Nobel Prize in Chemistry – 2018

Speeding Up Protein Evolution

Purnananda Guptasarma

The 2018 Nobel Prize for Chemistry rewards research on the use of bacteria and viruses to generate and screen highly diverse protein sequences for improved catalytic and ligand-binding function. One half of the Prize was awarded to Professor Frances Arnold of the California Institute of Technology (California, USA). The other half was awarded jointly to Professor George P Smith of the University of Missouri (Columbia, USA) and Professor Sir Gregory P Winter of the Medical Research Council's Laboratory of Molecular Biology (Cambridge, UK). The three winners have been amongst the tallest of stalwarts in combinatorial approaches to protein engineering.

Purnananda Guptasarma is Professor and Dean of Research and Development at the Indian Institute of Science Education and Research Mohali. He is also Director of the Centre of Excellence in Protein Science, Design and Engineering at IISER Mohali. He was the Founding Head of the Department of Biological Sciences. His work lies mainly in the areas of protein science and engineering, and protein design, using the approaches of biophysics, biochemistry, molecular and cellular biology and bioinformatics.

Keywords
Phage-display, bacteriophage, error-prone PCR, mutation, phagemid, combinatorial library, biopanning, gene amplification, antibodies, protein engineering.

1. Introduction

Frances Arnold laid the foundations of the field of directing and speeding-up the evolution of enzymes, using rapid and successive cycles of genetic (mutational) changes in genes encoding enzymes of interest. Her methods helped in increasing both the rate at which changes are introduced in the amino acid sequences of proteins/enzymes and also the speed with which these changes can be selected for, in the laboratory, in respect of the improvements they confer upon the functioning of an enzyme in a chemical process. These methods vastly increased the scope available for the development of robust enzymatic reagents for the synthesis of chemicals. They also helped in improving enzymes used to hydrolyse plant-based biomass or waste, with implications for biofuels.

George Smith was the first to demonstrate the possibility of combining and physically linking the genotype of a phage to its

Phage-display library selection made it possible for a selected variant displaying a binding capability to be biologically amplified by being 'grown' in large numbers within bacteria producing the variant 'clonally'.

phenotype through phage-display. His methods allowed the creation of libraries consisting of billions of phage particles, with each phage genome containing a gene encoding a variant amino acid sequence of a peptide, with the surface of the very same phage particle actually displaying the protein, or peptide, bearing the said variant amino acid sequence. The technique of screening of a vast library of such phage particles for variants possessing the ability to bind to a particular ligand revolutionized the field of combinatorial chemistry, because phage-display library selection made it possible for a selected variant displaying a binding capability to be biologically amplified by being 'grown' in large numbers within bacteria producing the variant 'clonally'.

Gregory Winter modified and adapted the technique of phage display to facilitate screening of vast libraries of large proteins (rather than short peptides). His creation of a human antibody library on phage revolutionized methods for the generation, and evolution, of antibodies for research, diagnostics and therapeutic applications.

An Interesting Link

All antibodies of therapeutic nature which have been used during the last two decades (or more), including the antibodies used for the treatment cancers, happen to be either derived from, or have benefited from, the screening of a phage-display antibody library.

There is an interesting link between this year's Nobel Prizes in Chemistry and Physiology or Medicine. The Prizes for the latter were awarded for the development of the novel concept of immune checkpoint therapy in cancer treatment. This approach involves the use of therapeutic antibodies which neutralize the function of immune checkpoint proteins (such as CTLA-4, or PD-1). These checkpoint proteins act as 'brakes' on the functioning of immune cells known as T-cells. When antibodies cognate to CTLA-4 or PD-1 are used to block their normal functioning, they no longer act as brakes on T-cell proliferation and development. This improves the body's natural immunity, facilitating better immune surveillance and removal of developing cancers. Now, it is interesting to note that all antibodies of therapeutic nature which have been used during the last two decades (or more), including the antibodies used for the treatment of every kind of cancer, happen to be either derived from, or have benefited from, the

screening of a phage-display antibody library, and from the selection of a part of their own amino acid sequence(s) through such screening. This can be directly traced back to the development of phage-display technology by George Smith and Greg Winter. Therefore, the work of the Chemistry Laureates can be notionally held to have increased the feasibility, or at least the application, of the work done by this year's Physiology or Medicine Laureates.

