KUNGL. VETENSKAPS-AKADEMIEN
THE ROYAL SWEDISH ACADEMY OF SCIENCES

3 OCTOBER 2018

Scientific Background on the Nobel Prize in Chemistry 2018

DIRECTED EVOLUTION OF ENZYMES AND BINDING PROTEINS

BOX 50005 (LILLA FRESCATIVÄGEN 4 A), SE-104 05 STOCKHOLM, SWEDEN
TEL +46 8 673 95 00, KVA@KVA.SE ⋆ WWW.KVA.SE

The Royal Swedish Academy of Sciences has decided to award the Nobel Prize for Chemistry 2018 with one half to **Frances H Arnold** “for the directed evolution of enzymes”, and the other half jointly to **George P Smith** and **Sir Gregory P Winter** "for the phage display of peptides and antibodies”.

Introduction

Natural evolution of enzymes has existed since the emergence of life on Earth. Genes have mutated and proteins have evolved to improve the fitness of an organism to tackle conditions in new environments. For thousands of years, humans have been breeding animals and plants through the selection of organisms with desired properties. For most of this time without even knowing they were doing it, humans evolved and optimised enzymes and binding proteins over many generations.

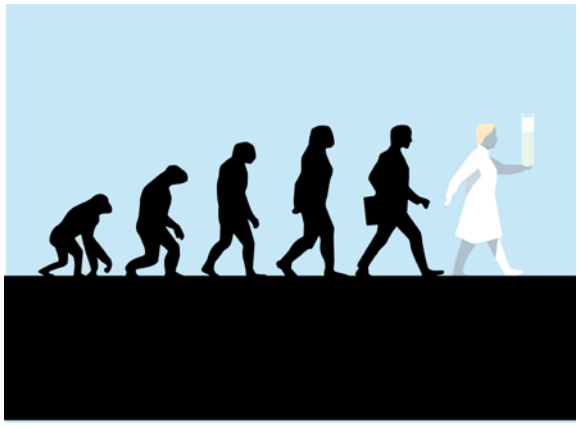

Figure 1. Evolution towards directed evolution

Directed evolution of enzymes and binding proteins is a manmade procedure built on molecular insights, which moves the evolution process into the laboratory and speeds it up. The procedure relies on intended variation of protein sequences at a defined level of randomness. This is coupled to engineered screening and selection strategies. Directed evolution is an iterative procedure which involves the identification of a starting state protein, diversification of its gene, an expression and screening strategy, re-diversification, re-screening, and so on until a satisfactory performance level in terms of enzymatic activity, binding affinity or specificity is reached.

Directed evolution of enzymes and binding proteins has become a widely used strategy in academic research as well as in the chemical and pharmaceutical industries. Directed evolution of enzymes tailors them to operate in new reaction conditions, optimises their catalytic activity towards new substrates, and makes them catalyse new chemical reactions. Directed evolution of enzymes has widely expanded the repertoire of useful biocatalysts. The evolved enzymes offer efficient and environmentally-friendly alternatives to metals and organic catalysts in chemical and biotechnical industries.

Directed evolution of binding proteins is an efficient way to identify variants with high affinity and selectivity for a given target, and to map the sequence requirements for high-affinity and high selectivity protein interactions. Directed evolution of human antibodies leads to useful therapeutics.

Directed evolution of enzymes and binding proteins – in theory

In 1984, Manfred Eigen published a theoretical paper outlining a possible work flow for directed evolution of enzymes (Figure 2, ref 1). Eigen noted that such optimisation becomes an interesting challenge because the genotype and the phenotype are dependent on different molecules. He reasoned that finding rare improved variants in large libraries would be hard if not impossible. Instead he proposed the use of smaller libraries and several generations of mutagenesis and screening as a procedure that would more likely lead forward. Eigen predicted that it would be possible to construct a stepwise iterative "evolutionary machine" to produce optimised enzymes.

