Directed evolution – bringing the power of evolution to the laboratory: 2018 Nobel Prize in Chemistry

The 2018 Nobel Prize in Chemistry has been awarded to Frances H. Arnold, George P. Smith and Gregory P. Winter for harnessing the power of Darwinian evolution into the laboratory. As rightly mentioned by Claes Gustafsson, Chair of the Nobel Committee for Chemistry during the award announcement, 'Our laureates have applied the principles of Charles Darwin in the test tubes, and used this approach to develop new types of chemicals for the greatest benefit of humankind'. This year's Prize is dedicated to methods that create new molecules using the biosynthetic machinery of nature: directed evolution for making new enzymes and phage display for therapeutic antibodies.

Arnold was recognized with half the share of the Prize for demonstrating the directed enzyme evolution to produce new enzymes, while Smith and Winter shared the remaining half of the Prize for developing a technology called phage display that utilizes microorganisms to produce specific peptides, proteins and antibodies. Both the technologies, i.e. 'enzyme evolution' and 'phage display' used the same principles – genetic change utilizing the power of molecular biology techniques followed by selection using functional assays to develop novel enzymes, proteins and antibodies. Enzymes produced by directed evolution are used in everything ranging from pharmaceuticals, eco-friendly detergents, taste enhancers to renewable fuels. Smith and Winter used a slightly different approach to evolve proteins and antibodies that can combat a large number of diseases.

Directed evolution of enzymes

Frances Arnold designed a way to direct evolution of proteins and enzymes inspired by the biological processes that drive natural selection. Her methods revolutionized the field of protein engineering and are now the widely used strategies in research as well as in the chemical and pharmaceutical industries. Arnold earned her Ph D in chemical engineering from the University of California (UC), Berkeley, USA in 1985. After completing postdoctoral research from UC Berkeley and the California Institute of Technology (Caltech), USA, she joined Caltech as faculty in the year 1987. As Arnold says, microbes are brilliant chemists sitting in our body. Her idea was if we could reprogramme these microbes to make new chemicals, they can serve as chemical factories of the future. Microbial factories can be engineered to convert the natural and waste material into chemicals and fuels that are required for our survival on this planet.

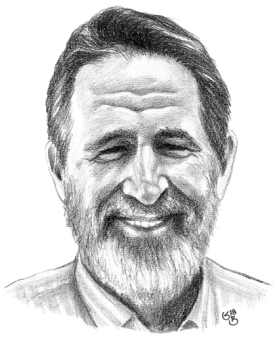
The idea of evolution had captivated Arnold to study the biological system and more specifically the field of protein engineering. Being from an engineering background, she has looked at all the biological problems from an engineering perspective. She engaged herself in the quest of engineering of enzymes by directing their evolution *in vitro*. The ability of enzymes to carry out complex chemical reactions is admired by all the chemists and biochemists. But the enzymes provided by nature are not adequate for all our requirements; therefore manoeuvering these complex structures to acquire novel properties was a logical option. Arnold utilized the capabilities of DNA manipulation technology using evolution as an information tool to create and innovate enzymes in the laboratory.

In nature, evolution has given rise to new proteins through mutations, recombination and selection to carry out well-suited biological tasks. There are many examples of enzyme evolution which led to new whole families of proteins sharing a common ancestral protein of the same structure. With the advent of technology at that time, it was possible to manipulate the biological macromolecules. The idea was to fine-tune the enzymes to make them functional under the desired conditions, rather than the situation in which they evolve. Directed evolution involves accumulation of valuable mutations during iterations of mutagenesis and then selection of useful variants.

The critical question was how to find the correct state of the enzyme that exhibits the desired property. A typical enzyme is made up of 200–300 amino acids and there can be 20 possible amino acids at every single position; hence astronomical combinations would be possible. Random sampling was not ideal and unlikely to lead to the desired protein. The alternative approach was to limit the search space in the protein by looking at how a particular protein evolved in nature, rather than searching a random space. Arnold integrated her knowledge of biochemistry to restrict the possibilities, focusing on introducing mutations that are expected to have a positive effect on activity and avoiding regions in which mutations would likely lead to detrimental effect on protein stability and folding.