2. Speeding Up Protein Evolution

Natural Evolution

A protein is essentially a string of amino acids that folds into a definite shape, known as its native structure. The native structure of every naturally occurring protein is exquisitely adapted to its specific function. Proteins that are enzymes bind to other molecules and catalyse their synthesis or breakdown. Examination of the amino acid sequences and structures of enzymes performing the same function in different organisms reveals that, between species, there are multiple differences in amino acid sequence(s) that are associated with subtle (and sometimes profound) changes in their microstructures, redox characteristics, and catalytic potentials, with enzymes in each organism being most well-adapted to their own characteristics and requirements. There is now much genetic, molecular and biochemical evidence to support the notion that evolutionary changes (mutations) and selection pressure drive the establishment of differences in the sequence(s) and structure(s) of proteins performing identical functions in different organisms. Basically, it is believed that mutations constantly occur at some basal frequency to alter the DNA encoding different enzymes and proteins. If such mutations result in changes in the folding rate, folded structure, or function of the encoded protein, and if the change is beneficial, then the individual in which the mutation has manifested develops an advantage over other members of his/her species.

There is now much genetic, molecular and biochemical evidence to support the notion that evolutionary changes (mutations) and selection pressure drive the establishment of differences in the sequence(s) and structure(s) of proteins performing identical functions in different organisms.

Over many generations of reproduction, such an advantage is passed on to more and more individuals because of the greater

opportunity for survival and passing on of the mutated gene. This is how a mutation in a gene manifests as a mutation in the amino acid sequence of an enzyme, and how the mutation then becomes established in a particular species. Conversely, when a mutation offers a growth disadvantage or a lethal outcome, it is eventually lost from the gene pool of the species without being passed on to succeeding generations. This is how proteins and enzymes are believed to evolve and establish themselves in different organisms.

The Genius of Frances Arnold

While others were still busy analysing protein evolution by studying the evidence of its occurrence in the genomes of organisms, and using the evidence and the analyses to engineer proteins and enzymes in rational ways to test out various hypotheses concerning the relationship between a protein's sequence and its structure, Arnold was impatient and decided to take the 'bull by the horns'. She tried speeding up the two most important aspects of protein evolution in her laboratory, the first being the rate at which sequence variants are generated through mutations, and the second being the rate of selection of the fittest variants generated.

To speed up the process of protein evolution, Frances Arnold took enzyme-encoding genes out of their natural contexts and placed them within bacteria.

The natural rate of generation of variants in a population of organisms is quite slow, with mutations only seen to occur very rarely, once in many generations. To speed this up, Arnold decided to take the enzyme-encoding genes out of their natural contexts and place them within bacteria. This was possible, given the ease with which proteins can be produced in bacteria in recombinant form as a consequence of the revolution in molecular and cellular biology over the previous several decades. She then took these genes and amplified them using a Nobel Prize-winning technique, called the polymerase chain reaction (PCR)[1]. However, Arnold took care to cripple the technique by deliberately causing it to introduce errors during DNA amplification. It may be noted that by then, PCR was already being widely used by molecular biologists to clone, amplify, and play with protein-encoding genes of all kinds. Most scientists bemoaned the low-fidelity of the DNA polymerase enzymes used in the PCR technique initially since

[1] Simarjot Singh Pabla and Sarabjot Singh Pabla, Real-Time Polymerase Chain Reaction – A Revolution in Diagnostics, *Resonance*, Vol.13, No.4, pp.369–377, 2008.

these happened to introduce all kinds of unwanted mutations into the genes that were being cloned. Thus, there was a great focus on searching for enzymes with higher fidelity rates of copying DNA.

Arnold decided to take the opposite direction and explore instead the reasons for low fidelity of DNA-copying enzymes, and apply/exploit these factors in order to deliberately introduce mutations, where there were none to begin with. She soon discovered the reaction conditions under which low-fidelity DNA polymerases introduced genetic mutations at higher rates than normal, and Arnold developed this discovery into a method called 'error-prone PCR'. Using this method, she introduced all kinds of mutations in enzyme-coding genes of her interest. All the mutations were essentially introduced randomly rather than through specific design, exactly as is believed to occur during natural evolution (except that mutations were introduced in each cycle of the error-prone PCR reaction at much higher rates than what occurs during natural evolution). Arnold thus managed to rapidly create libraries of sequence variants of the same gene within the same reaction, covering a large part of the diversity that nature would produce with a much lower frequency, and over a larger group of organisms, over evolutionary time (hundreds of millions of years). This was the first aspect of evolution that she mimicked in the lab, essentially by tweaking the PCR reaction to be more (rather than less) prone to making errors while copying DNA.

The first aspect of evolution that Frances Arnold mimicked in the lab was based on tweaking the PCR reaction to be more (rather than less) prone to make errors while copying DNA.