Another theoretical prediction is found in a patent (2) that describes evolution of binding proteins through iterative diversification of libraries between selection rounds. Directed evolution of enzymes was also briefly introduced as a possibility in an abstract of a more classical protein-engineering study of enzyme optimisation (3).

Controlled optimization by the above procedure really becomes an interesting challenge when genotype and phenotype are different molecules.

Let us therefore expand the procedure as follows -

10 PRODUCE A MUTANT SPECTRUM OF SELF-REPRODUCING TEMPLATES
20 SEPARATE AND CLONE INDIVIDUAL MUTANTS
30 AMPLIFY CLONES
40 EXPRESS CLONES
50 TEST FOR OPTIMAL PHENOTYPES
60 IDENTIFY OPTIMAL GENOTYPES
70 RETURN TO 10 WITH A SAMPLE OF OPTIMAL GENOTYPES

Figure 2. Extract from Eigen's theoretical proposal of directed evolution (1).

Directed evolution of enzymes – in practise

One decade after Eigen's theoretical work (1), the first experimental work appeared that described successful implementation of directed evolution of enzymes in a laboratory setting to improve enzyme function and versatility (4). **Frances H Arnold** reported the directed evolution of subtilisin E to obtain an enzyme variant which was active in a highly unnatural (denaturing) environment, i.e. at high concentrations of the polar organic solvent dimethylformamide (DMF). After four sequential rounds of mutagenesis and screening in the presence of DMF, an enzyme variant with 256-fold higher activity than the wild-type enzyme in 60% (v/v) DMF was created (4,5).

In the seminal paper (4), Arnold had mastered the whole work flow for directed evolution of enzymes, a methodology relying on several parts: 1) identification of a suitable starting enzyme for the chosen task, 2) DNA-sequence library construction to cover well-chosen subsets of

sequence space, 3) identification of selection criteria that will lead to enhanced or new functions and methods for selection of optimised enzyme variants, 4) re-diversification of the genes to create new DNA-sequence libraries around the sequences from the first selection to cover new subsets of sequence space, 5) setup of selection criteria with increased stringency, and so on for as many rounds as needed to reach the target level of enzyme performance. Each of these five steps has since been further developed and optimised over the years in the Arnold lab and several other labs.

In addition to a first set of four combined single mutations, the first work (4) used error-prone PCR to create and re-diversify DNA-sequence libraries through three rounds of random mutagenesis and screening to evolve subtilisin E. The selection criterion was hydrolysis of the milk protein casein. Active enzyme variants created visible halos on agar plates with casein. Enzymes secreted by bacterial colonies were thus transferred to agar plates containing both DMF and casein, to enable identification of the most active enzyme variants in the presence of the organic solvent. Plasmid DNA was isolated from clones secreting an enzyme variant that produced a halo larger than those surrounding the parent enzyme, and subjected to further rounds of mutagenesis.

The directed evolution of subtilisin E to improve its activity in a polar organic solvent was a benchmark achievement that opened the field of directed evolution of enzymes. This work became the starting point for continued technical development of the methodology for directed evolution. The field expanded towards improving and reshaping enzymes for numerous chemical reactions, old and new, leading to applications of importance for research in organic synthesis, as well as for the chemical and pharmaceutical industry and beyond.

Molecular insights guide library design

It is not possible to randomise every position in an enzyme, the typical size of which is 200-300 amino acid residues or more. Indeed, only a small fraction of the amino acid positions can be varied if the aim is a library with full sequence coverage. The reason is simple combinatorial mathematics and the quickly growing number of variants relative to the number of clones that can be handled in any laboratory setting, or even using the joint capacity of all laboratories in the world. Still, a wealth of studies makes it clear that mutations in and nearby the active site as well as more distant substitutions on the enzyme surface may contribute to optimised catalytic activity. Arnold and co-workers have shown by many examples that library design must be based on molecular insight and knowledge-based choices of which amino-acid positions to vary, combined with some element of added randomness, e.g. through error-prone PCR.