Arnold and her colleagues used an enzyme subtilisin E, which was desirable for working with polar organic solvents like dimethylformamide (DMF). The enzyme was highly stable in such solvents, but its catalytic activity was very low[1]. She began with a question whether the enzyme could be manipulated to perform well even in DMF. She started by preparing a library of subtilisin mutants with random amino acid mutations and then screen the mutant library for progeny functionally better than the parent. Random mutagenesis involves sequential

Frances H. Arnold



George P. Smith

Gregory P. Winter

random mutations in the enzyme and screening of mutant library for identification of better function. The powerful polymerase chain reaction (PCR) technique was utilized to create mutants by introducing random base substitutions in the gene under sub-optimal conditions. Mutants obtained were then transformed into bacteria and checked for their functions. After each round of mutagenesis, the mutant enzymes were screened for their ability to hydrolyse the milk protein casein. Enzyme secretion was tested by growing the bacteria on agar plates in the presence of casein and DMF. The bacteria would hydrolyse the casein and produce a visible halo if functional. Active enzyme variants that created the biggest halos were selected and their DNA was isolated for further rounds of mutagenesis. Substitution of three amino acids led to significant improvement in enzyme activity compared to wild-type subtilisin. Random mutagenesis combined with site-directed mutagenesis resulted in overall 40-fold more enzyme activity than wild type in 60% DMF. A further six generations of mutagenesis and screening created an enzyme which was 500-fold more active in 60% DMF. The directed evolution of subtilisin E was a benchmark achievement to design enzymes that can be used for a wide variety of applications like renewable fuels and pharmaceutical products. This work became the starting point and led to the improvement and reshaping of enzymes used in chemical and organic synthesis by directed evolution.

Arnold developed different selection criteria in order to optimize enzymes for different functions using directed evolution in combination with other techniques. An alternate approach is recombination, i.e. the cutting and pasting of pieces of genes to accelerate molecular evolution. Recombination is an effective method for generating novel functional diversity. Arnold recombined the genes from different sources having certain homology with favourable mutations to produce chimeras that still maintained fold and function. Directed evolution of an enzyme para-nitrobenzyl esterase to exhibit a new function using the combined strategy of random mutagenesis followed by recombination was achieved resulting in around 470-fold increase in enzyme activity over wild type[2]. Arnold developed computational algorithms to predict the parts that can be combined without disrupting their original structure and applied directed evolution to further mutate the chimeras to optimize their functions[3]. These chimeric libraries are highly diverse, with a large number of mutations ranging from ten to even hundreds, that fold and are highly active.

Directed evolution has revolutionized the field of enzyme engineering by catalysing a set of reactions for which no specific enzyme was previously available[4]. These include cyclopropanation, nitrene and carbene transfer reactions, arsenate detoxification and production of strained carbocycles. One of the applications of directed evolution was optimization of thermostability of an enzyme. Six rounds of random mutagenesis, recombination via DNA shuffling, and screening were done to considerably stabilize the *Bacillus subtilis p*-nitrobenzyl esterase considerably without affecting its catalytic activity at lower temperatures[5].

The major challenge for humanity in today's world is to find a suitable substitute for non-renewable sources that can be sustainably produced. Among the lead candidates for biofuel is isobutanol and the biosynthetic pathway for isobutanol production involves the use of reduced nicotinamide adenine dinucleotide phosphate (NADPH). However, in the normal metabolism of bacteria, nicotinamide adenine dinucleotide (NADH) is produced. To resolve this, Arnold utilized directed evolution to alter the enzyme specificity of *Escherichia coli* ketol-acid reductoisomerase (KARI) enzyme by constructing NADH-dependent pathway using engineered enzymes, which led to the production of isobutanol in *E. coli* suitable for biofuel production[6]. Directed evolution of enzymes is a highly versatile and efficient technique for the development of novel biocatalysts, optimization of new and known enzymes, biofuel production and eco-friendly synthesis of chemicals. This technology has enormous potential and it has truly revolutionized the world of enzyme engineering.