The second aspect of evolution that she manipulated is the process by which there is a selection of the fittest variant. Here, of course, one is dealing with the selection of the enzyme which is the fittest in the context of a particular chemical reaction carried out in the test tube, or on a multi-well test plate, rather than the fittest enzyme in the context of a living cell, or a whole organism, which is a much more complex process. Arnold developed rapid-assaying formats and protocols for quickly estimating which mutations had the desired effect of being more catalytically-efficient, using colour-generating enzymatic reactions that could be examined visually, or spectrophotometrically on test plates. Each well in the test plates was populated by the growth and division of

a single progenitor bacterium producing a particular amino acid sequence variant of an enzyme – either secreting it out into the medium or being lysed to release it into the medium. In a single readout, it became possible for her to size-up and assess the mutated enzymes and compare them to the wild-type (unmodified) enzyme. She discarded the inactive enzymes and the ones with reduced catalytic efficiency and subjected the rest to repeated cycles of error-prone PCR and test-plate-based testing of bacteria expressing the variants thus generated.

Using the twin approaches of generating sequence diversity rapidly, and screening for the fittest enzymes using test plates, it was possible to develop many robust enzymes that were exquisitely well-adapted to functioning in particular environments mimicking the chemical industry.

Using these twin approaches of generating sequence diversity rapidly, and screening for the fittest enzymes using test plates, Arnold developed many robust enzymes that were exquisitely well-adapted to functioning in particular environments mimicking the chemical industry (e.g., in the presence of organic solvents, or at a particular temperature, or in the presence of some acids or alkalis). Her Nobel Prize is thus a reward for her development and demonstration of techniques allowing the adaptation of presumed evolutionary principles and showing how they can be used to obtain enzymes with desired characteristics in real-time, by speeding up their evolution.

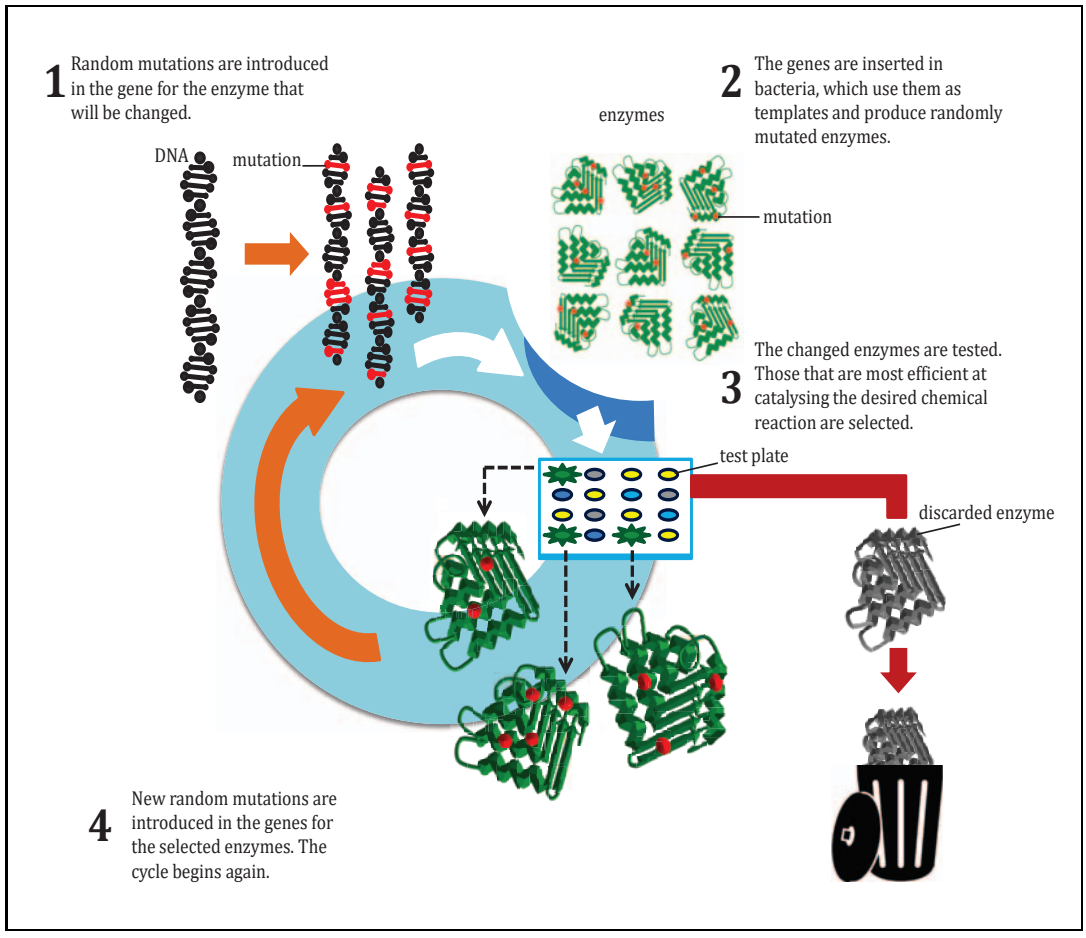
The entire concept is summarized in the schematic (*Figure* 1) which is adapted and re-drawn from the illustrations released by www.nobelprize.org for non-commercial usage.

