith, 1990), simultaneously with analogous results from two other labs (Cwirla et al., 1990; Devlin et al., 1990).

IN VITRO EVOLUTION OF S-PEPTIDE ANTAGONISTS BY AFFINITY MATURATION

In 1992, David Schultz and John Ladbury brought the ribonuclease S protein/S peptide system to my lab (Smith et al., 1993). In the 1950s, Fred Richards at Yale (Steitz, 2009) found that partial digestion of bovine pancreatic ribonuclease A with the protease subtilisin cleaves the ribonuclease polypeptide chain at a single position, between amino acids 20 and 21. The cleaved protein retains its enzymatic activity, but when the two fragments – the 20 amino acid S peptide and the 104 amino acid S protein – are separated, as shown in Figure 7, neither one is enzymatically active. When the two fragments are mixed, however, S-peptide binds to S-protein, and enzymatic activity is restored.
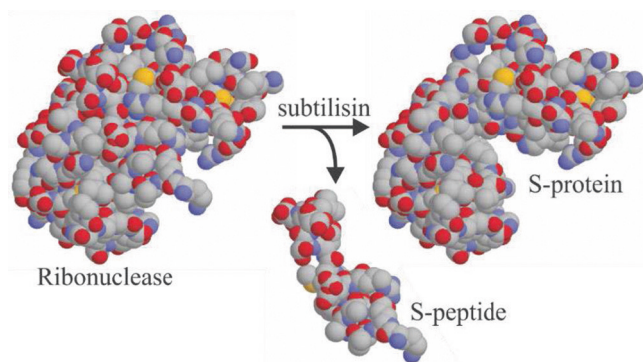
Figure 7. Cleavage of ribonuclease with subtilisin and separation of S-protein and S-peptide moieties. S-protein served as the model "receptor" and S-peptide as its natural ligand (e.g., "hormone") in Jinan Yu's affinity maturation study. Space-filling CPK models created from X-ray crystallographic coordinates in Protein Data Bank accession 1KF5.

For several years, S-protein was our selector molecule of choice in further development of phage-display technology. We presented it as a model "receptor," and S-peptide as its cognate natural ligand (e.g., "hormone"). The "physiological effect" that ensued when the "hormone" bound its "receptor" was restoration of enzyme activity. Peptides affinity-selected from a random peptide library by immobilized S-protein selector would be candidate "drugs" that agonize or antagonize the S-peptide "hormone," or in some other way modulate physiological activity. "In short," we wrote (Smith et al., 1993), "the S-protein/S-peptide model preserves the essential features of a pharmacologically significant receptor/ligand couple but is convenient for methodological development since both components can be purchased or prepared inexpensively in large quantities."

My graduate student Jinan Yu used this model system in an extensive study of *affinity maturation*, a strategy for *in vitro* evolution of high-affinity ligands from random peptide libraries (Yu and Smith, 1996). Her starting library, f3-15mer (GenBank Accession AF246445), generously provided by Japanese researchers (Nishi et al., 1996), has 250 million clones, and presumably displays the same number of different 15 amino acid random peptides (15mers). Those 15mers are an exceedingly sparse sampling of all 3.3×10^{19} possible 15mers. A library's "champion" – the displayed 15mer with the highest affinity for the S-protein selector – would almost certainly have much lower affinity than the best possible 15mer.

Affinity maturation is a strategy to overcome extreme sparseness in input random peptide libraries. It entails introducing random mutations into the DNA inserts during the amplification step (step 5 in Figure 6) of

each round of selection except the last, while gradually increasing stringency in successive rounds of selection. Mutation creates a "clan" of closely-related mutant peptides from each peptide in the unamplified output, thus strategically increasing the number of peptides represented in the input to the next round of selection. Meanwhile, increasing stringency gradually, rather than immediately after the first round, opens up the possibility of discovering a "dark horse" peptide: a peptide with higher affinity than any peptide that arises through mutation of the initial champion.

Addition of ongoing mutation sharpens the parallelism between affinity selection of phage-displayed peptides and natural selection of organisms in the living world. The resulting peptide ligands can justifiably be said to have "evolved" from their peptide ancestors in the starting unselected library.

Affinity maturation is a term borrowed from immunology. In the course of the natural antibody response to an antigen, the affinity of circulating antibodies for the antigen improves. This improvement stems from hypermutation specifically targeted to the antibody genes in each B-cell clone that responds to the antigen, along with gradually increasing strength of selection in favor of clones with high-affinity antibodies as the concentration of antigen in the body steadily declines. To me, as to many other practitioners familiar with basic immunology, the extension from B-cell clones to phage-display clones was obvious. Unsurprisingly, this was especially true of researchers such as my co-laureate Greg Winter and his colleagues who were affinity selecting antibodies from phage-antibody libraries (Low et al., 1996).