Phage display of peptides and proteins

George Smith is best known for his pioneering work on phage display technology in 1985 (ref. 7). He utilized the available genetic tools and exploited bacteriophages for linking genotype to phenotype, which was undoubtedly a breakthrough. Smith received his Ph D in bacteriology and immunology from Harvard University, USA in 1970. Following a post-doctoral research stint at the University of Wisconsin, he joined as faculty at the University of Missouri in Columbia, USA in 1975. The idea of evolution greatly influenced Smith in his early career. He considered that replicability and mutability are two essential characteristics of evolving organisms. To test this concept *in vitro*, a genetic machinery was required. So to mimic the idea of *in vitro* evolution of chemicals, he developed the technology of phage display which, in principle, provides for artificial chemical evolution.

Phages are viruses that infect the host bacterium *E. coli.* The main characteristic of phage vectors is that they can incorporate foreign DNA fused along with their own genome without losing infectivity. As the phage replicates in the host *E. coli*, the foreign insert also replicates with it as a messenger. During translation, the foreign protein is packaged with phage endogenous coat protein in such a way that the phage displays it on its surface in immunologically accessible form.

Smith inserted the foreign DNA in the middle of *gene III* of bacteriophage to create a fusion protein between the two domains[7].The displayed peptide used by him for insertion in bacteriophage was a 171 bp fragment of *Eco*RI endonuclease gene. After confirming the presence of inserts using restriction analysis, Smith used the *Eco*RI-specific antibody to select the cognate fusion phage from the broad background of non-cognate phages. *Eco*RI-specific antibody was coated onto polystyrene dishes and a single round of affinity purification was done to carry out the selection. A suspension containing phages specific to foreign gene and an excess of non-specific phages were added to the dish. After extensive washing, phages were eluted, and subsequent titration experiments revealed the enrichment of cognate fusion phages against the background phages. The success of this experiment has opened new avenues, and it would be possible to isolate desired clones from a library of random inserts using any receptor of interest as bait.

In 1988, Parmley and Smith[8] introduced a much-improved procedure for the phage display technology. Cloning of foreign insert was shifted downstream of

the signal peptide so that *gene III* is uninterrupted by the insert. A new affinity purification procedure was introduced using the strong chemistry of biotin–streptavidin bond. The phages screened against the biotinylated antibody were directed against the streptavidin-coated petri plate. Smith introduced the word 'biopanning' for purification of phages displaying the target insert against the background phages and following this approach, a 10^8-fold enrichment of phages was achieved to display a peptide of anti-β-galactosidase antibody.

Following Smith's groundbreaking discovery, a large number of research groups started working on phage display technology to display random inserts on the phage surface. Phage constructs of random peptide length ranging from 6 to 30 mer were constructed and selected against diverse antigens. In 1990, a library of 2×10^7 random 15-residue peptide sequence was constructed and expressed as the fusion peptide on the surface of the phage in the year 1990 (ref. 9). The peptide libraries were constructed using degenerate oligonucleotides, containing a mixture of nucleotides resulting in enormous diversity. This vast diversity of displayed peptides on the surface of phages leads to isolation of specific peptide to almost any type of targets.

Scott and Smith[10] reported a randomized epitope peptide library of approximately 4×10^7 different hexapeptide fragments. The epitope library was then screened against two monoclonal antibodies (mAbs), A2 and M33, which are specific for the hexapeptide DFLEKI epitope of myohemerythrin. They observed that even if the A2 and M33 antibodies were confronted with a milieu of random peptides, mAbs still preferentially selected a peptide similar to DFLEKI epitope.

The phage display peptide libraries have been used widely to select epitopes for antibodies. The technology has been utilized to understand how the immune system generates requisite paratope diversity to be able to recognize any and every epitope that the organism may encounter[11]. A new mechanism for expanding the primary antibody repertoire by differential ligand positioning was reported utilizing random dodecapeptide phage-display library[12,13]. The technology has also facilitated discovery of the molecular mimics[14–17] and potential new drug leads that might act as peptide agonists or antagonists[18–20].