3. Phage Display Technology

Protein Display on Phages

So, what exactly is phage-display? Well, first of all, the word 'phage'[2] here generally refers to a virus that infects bacteria, and specifically to something called a filamentous bacteriophage. Filamentous bacteriophages, such as M13 or fD are long filament-shaped viruses made up of rolled-up, single-stranded, circular DNA and a few proteins that coat this DNA (known as coat proteins). The phages are capable of interacting with bacteria through a structure on the bacterial surface known as the 'pilus'. Fila-

[2] For a general introduction to bacteriophages, please see: K L Sebastian, Mechanochemistry – The Amazing Viral DNA Packaging Molecular Motor, *Resonance*, Vol.12, No.5, pp.48–59, 2007.



mentous bacteriophages manage to insert their genomic contents (DNA) into bacteria such as *Escherichia coli* essentially by interacting with this pilus, in a process that is called 'infection'. Once an individual bacteriophage has infected a bacterium, it no longer supports infection by another bacteriophage. The bacteriophage-derived DNA which has been inserted into the bacterial cell then uses that cell's own machinery to make multiple copies of the phage genome and phage coat proteins, passing through a stage of double-stranded existence very much like the plasmid DNA that genetic engineers work with (known as a phagemid).

The result of this virtual 'take-over' of the bacterium through a

Figure 1. Enzyme Evolution. A schematic explaining the key aspects of the techniques of error-prone PCR to generate random mutations, and rapid multi-well plate-based testing to screen for the fittest enzymes which are most efficient, and subjecting them to more cycles of mutations.

bacteriophage infection is that the infected cell continues to create and secrete, a virtually endless series of copies of the bacteriophage. Since each bacteriophage is secreted by the originally infected *E. coli* cell into the environment, the cell survives without being killed[3]. While the secretion of filamentous bacteriophages slows down the ordinary metabolism of the infected cell, it allows the infected cell to continue to grow and divide slowly, thus creating more bacteria containing the phage genome, which continues to create and secrete even more bacteriophages *ad infinitum*, or at least until the nutritional resources supporting the bacterial growth last. The process also allows the secreted bacteriophages to infect other uninfected bacterial cells in the vicinity.

[3]Note that in this respect, filamentous bacteriophages differ from lytic phages which destroy the cell within which they replicate.

Thus, when one starts off the process of infection, using a few bacteriophages and some uninfected *E. coli* cells, the cells continue to multiply through growth and division into a population of slowly-growing bacteria that are infected to levels of saturation, and which continuously produce and secrete copies of the infecting phage.

4. The Genius of George P Smith

Peptide Phage Display

George P Smith realized that he could genetically manipulate the amino-terminal end of one of the phage coat proteins, which is produced by the third gene (*gene III*) in the phage repertory of roughly ten genes.

George Smith's genius lay in several things. First of all, he realized that he could genetically manipulate the amino-terminal end of one of the phage coat proteins, which is produced by the third gene (*gene III*) in the phage repertory of roughly ten genes. Isolating the phage genome as a double-stranded phagemid (akin to a plasmid), Smith fused a piece of DNA encoding a 'foreign' peptide sequence to *gene III*, such that the phage genome would now produce a new and modified version of the encoded protein (gene III protein, or gIIIp) as well as display this modified gIIIp (bearing a new peptide sequence attached to its amino-terminus) upon its own surface. There happen to be about five copies of gIIIp at the tip of each phage, and the tip of each phage is involved in the process of infection, with gIIIp playing a role in the infection by interacting with the bacterium's pilus. One of the

things that George Smith showed is that the insertion of an additional sequence of a few (say seven to ten) amino acids at the amino-terminus of gIIIp, in all five copies of the protein at the tip of the phage, does not affect its ability to infect bacteria. This fact allowed Smith to create legions of bacteriophages, all displaying the same (desired) peptide sequence at its tip, by growing the phages in a bacterial culture.