A prominent early contributor to the development and implementation of methodology for directed evolution was the late William (Pim) Stemmer (†2013). Stemmer introduced a DNA recombination strategy termed “DNA shuffling” to the evolution of enzymes. This was an efficient way to propagate beneficial mutations while increasing the size of a DNA library

through random fragmentation and re-assembly of genes (6,7). He showed that the use of DNA shuffling, i.e. recombination of DNA from similar genes from several organisms, introduces more variation than many other methods and can thus improve the chances to reach a substantial activity increase in the evolved variants. In a proof-of-principle study Stemmer and co-workers set out to increase the activity of the enzyme β-lactamase (an enzyme responsible for antibiotic resistance); three cycles of DNA shuffling and screening on plates with successively higher concentration of the antibiotic cefotaxime led to evolution of an enzyme with significantly increased activity (6,7). Gene shuffling had been reported as a means to increase variation and improve antibody affinity for a target, sometimes called affinity maturation, and provided early examples of directed evolution of binding proteins (8-10). In these cases the shuffled gene segments corresponded to light chains (10) or to the variable loops of heavy and light chains of immunoglobulins (8,9).

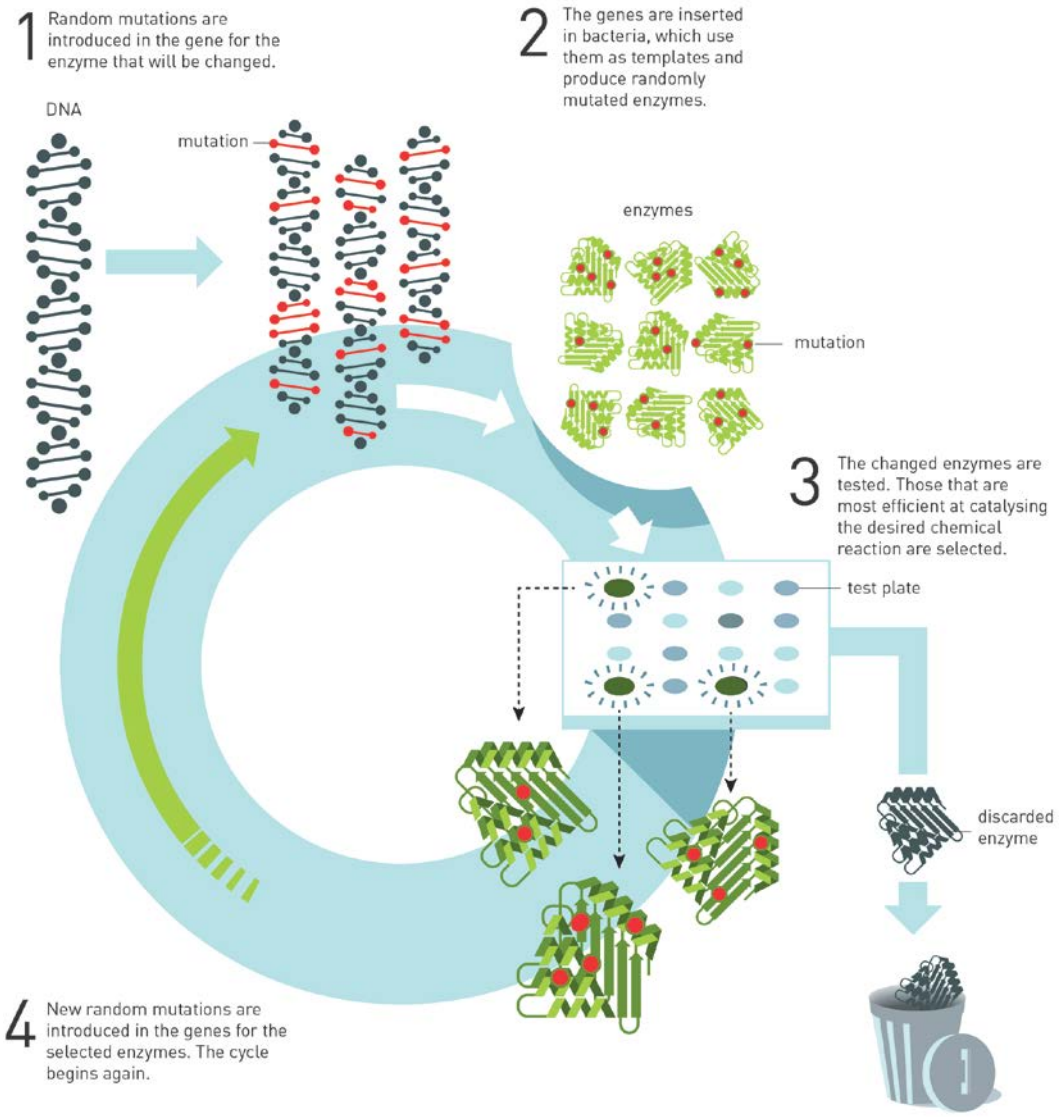

Figure 3. The work flow for the directed evolution of enzymes.

DNA shuffling, staggered extension process (StEP) and other methods for library generation were further developed in Arnold's and Stemmer's laboratories for use in the directed evolution of enzymes during the second half of the 1990s (11-26). Since the end of last century, the progressively decreasing costs for *de novo* synthesis of genes with degenerate codons, or fully designed DNA libraries, have opened a new path towards efficient and affordable production of sequence libraries with tailored diversity.

Selection criteria and screening techniques

The selection criteria and screening techniques must be adapted for each enzyme optimisation endeavour. Selection may be coupled to a cellular survival function; for example, the desired enzymatic activity may detoxify a compound which otherwise inhibits growth. Selection may also be coupled to a spectroscopic enzyme assay or other means of optical or ocular probing for enhanced activity.