Phage display of antibodies

In 1990, several research groups started to use phage display developed by Smith to make new biomolecules. One of them who adopted and modified this technique was Gregory Winter. He was able to display antibodies on the surface of the bacteriophage and could use phage display in the directed evolution of antibodies. Winter obtained his Ph D working at MRC Laboratory of Molecular Biology, Cambridge, UK in 1977. After completing his postdoctoral research, he became a Group Leader at the same institution in 1981. Winter is best known for his revolutionary work in the field of phage display antibody technology. He was very much inspired by the work of Georges Köhler and César Milstein, who discovered the hybridoma technology which eventually led to a Nobel Prize in 1984. However, the administration of mAbs raised in mice has immunological complications in humans. So to bypass the animal immunization, Winter utilized the phage display technology developed by Smith to make humanized antibodies in the bacteria. As the size of full-length antibodies is quite large, Winter used single-chain Fv fragments (scFv) and displayed them on the surface of filamentous phage.

In 1990, Winter and his colleagues demonstrated the assembly of scFv prepared from heavy- and light-chain domains of anti-hen egg lysozyme antibody (D1.3)[21]. The scFv was cloned into a phage vector fused with gIII protein of filamentous phage. Interestingly, when the binding of phage was checked with lysozyme by ELISA, it showed the same reactivity as against the original D1.3 full-length antibody. The binding was so specific that scFv did not recognize the turkey lysozyme which differs from hen by only seven amino acids.

The first phage display library was prepared by Winter by amplifying variable heavy-chain (V_H) and variable light-chain (V_k) genes from mRNA isolated from mice spleen immunized with a hapten 2-phenyloxazol-5-one (phOx)[22]. The amplified diverse V_H and V_k genes were then linked together to form scFv. The assembled gene was then fused with the fd bacteriophage coat protein *gene III* and electroporated in *E. coli*. The initial library had a size of around 2×10^5 clones. The library was then passed onto the immobilized phOx column and eluted with hapten to obtain specific phages against it. Thirteen per cent of bound phages were found to be specific for phOx with varied binding range and binding could be increased upon successive rounds of selection. In addition, Winter found that the promiscuous reshuffling of V_H and V_L genes led to new combinations resulting in high-affinity antibodies which were not present naturally in the immunized mouse. Using this combinatorial approach, the size of the library could be increased to 4×10^7 and a larger number of high-affinity phOx binders were recovered.

The next challenge was to mimic the strategy of the immune system to create a diverse antibody repertoire by making human antibodies in bacteria. In order to apply phage display technology for making human antibody library, the most important step was the creation of the antibody repertoire of human origin. Antibody diversity was created by randomly combining the variable domain of the heavy and light chains derived from peripheral blood lymphocytes of unimmunized humans[23]. The combinatorial library of around 10^7 scFv was then expressed on the phage surface. The biopanning of human antibody library was done against certain randomly selected antigens to obtain high-affinity antibodies, as the library was derived from unimmunized B cells. With this, Winter and his colleagues established that it is possible to isolate human antibodies against any antigen from a single phage library bypassing the need of hybridoma technology.

Affinity maturation is the process by which antibody gains the affinity for its cognate antigen by undergoing somatic hypermutations and clonal selection in the complementary determining regions (CDRs). *In vitro*, affinity maturation can be obtained by introducing mutation and selection. So to mimic the affinity maturation process of the immune system, Winter and co-workers[24] introduced random mutations in the antibody gene using random mutagenesis by error-prone PCR. In addition to increasing antibody diversity, this has led to dramatic improvements in antibody affinity. One such library, which is available commercially is the Tomlinson I and J

library. This library is based on a single framework region and the diversity is incorporated at specific sites in CDRs that are already known to play an important role in antigen binding. Tomlinson I and J library has been widely used both for academic research and for development of antibody-based therapeutics.

The first approved human therapeutic antibody derived using phage display was adalimumab in 2002 (ref. 25). Humira, the commercial brand for adalimumab, commands a USD 20 billion market. It binds with very high affinity to the TNF-α, a pro-inflammatory cytokine, and is used in the treatment of rheumatoid arthritis, psoriasis and inflammatory bowel diseases. Another therapeutic antibody, belimumab, was discovered using phage display technology and approved in 2011 for treatment of systemic lupus erythematosus. Raxibacumab is yet another human monoclonal antibody produced using the same technology as for treatment and prevention of anthrax. The phage display platform has indeed provided several useful therapeutic candidates for clinical development, and can be used in tailor-made antibodies for the desired characteristic.