Combinatorial Peptide Phage Display Libraries

George Smith also made a second fascinating discovery. He realized that if he were to replace the DNA encoding a single peptide sequence with a combinatorially-diverse library of DNA encoding many different amino acid sequences, he could theoretically create, grow, and propagate a large library of bacteriophage genomes. Each individual bacteriophage produced through the propagation of such a library of phage genomes within *E. coli* would then be characterized by a single genome incorporating a single DNA sequence variant which would be unique to that bacteriophage. This DNA sequence variant would thus specify the inclusion of a certain specific amino acid sequence variant peptide at its tip (attached to its own gIIIp protein) appearing on the tip of any copy of that phage produced by that particular genome, after its infection of any bacterium, by the cellular machinery of that particular infected bacterium. Since there is virtually no likelihood that two phages could ever infect the same bacterium, the whole process of amplification of the phage (i.e. its growth and multiplication in a culture of growing and multiplying bacteria) would offer an iron-clad guarantee that there would be no 'mixing-up' of different DNA (genomes) and proteins (gIIIp) during the creation of bacteriophages. This would ensure that any single bacterial cell would produce and secrete only a single phage variant. This would also ensure that each such phage would display (at its own tip) the very same amino acid sequence which is encoded by the modified DNA present within its own genome, lying in fusion with its own gene III-encoding protein.

George Smith also realized that if one was to replace the DNA encoding a single peptide sequence with a combinatorially-diverse library of DNA encoding many different amino acid sequences, one could theoretically create, grow, and propagate a large library of bacteriophage genomes.

George Smith realized that this would work even if a billion

different phage variants were all simultaneously grown and propagated within the same flask supporting an *E. Coli* culture infected by a phage-displayed library of billions of bacteriophages, displaying a billion different variant amino acid sequences at the tip of gIIIp, each present initially in numbers of a few tens or hundreds of copies, but growing to many times that number during amplification through growth in the culture.

Screening of Phage Display Libraries

The third great thing that Smith realized and implemented was that the phage-displayed library offers great scope for competitive assays for binding. Here is how this works. Imagine that you have a library of approximately a billion different variant peptides which are seven amino acids long, all displayed at the tips of bacteriophages, with each variant present in a few tens to a few hundreds of copies[4]. Now imagine that you have a test-plate surface functionalized with several million copies of a particular protein, X. You are trying to find a peptide that can bind to this protein, X. The process involves multiple cycles of bio-panning and amplification of bacteriophages.

[4] At each position in a seven amino acid residues-long peptide, one can have any one of the twenty different naturally-occurring amino acids. Therefore, with twenty amino acids being possibly located at each of the seven positions, one can create 20^7 or about 1.28 billion peptide variants.

Biopanning: Firstly, we let a suspension of the billion phage variants (each present in multiple copies) slosh-around over the surface containing a few million copies of protein X. Obviously, not all the 1.28 billion seven residues-long peptide variants displayed at the tips of the bacteriophages will have the ability to bind to protein X. Now let us imagine that there are 50 variants out of the 1.28 billion which bind to protein X with different affinities, varying over several orders of magnitude of binding strength, i.e. with the binding of some being thousands of times stronger than those of others. Because there are several million copies of protein X and only a few hundred copies of each of the 50 variants, no more than 5000 copies of protein X will be bound by any peptide-bearing phage, leaving most copies of protein X unbound. In other words, at this stage, there is no competition. Now, we extract these 5000 phages from the surface. What we have done is to start with a billion peptide variants and identify 50 variants that

can bind to the ligand of our interest (protein X).

Amplification and Repeated Biopanning: Now we take the extracted bacteriophages (5000 (5×10^3), or fewer), collectively representing 50 peptide variant sequences, each present in multiple copies. We add these to a growing *E. coli* culture, allowing them to infect the bacterial cells and amplify in numbers. Since each millilitre of the culture can accommodate up 10^{13} phage particles in a stationary-phase bacterial culture, by the end of this culture's growth one has effectively raised what was once a population of 5×10^3 bacteriophages to 1×10^{13} bacteriophages. This means that each of the 50 variants is now present in ten billion copies each, instead of only ten to a hundred copies each. This offers scope for immense competition. If this population is once again allowed to slosh-over an identical surface of a few million copies of protein X, only the very strongest binder will bind to all the copies of protein X, elbowing-out the next-best binder if the two bind with very different strengths. If they bind with very similar strengths, they will end up occupying binding sites on protein X in proportion to their binding strengths. Thus, when one extracts phages from this second round of biopanning, one obtains several million bacteriophages representing only a few peptide sequence variants (1–4, but usually no more).