When using directed evolution to improve enzyme activity under non-native conditions, such as elevated temperature or high concentration of a toxic or denaturing substance or organic solvent, the conditions may have to be introduced stepwise with gradually increasing stringency of selection pressure. This ensures that the enzyme template in each round has at least some rudimentary starting activity under the conditions used in that round. Stepwise increase of the stress factor, with intervening diversification between selection rounds, makes it possible to derive successively more effective and more tolerant enzymes or enzymes with new catalytic properties. Selection may be performed on agar plates (4-6), using a filter-lift assay (4) or using flow cytometry (27,28). Dan Tawfik showed that directed evolution of enzymes can be set up without use of living cells, for example, using *in vitro* compartmentalization in water-in-oil emulsion droplets containing ribosomes and library mRNA (29).

New reaction conditions

The early applications of directed evolution of enzymes aimed to optimise the stability and the performance under new reaction conditions such as high fractions of organic solvents (4), further rounds of directed evolution were added to reach a 471-fold activity increase over wild-type (30). Another example from the Arnold lab concerned the optimisation of a para-nitrobenzyl esterase for activity in the presence of a (31).

Many methods exist for increasing the thermostability of enzymes and other proteins. When directed evolution of enzymes is used with an aim to increase their thermal stability, the evolutionary process may be set up as interleaved heat treatment and activity assays, or alternatively the activity assays may be performed at elevated temperature. In one study, the thermostability of *Bacillus subtilis p*-nitrobenzyl esterase was increased by over 14°C (increase in T_m) after six generations of random mutagenesis, recombination via DNA shuffling, and

screening with interleaved heat treatment and activity assays (32). This work showed that it is possible to improve the thermal stability of an enzyme without compromising its catalytic activity at lower temperatures, if both properties are constrained. If not, the evolution of one property may come at the cost of the other, regardless of whether the two properties are inversely correlated or not correlated at all (32). Nature usually provides organisms adapted to cold or warm environments with two different enzymes having optimal catalytic properties at low or at high temperatures, respectively. Arnold showed that directed evolution can produce a single enzyme with high catalytic activity at both high and low temperatures (32-35). Another directed evolution strategy relied on structural information in the form of crystallographic B-factors, a measure of which regions are more or less ordered in a crystallised protein. By focusing the library of mutations to the 10 positions with highest B-factors, a large increase in enzyme stability was achieved (36).