The methods developed by the 2018 Nobel laureates in chemistry – Frances Arnold, George Smith and Greg Winter led to the start of a new era in biochemical sciences. The approaches used by all three laureates, in principle, are the same – the directed evolution of enzymes, proteins and antibodies. This technology of directed evolution has limitless advantages and one can use it based on the desired objective to make specific molecules. To conclude, the impact of the technology as stated by the Nobel Committee members is 'we are in the early days of directed evolution's revolution which, in many different ways, is bringing and will bring the greatest benefit to humankind'.

1. Arnold, F. H., *Chem. Eng. Sci.*, 1996, **51**, 5091–5102.
2. Moore, J. C. and Arnold, F. H., *Nature Biotechnol.*, 1996, **14**, 458–467.
3. Meyer, M. M. *et al.*, *Protein Sci.*, 2003, **12**, 1686–1693.
4. Arnold, F. H., *Angew. Chem. Int. Edit.*, 2018, **57**, 4143–4148.
5. Giver, L., Gershenson, A., Freskgard, P.-O. and Arnold, F. H., *Proc. Natl. Acad. Sci. USA*, 1998, **95**, 12809–12813.
6. Bastian, S., Liu, X., Meyerowitz, J. T., Snow, C. D., Chen, M. M. Y. and Arnold, F. H., *Metab. Eng.*, 2011, **13**, 345–352.
7. Smith, G. P., *Science*, 1985, **228**, 1315–1317.
8. Parmley, S. F. and Smith, G. P., *Gene*, 1988, **73**, 305–318.
9. Devlin, J. J., Panganiban, L. C. and Devlin, P. E., *Science*, 1990, **249**, 404–406.
10. Scott, J. K. and Smith, G. P., *Science*, 1990, **249**, 386–390.
11. Kaur, H. and Salunke, D. M., *IUBMB Life*, 2015, **67**, 498–505.
12. Manivel, V., Bayiroglu, F., Siddiqui, Z., Salunke, D. M. and Rao, K. V. S., *J. Immunol.*, 2002, **169**, 888–897.
13. Sethi, D. K., Agarwal, A., Manivel, V., Rao, K. V. S. and Salunke, D. M., *Immunity*, 2006, **24**, 429–438.
14. Kaur, K. J., Khurana, S. and Salunke, D. M., *J. Biol. Chem.*, 1997, **272**, 5539–5543.
15. Jain, D., Kaur, K., Sundaravadivel, B. and Salunke, D. M., *J. Biol. Chem.*, 2000, **275**, 16098–16102.
16. Goel, M. *et al.*, *J. Biol. Chem.*, 2001, **276**, 39277–39281.
17. Goel, M., Krishnan, L., Kaur, S., Kaur, K. J. and Salunke, D. M., *J. Immunol.*, 2004, **173**, 7358–7367.
18. Brown, K. C., *Curr. Pharm. Des.*, 2010, **16**, 1040–1054.
19. Demartis, A. *et al.*, *Sci. Rep.*, 2018, **8**, 585.
20. Lau, J. L. and Dunn, M. K., *Bioorg. Med. Chem.*, 2018, **26**, 2700–2707.
21. McCafferty, J., Griffiths, A. D., Winter, G. and Chiswell, D. J., *Nature*, 1990, **348**, 552–554.
22. Clackson, T., Hoogenboom, H. R., Griffiths, A. D. and Winter, G., *Nature*, 1991, **352**, 624–628.
23. Marks, J. D., Hoogenboom, H. R., Bonnert, T. P., McCafferty, J., Griffiths, A. D. and Winter, G., *J. Mol. Biol.*, 1991, **222**, 581–597.
24. Hawkins, R. E., Russell, S. J. and Winter, G., *J. Mol. Biol.*, 1992, **226**, 889–896.
25. Nixon, A. E., Sexton, D. J. and Ladner, R. C., *mAbs*, 2014, **6**, 73–85.

Sheenam Verma, Regional Centre for Biotechnology, NCR Biotech Science Cluster, 3rd Milestone, Faridabad 121 001, India; **Dinakar M. Salunke*,** International Centre for Genetic Engineering and Biotechnology, Aruna Asaf Ali Marg, New Delhi 110 067, India.
*e-mail: dinakar.salunke55@gmail.com