Identification of Binding Peptide Variants:

One can individually isolate and sequence the DNA of the phages identified from the above step to immediately decipher exactly which DNA sequence (and, therefore, which peptide variants) constitute the best binding variants from amongst over a billion variants. This is really the most interesting aspect of phage-display. A phage with a particular binding quality (owing to its displayed variant peptide/protein) can be easily isolated, grown and sequenced. This helps in the identification of the genetic sequence encoding the variant and, therefore, its chemical structure. This physical link between the 'genotype' and the 'phenotype' is what makes phage display so interesting and useful.

One can individually isolate and sequence the DNA of the phages identified from the above step to immediately decipher exactly which DNA sequence (and, therefore, which peptide variants) constitute the best binding variants from amongst over a billion variants.

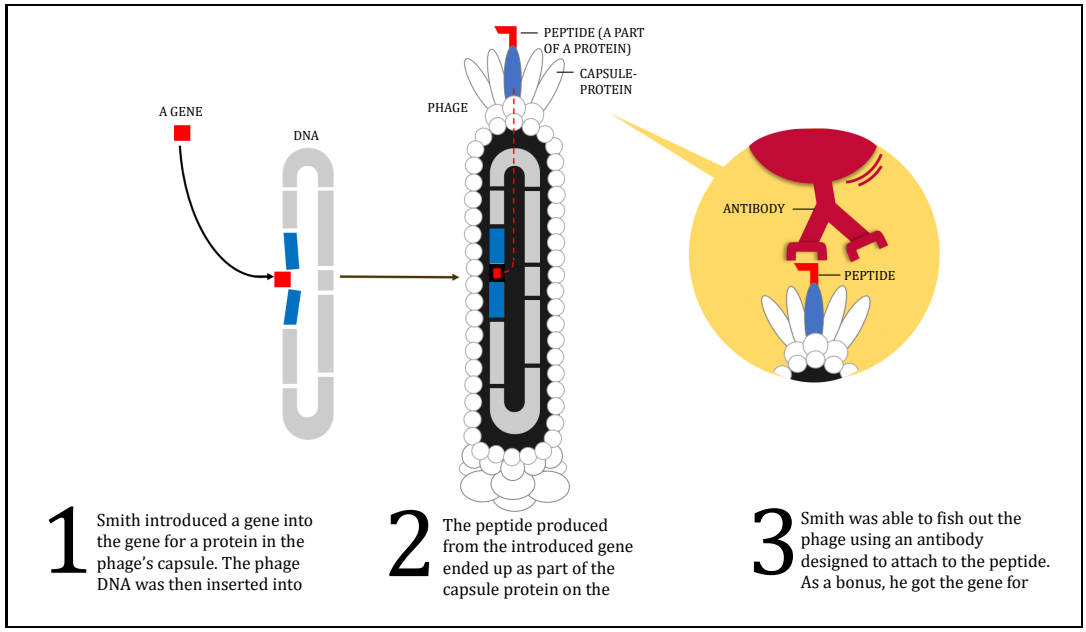


Figure 2. A schematic explaining the key aspects of the technique of phage display.

Why is This Remarkable?

Imagine that you were to try the same process with a heterogeneous library of billions of chemicals, instead of phage-displayed peptide sequence variants. In the first round of biopanning, you would recover the chemicals that engage in binding. However, they would be present in trace amounts. You would not be able to increase their numbers without identifying them *a priori* and resynthesizing them because unlike peptides their quantity cannot be amplified using a genetic system and propagation within microorganisms. Therefore, you would be hard put to either identify the binding variants or create the kind of white-heat competition for binding that becomes critical in identifying the best binders.

The entire concept is summarized in the schematic (*Figure* 2) which is adapted and re-drawn from the illustrations released by www.nobelprize.org for non-commercial usage.

5. Phage Display and the Genius of Gregory P Winter

Antibody scFv Libraries

Now, one has to understand that to incorporate small peptides in fusion with gIIIp is one thing, but to put large proteins in fusion with gIIIp is quite another. A large protein physically occludes the domains of gIIIp. This interferes with the process of bacterial infection because there is a steric hindrance in the interaction of the phage with the bacterium, when the domains of gIIIp are occluded in all five copies of the protein at the tip of the bacteriophage.

In the discussion above, we have seen how critical it is for the phage displaying some peptide or protein variant, and present in a phage-display library, to retain its infective capability. This is because phage(s) isolated from one round of biopanning must be amplified through infection of *E. coli* before the second round of biopanning can be undertaken with the amplified copies of the same phage(s).