Affinity maturation substantially increases the labor investment in an affinity selection project. Winter reminds us in his Nobel lecture that the need for affinity maturation (with its attendant cost in labor) stems from the limited number of clones in the initial library (Winter, 2018). He and his colleagues have now constructed phage-antibody libraries that are so large that high-affinity antibodies can be directly affinity-selected with many antigens without the need for affinity maturation.

Yu's affinity maturation experiments did not reveal any dark horse peptides, but they did reveal mutants of the library's initial champion with substantially higher affinity for the S-protein selector than the champion itself (Figure 8). All the affinity-selected peptides could be aligned with the natural ligand, S-peptide, such that the four amino acids in S-peptide that make close contact with S-protein in the S-peptide/S-protein complex are matched in the selected peptides. This is not unexpected, of course, but it's important to understand that the experiment might well have turned out differently: it might have revealed peptides that bind the S-protein selector in an entirely different way than does the natural ligand, and that don't have any discernible sequence similarity to that natural ligand.

	Buried residues	K_D
	| |||	
S-peptide (natural ligand)	**KETAAAKFERQHMDSSTSAA**	6 nM
Champion peptide	**NR.WSEFLWQHLAPV**	500 nM
Affinity-matured peptide 72	**..G........H...**	30 nM
Affinity-matured peptide 88	**.....V.Q...I..A**	12 nM

Figure 8. Amino acid sequences using one-letter abbreviations for S-protein-binding peptides from Jinan Yu's affinity-maturation study, aligned with S-protein's natural ligand S-peptide. The "champion" peptide was affinity-selected from the f3-15mer library without mutagenesis; peptides 72 and 88, which are presumably mutant descendants of the champion peptide, were obtain through affinity maturation, as described in the text. Amino acids in peptides 72 and 88 that match those in the champion peptide are rendered as dots (·). The four amino acids in S-peptide that make close contact with S-protein in intact ribonuclease are colored red, as are the corresponding letters or dots in the affinity-selected peptides. L (leucine) and I (isoleucine) are considered matches to the M (methionine) in S-peptide, since S-peptide variants with those substitutions are still able to bind S-protein. Approximate dissociation equilibrium constants K_D for binding to S-protein are given for each peptide; these constants are inversely related to affinity.

The ligands Yu discovered have no known practical worth apart from their heuristic value. But ligands for other, pharmacologically significant selectors have been affinity-selected from the same library, including in the lab of my long-time University of Missouri colleague Sue Deutscher (Glinsky et al., 2000; Kumar et al., 2007; Landon and Deutscher, 2003; Landon et al., 2004a; Landon et al., 2004b; Larimer and Deutscher, 2014; Peletskaya et al., 1997; Peng et al., 2017; Soendergaard et al., 2014a, b; Zou et al., 2004). Indeed, is there any limit to the number of different ligands, some perhaps of great practical worth, that might evolve from the very same library by imposing different selection regimes, without the need for deep prior knowledge? Members of the phage-display community have constructed many additional libraries of innovative design, from which many valuable ligands have been selected, including the therapeutic antibodies described in Winter's Nobel lecture (Winter, 2018).

A COMMUNITY ACHIEVEMENT

The phage-display concept evolved gradually, and continues to evolve, in a phage-display community of which I'm only one of numerous members. At each step along my own lab's evolutionary path, an idea or experimental result that I and my coworkers had access to by virtue of our membership in overlapping global scientific communities – phage biology, immunology, molecular biology, evolutionary biology, others – brought us within sight of some incremental advance. The same is surely true of the

other members of the community. At no point – certainly not when my first article was published (Smith, 1985) – was the significance of phage display revealed in anything like full clarity, as I have detailed in my own case for the year 1988. We did experience *eureka* moments of exultation, such as when Scott and I saw the results of our first large-scale affinity selections (Scott and Smith, 1990). But evolution of the phage-display concept certainly didn't stop there. Indeed, its qualification as a discovery or improvement meeting the standard in Alfred Nobel's Testament has been the collective achievement of the phage-display community, who have applied and modified it in countless innovative ways, very few of which I or any other individual member could possibly have foreseen. It is accordingly on behalf of that community that I have accepted this great honor.