In addition to deriving improved and novel biocatalysts, directed evolution studies contribute to our general understanding of natural protein evolution process and determinants of enzyme action, although the selection pressures operate over totally different time scales, population sizes, mutation rates, strength of selection, etc. Arnold and others have shown the importance for protein evolution of factors such as thermostability (37,38), the relative effects of random mutations and recombination (39), the importance of neutral drift for the evolution of protein function (40,41) and correlations between the rates of protein expression and evolution (42).

Choice of starting state

Arnold and co-workers have repeatedly shown that it is possible to evolve enzymes to improve their activity under new conditions in terms of solution composition, temperature, etc., and to change their catalytic activity towards new substrates and reactions. This is possible as long as the enzyme that is chosen as a starting point has at least some low level of activity for the intended reaction, i.e. some level of catalytic promiscuity (Figure 4, reviewed in e.g. 43-46). An inactive scaffold is not a suitable choice; directed evolution requires some low level of activity. Even a very low activity level towards the intended reaction provides a starting state to optimise through evolution. Often just a few mutations are required to boost up the new activity.

If an enzyme has a low level of activity for an intended reaction, but much higher activity for a natural one, it may be fruitful to first lower the natural activity before starting the directed evolution efforts towards the new intended reaction.

New chemical reactions

As a recent example of this latter strategy, the activity of tryptophan synthetase from *Pyrococcus furiosus* was first reduced by 95% through the removal of the non-catalytic domains of the enzyme. The isolated catalytic domain was subject to three rounds of directed evolution to introduce new catalytic activities towards synthesis of tryptophan analogues (47-51).

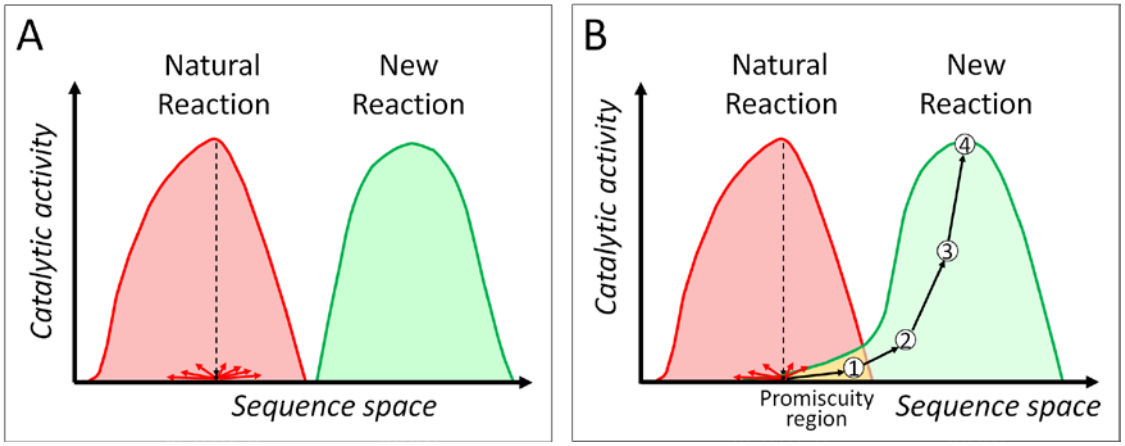

Figure 4. A: A starting point with no activity for the intended reaction is useless since no sequence variations (red arrows) create the new reactivity. B: A promiscuous enzyme with at least low activity for the intended reaction is a suitable staring point. Some combinations of random mutations may improve the new reactivity (black arrow). The first variant (1) serves as a starting state for sequential rounds of variation and screening →(2)→(3)→(4) for improved variants. Only a small number of cycles and are typically needed to boost up the new reactivity.