Gregory Winter had created a DNA library of about a billion different gene sequences encoding the variable heavy (VH) and variable light (VL) domains of human antibodies[5]. The sequences had been naturally produced through a process called 'V-D-J recombination' which occurs within immune cells. Winter had already cloned the entire repertoire of V–D–J recombined VH and VL domains into a genetic library using the PCR reaction. He wanted to fuse the VH and VL domains further, through combinatorial approaches, and then mount these VH–VL fusion variants on bacteriophages, using George Smith's phage display approach. His purpose was to create a library of human-derived antibody-like proteins that could be screened in the laboratory for their ability to bind to specific ligands or antigens.

[5] Desirazu N Rao and Bharath Wootla, Catalytic Antibodies: Concept and Promise, *Resonance*, Vol.12, No.11, pp.6–21, 2007.

Obviously, the fruits of creating such a library were many. The very existence of such a library would make it potentially unnecessary for experimenters to follow the much longer normal protocol for generating antibodies, involving injecting animals with

protein antigen(s), isolation of immune cells, generation of hybridomas, isolation of monoclonal antibodies, and testing of such antibodies for binding efficacy and/or neutralization efficacy.

Winter had already verified that a VH-VL fusion constitutes a single chain Fv region of an antibody (or an scFv). Of course, scFvs are not whole antibodies as they lack the constant or Fc region present in an antibody, and unlike an ordinary antibody, they are monovalent[6] and not divalent. Still, Winter recognized that scFv antibodies could be very useful. He understood that an scFv antibody library could be used to fish out the V-D-J recombined sequences of the VH and VL domains, with all the work of recombining different CDR (complementarity determining region) sequences in each VH or VL domain having already occurred within the cells of a human system, once and for all, before the creation of such a library. Therefore, once an scFv was identified, and its encoding gene sequenced, it would become a relatively simple task to use this information to generate a completely human antibody for therapy, or for diagnostic applications of various kinds.

[6] Possessing a single binding site.

Gregory Winter aimed to create a library of human-derived antibody-like proteins that could be screened in the laboratory for their ability to bind to specific ligands or antigens.

Winter very much wanted to mount these scFv antibodies on phage, but he realized that the scFv antibodies would not infect bacteria, because of the steric occlusion affecting gIIIp. So, he identified and implemented a new approach. He truncated the gIIIp gene and removed one of its domains, to make space to incorporate the scFv, and ensure that the total size of the modified, variant gIIIp would remain the same. He also created a bacterial host system that would produce a wild-type (unmodified) gIIIp alongside the modified, variant gIIIp from the incoming (infecting) phage, in which one domain of the gIIIp had been removed to insert the scFv variant(s). The resultant phage (called 'hybrid phage', as opposed to Smith's phage which was called 'recombinant phage') would thus have some copies each of both normal, unmodified gIIIp molecules and modified, variant gIIIp molecules at the tip, amongst the five copies of the gIIIp present.

The idea was that the normal gIIIp molecules incorporated randomly during phage assembly would help the scFv-displaying