REFERENCES

't Hoen, P.A., Jirka, S.M., Ten Broeke, B.R., Schultes, E.A., Aguilera, B., Pang, K.H., Heemskerk, H., Aartsma-Rus, A., van Ommen, G.J., and den Dunnen, J.T. (2012). Phage display screening without repetitious selection rounds. *Analytical Biochemistry* **421**, 622–631.

Armstrong, J., Perham, R.N., and Walker, J.E. (1981). Domain structure of bacteriophage fd adsorption protein. *FEBS Lett* **135**, 167–172.

Bird, R.E., Hardman, K.D., Jacobson, J.W., Johnson, S., Kaufman, B.M., Lee, S.M., Lee, T., Pope, S.H., Riordan, G.S., and Whitlow, M. (1988). Single-chain antigen-binding proteins [published erratum appears in *Science* 1989 Apr 28; **244** (4903): 409]. Science **242**, 423–426.

Crissman, J.W., and Smith, G.P. (1984). Gene-III protein of filamentous phages: evidence for a carboxyl-terminal domain with a role in morphogenesis. *Virology* **132**, 445–455.

Cwirla, S.E., Peters, E.A., Barrett, R.W., and Dower, W.J. (1990). Peptides on phage: A vast library of peptides for identifying ligands. *Proceedings of the National Academy of Sciences of the United States of America* **87**, 6378–6382.

de la Cruz, V.F., Lal, A.A., and McCutchan, T.F. (1988). Immunogenicity and epitope mapping of foreign sequences via genetically engineered filamentous phage. *Journal of Biological Chemistry* **263**, 4318–4322.

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

Dower, W.J., Miller, J.F., and Ragsdale, C.W. (1988). High efficiency transformation of E. coli by high voltage electroporation. *Nucleic Acids Research* **16**, 6127–6145.

Fieser, T.M., Tainer, J.A., Geysen, H.M., Houghten, R.A., and Lerner, R.A. (1987). Influence of protein flexibility and peptide conformation on reactivity of monoclonal anti-peptide antibodies with a protein alpha-helix. *Proceedings of the National Academy of Sciences of the United States of America* **84**, 8568–8572.

Geysen, H.M., Rodda, S.J., and Mason, T.J. (1986a). The delineation of peptides able to mimic assembled epitopes. *Ciba Foundation Symposium* **119**, 130–149.

Geysen, H.M., Rodda, S.J., and Mason, T.J. (1986b). A priori delineation of a peptide which mimics a discontinuous antigenic determinant. *Molecular Immunology* **23**, 709–715.

Glinsky, V.V., Huflejt, M.E., Glinsky, G.V., Deutscher, S.L., and Quinn, T.P. (2000). Effects of Thomsen-Friedenreich antigen-specific peptide P-30 on beta-galactoside-mediated homotypic aggregation and adhesion to the endothelium of MDA-MB-435 human breast carcinoma cells. *Cancer Res* **60**, 2584–2588.

Gray, C.W., Brown, R.S., and Marvin, D.A. (1981). Adsorption complex of filamentous fd virus. *Journal of Molecular Biology* **146**, 621–627.

Haas, S.J., and Smith, G.P. (1993). Rapid sequencing of viral DNA from filamentous bacteriophage. *Biotechniques* **15**, 422–424, 426–428, 431.

Kumar, S.R., Quinn, T.P., and Deutscher, S.L. (2007). Evaluation of an 111In-radiolabeled peptide as a targeting and imaging agent for ErbB-2 receptor expressing breast carcinomas. *Clin Cancer Res* **13**, 6070–6079.

Ladner, R.C., Glick, J.L., and Bird, R.E. (1987). Method for the preparation of binding molecules (https://patentimages.storage.googleapis.com/9a/70/56/d588700bb3fc66/WO1988006630A1.pdf)

Landon, L.A., and Deutscher, S.L. (2003). Combinatorial discovery of tumor targeting peptides using phage display. Journal of cellular biochemistry 90, 509–517.

Landon, L.A., Harden, W., Illy, C., and Deutscher, S.L. (2004a). High-throughput fluorescence spectroscopic analysis of affinity of peptides displayed on bacteriophage. *Anal Biochem* **331**, 60–67.

Landon, L.A., Zou, J., and Deutscher, S.L. (2004b). Is phage display technology on target for developing peptide-based cancer drugs? *Current drug discovery technologies* **1**, 113–132.

Larimer, B.M., and Deutscher, S.L. (2014). Development of a peptide by phage display for SPECT imaging of resistance-susceptible breast cancer. *Am J Nucl Med Mol Imaging* **4**, 435–447.

Low, N.M., Holliger, P.H., and Winter, G. (1996). Mimicking somatic hypermutation: Affinity maturation of antibodies displayed on bacteriophage using a bacterial mutator strain. *Journal of Molecular Biology* **260**, 359–368.