In a series of studies, Arnold and co-workers changed the activity of cytochrome P450 to catalyse a set of reactions for which no specific enzyme was previously available, for example, cyclopropanation. Cytochrome $P450_{BM3}$ has a catalytic promiscuity and an ability to catalyse, with very low efficiency, the cyclopropanation of styrene by ethyl-diazoacetate (EDA). Much more specific and efficient enzymes were evolved and only a small fraction (0.2%) of the amino acids in the enzyme needed to be changed to optimise the new catalytic activity (52-54). This included a change of the iron-ligating residue from Cys to Ser or His, leading to a shift in the characteristic 450-nm Soret peak in the absorbance spectrum of the enzyme to 411 nm. Therefore, the evolved enzymes were called cytochrome P411.

Other examples of reactions for which no natural enzymes have evolved are nitrene transfer reactions. In one case, Arnold and co-workers started from a cytochrome P411 variant performs azide reduction about 100 times more efficiently than nitrene transfer to sulphide. Using directed evolution they produced an enzyme variant that instead efficiently promotes the desired nitrene transfer process (55). There are several other examples of directed evolution of enzymes for carbene and nitrene transfer reactions (see for example 56,57).

Reactions with aliphatic and aromatic CH bonds are another tractable goal. Using directed evolution of cytochrome P450 monooxygenase, an enzyme was created that catalyses intermolecular amination of benzylic C–H bonds. The biocatalyst is enantioselective and lasts

for up to 1,300 turnovers, thereby providing an efficient biocatalyst for synthesis of valuable benzylic amines (58).

Other examples of evolution towards new reactions include generation of enzymes that catalyse arsenate detoxification (14), the production of highly strained carbocycles (59), and the switching of an enzyme from a galactosidase to a fucosidase (15).

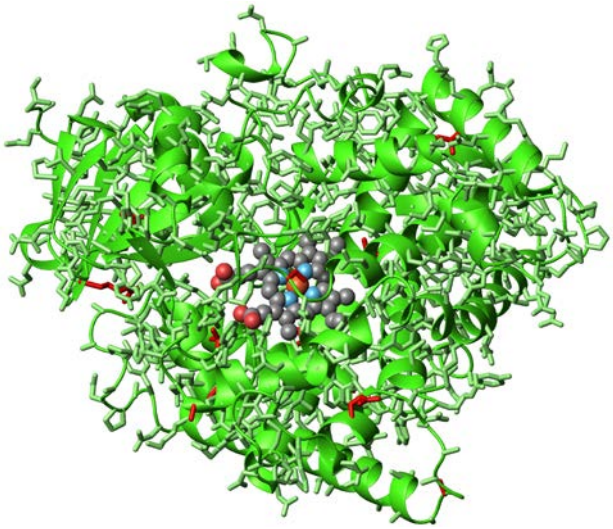

Figure 5. An evolved biocatalyst for cyclopropanation. The cytochrome P411 variant of cytochrome P450 (ref. 52) with the protein backbone shown as ribbon representation and side-chains as sticks. Side-chains that were mutated in engineered variants are shown in red.

Metabolic pathways

A strength of the directed evolution methodology is the ability to co-evolve enzymes in biosynthetic pathways. In one example Arnold and co-workers evolved a multi-enzyme pathway for carotenoid production in E. coli (60). Her lab also showed how whole-cell biocatalysts can be developed for the production of valuable chemicals by using directed evolution to enable the production of L-methionine in *E. coli* (61).

Biofuels

One challenge for mankind is finding suitable replacements or supplements for fossil fuels, which can be produced in a sustainable and environmentally-friendly manner. Here, one seeks to produce alcohols from short-chain alkanes (62) and a leading candidate biofuel is 2-methylpropan-1-ol (isobutanol). Isobutanol can be produced using a biosynthetic pathway in recombinant *Escherichia coli*. Two enzymes in the pathway, however, require reduced nicotinamide adenine dinucleotide phosphate (NADPH) as a cofactor, while glycolysis, the normal metabolism during growth of *E. coli* produces reduced nicotinamide adenine dinucleotide (NADH). To resolve this obstacle, Arnold and co-workers used directed evolution to alter the co-factor dependence of the enzymes so they can instead rely on NADH, making the enzymes and thereby the organism suitable for biofuel production (63).