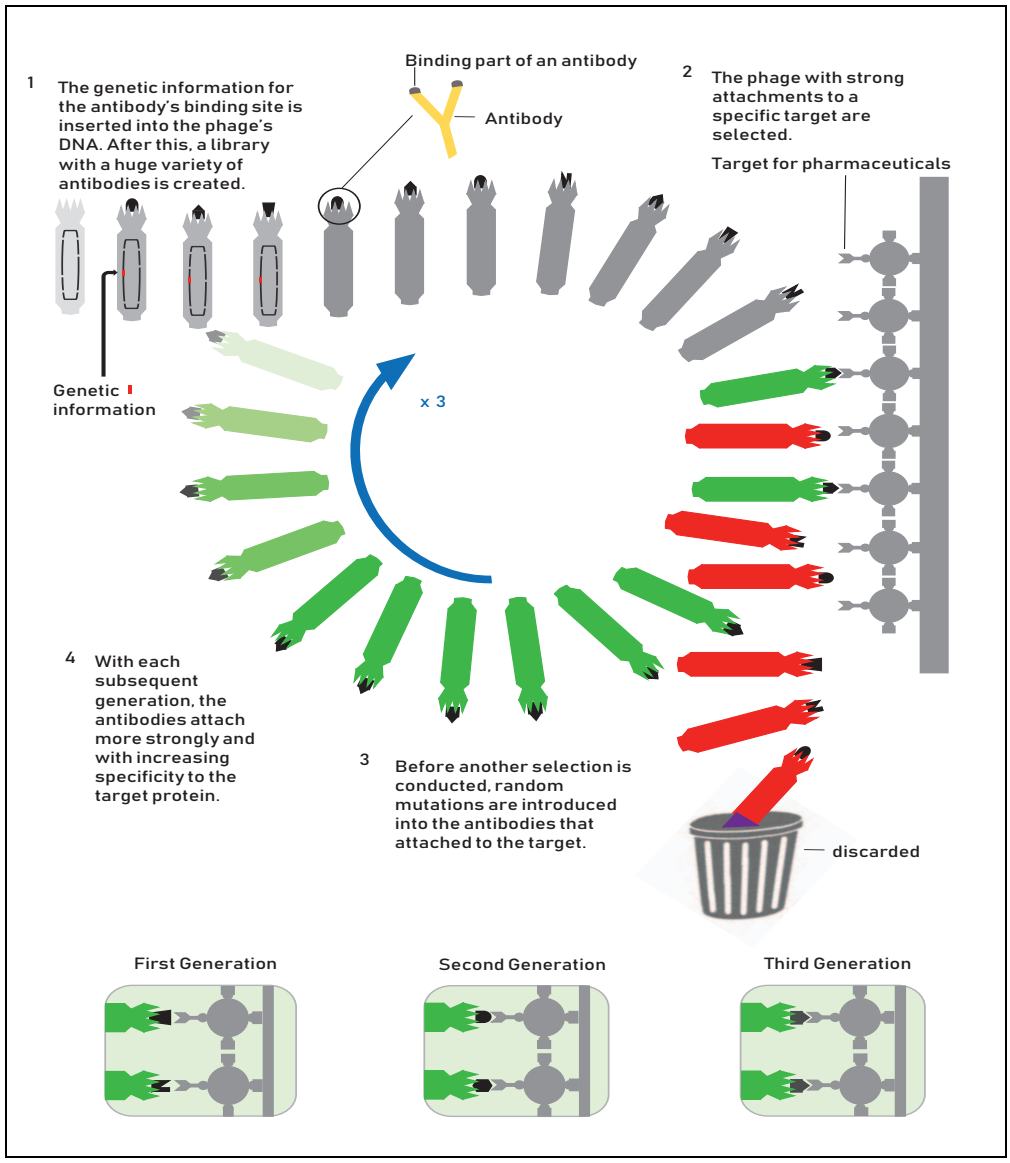


Figure 3. A schematic explaining the key aspects of the use of the technique of phage-display for screening of a phage-display antibody library.

phage to infect bacteria, while the modified, variant gIIIp molecules bearing the scFc would help the phage participate in biopanning. Any phage incorporating only the normal gIIIp coat proteins (all five copies) would be able to infect bacteria, but fail to participate in biopanning. Any phage incorporating only modified gIIIp coat proteins would participate in biopanning, but fail to infect bacte-

ria or become amplified. Thus these would be lost from the gene pool of the portion of the phage library used for any experiment. The hybrid phages bearing some normal and some modified gIIIp would go through the cycles of biopanning and amplification.

The concept is summarized in the schematic (*Figure* 3) which has been adapted and re-drawn from the illustrations released by www.nobelprize.org for non-commercial usage.

This brilliant tactic allowed Winter's group to mount the entire population of one billion scFv antibodies on phage, and demonstrate that it is possible to screen this library for scFv variants for binding to any ligand.

The rest, as is commonly said, is history. The use of phage-display technology for proteins (rather than just for small peptides) and its application in the generation of libraries of scFv (and now many other types of) antibodies has led to it becoming a tool that now lies at the heart of all combinatorial approaches to protein engineering.

Address for Correspondence
Purnananda Guptasarma
Centre for Protein Science, Design and Engineering (CPSDE)
Department of Biological Sciences
Indian Institute of Science Education and Research (IISER) Mohali
Sector-81 SAS Nagar
Punjab 140 306, India.
Email:
guptasarma@iisermohali.ac.in

Acknowledgements

Purnananda Guptasarma thanks his doctoral students, Arpita Mrigwani and Arpita Sarkar, and his masters' thesis student, Arun Chikkara, for adapting and re-drawing the illustrations available on www.nobelprize.org for better reproducibility.

Suggested Reading

[1] L Pritchard *et.al*., A General Model of Error-prone PCR, *J Theor Biol*., 234(4), pp.497–509, 21 June 2005.

[2] Justine Bazan, Ireneusz Calkosinski, and Andrzej Gamian, Phage Display – A Powerful Technique for Immunotherapy, *Hum Vaccin Immunother*., 8(12), pp.1817–1828, 1 Dec 2012.