Mierendorf, R.C., Percy, C., and Young, R.A. (1987). Gene isolation by screening lambda gt11 libraries with antibodies. *Methods in Enzymology* **152**, 458–469.

Nelson, F.K., Friedman, S.M., and Smith, G.P. (1981). Filamentous phage DNA cloning vectors: a noninfective mutant with a nonpolar deletion in gene III. *Virology* **108**, 338–350.

Nishi, T., Budde, R.J., McMurray, J.S., Obeyesekere, N.U., Safdar, N., Levin, V.A., and Saya, H. (1996). Tight-binding inhibitory sequences against pp60(c-src) identified using a random 15-amino-acid peptide library. *FEBS Lett* **399**, 237–240.

Parmley, S.F., and Smith, G.P. (1988). Antibody-selectable filamentous fd phage vectors: Affinity purification of target genes. *Gene* **73**, 305–318.

Peletskaya, E.N., Glinsky, V.V., Glinsky, G.V., Deutscher, S.L., and Quinn, T.P. (1997). Characterization of peptides that bind the tumor-associated Thomsen-Friedenreich antigen selected from bacteriophage display libraries. *J Mol Biol* **270**, 374–384.

Peng, Y., Prater, A.R., and Deutscher, S.L. (2017). Targeting aggressive prostate cancer-associated CD44v6 using phage display selected peptides. *Oncotarget* **8**, 86747–86768.

Salivar, W.O., Tzagoloff, H., and Pratt, D. (1964). Some physical-chemical and biological properties of the rod-shaped coliphage M13. *Virology* **24**, 359–371.

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

Smith, G.P. (1985). Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. *Science* **228**, 1315–1317.

Smith, G.P. (1996). 1988 – A year of discovery. *In Phage Display of Peptides and Proteins: A Laboratory Manual*, B.K. Kay, J. Winter, and J. McCafferty, eds. (New York, Academic Press), pp. xvii–xix.

Smith, G.P. (2018). Nobel lecture in Chemistry, 2018: Phage display: simple evolution in a Petri dish (https://www.nobelprize.org/prizes/chemistry/2018/smith/lecture/) (Stockholm, Sweden, Nobel Foundation).

Smith, G.P., and Petrenko, V.A. (1997). Phage Display. *Chemical Reviews* **97**, 391–410.

Smith, G.P., Schultz, D.A., and Ladbury, J.E. (1993). A ribonuclease S-peptide antagonist discovered with a bacteriophage display library. *Gene* **128**, 37–42.

Soendergaard, M., Newton-Northup, J.R., and Deutscher, S.L. (2014a). In vitro high throughput phage display selection of ovarian cancer avid phage clones for near-infrared optical imaging. *Combinatorial chemistry & high throughput screening* **17**, 859–867.

Soendergaard, M., Newton-Northup, J.R., and Deutscher, S.L. (2014b). In vivo phage display selection of an ovarian cancer targeting peptide for SPECT/CT imaging. *Am J Nucl Med Mol Imaging* **4**, 561–570.

Steitz, T.A. (2009). Retrospective. Frederic M. Richards (1925–2009). *Science* **323**, 1181.

Thomas, W., and Smith, G. (2010). The case for trypsin release of affinity-selected phages. *Biotechniques* **49**, 651–654.

Thomas, W.D., Golomb, M., and Smith, G.P. (2010). Corruption of phage display libraries by target-unrelated clones: Diagnosis and countermeasures. *Anal Biochem* **407**, 237–240.

Winter, G.P. (2018). Nobel Lecture in Chemistry, 2018: Harnessing evolution to make medicines (https://www.nobelprize.org/prizes/chemistry/2018/winter/lecture/) (Stockholm, Sweden, Nobel Foundation).

Wysocki, L.J., and Sato, V.L. (1978). "Panning" for lymphocytes: A method for cell selection. *Proceedings of the National Academy of Sciences of the United States of America* **75**, 2844–2848.

Young, R.A., and Davis, R.W. (1983). Efficient isolation of genes by using antibody probes. *Proceedings of the National Academy of Sciences of the United States of America* **80**, 1194–1198.

Yu, J., and Smith, G. (1996). Affinity maturation of phage-displayed peptide ligands. *Methods in Enzymology* **267**, 3–27.

Zou, J., Dickerson, M.T., Owen, N.K., Landon, L.A., and Deutscher, S.L. (2004). Biodistribution of filamentous phage peptide libraries in mice. *Mol Biol Rep* **31**, 121–129.