New chemical bonds

Carbon-silicon bonds are common in human-made chemicals but absent in biology. Nature has not evolved enzymes that catalyse the formation of carbon-silicon bonds. However, directed evolution can be used as a strategy to ensure that such chemistry invented by humans can also be conducted by help of enzymes. Arnold and co-workers noted that haem proteins can catalyse non-natural carbene-insertion reactions. After screening a number of haem proteins from various organisms, they decided to use cytochrome c from *Rhodothermus marinus* as a starting point. This protein catalyses the formation of carbon-silicon bonds with low efficiency, but with 97% enantiomeric excess (ee; 64). A small library of variants was screened during heat treatment and in catalytic activity assays, and the best candidate was subjected to further mutagenesis and screening. The result of this work is an enzyme that catalyses silicon-carbon bond formation 40 times better than the starting enzyme and with 99% ee (64). The evolved enzyme had 15 times higher turnover number than the best non-enzyme catalyst known for the same reaction. This example shows that it is possible to expand the scope of enzyme-catalysed reactions in terms of which kinds of bonds are formed by the engineered enzyme.

Other examples of bonds and reactions not catalysed by any enzyme found in nature, but for which directed evolution was used to create efficient enzymes, are carbon-borane bonds (65) and enantio-selective intramolecular C-H amination (66).

Enantio-selectivity

Directed evolution is an efficient way to improve the enantio-selectivity of enzymes, i.e., enhancing their performance in asymmetric catalysis. The evolved enzymes are used in the production of chiral substances with high enantiomer purity. An early example of directed evolution with the aim of improving the enantio-selectivity of an enzyme was reported by Matcham and Bowen concerning transaminases in the catalysis of chiral amine production (67). This work started with an enzyme with low level of S-selectivity (65% ee) in the conversion of the ketone β-tetralone to aminotetraline, the corresponding amine. A library of mutants was generated and screened for enhanced activity on the S-isomer but not the R-isomer. The result was a biocatalyst that produced the S-aminotetraline with greatly enhanced selectivity (94% ee), which was further improved by additional rounds of mutagenesis and screening (67). Manfred Reetz and co-workers reported another early example that has led to improved enantio-selectivity of lipases in ester hydrolysis (68). Through directed evolution via four cycles of random mutagenesis, the selectivity factor of a bacterial lipase from *Pseudomonas aeruginosa* was first increased from 1.1 to 11 (68) and then to 35 after further diversification of the library (69). Other early examples are found in (70-73).

Directed evolution in organic synthesis and industry

Directed evolution quickly made its way from the academic setting to industrial applications (74-77). Enzymes developed using directed evolution are used in industry in the production of

biofuels, materials, bulk and fine chemicals, detergents, consumer products, laboratory reagents and pharmaceuticals, as well as intermediates for the pharmaceutical industry. Several of the enzymes developed in the Arnold lab are used in industry. Many companies have their own scientific teams applying directed evolution strategies to improve catalysts or protein-based therapeutics in terms of stability, activity, specificity or other properties. Specific examples of evolved enzymes and products are taste enhancers, drugs against diabetes and vascular plaques, as well as lipid-lowering pharmaceuticals. Some enzymes produced by directed evolution are made on very large scale. This includes lipases used in detergents. Industrial chemicals are made in enormous quantities with the help of biocatalysts produced by directed evolution.

A green alternative

Directed evolution provides enzymes with unique specificity, thereby offering environmentally-friendly biocatalysts. Enzymes developed using directed evolution have replaced harsh industrial processes with milder biotechnology that does not require toxic metals or large amounts of organic solvent. Industrial use of enzymes developed using directed evolution have, for example, replaced chemical catalysts in asymmetric synthesis and provide a green alternative that leads to lower consumption of organic solvents and lower amounts of side products and waste.

Directed evolution in protein design

Directed evolution seeks to alter the activities of already-existing enzymes based on molecular insights combined with a large element of randomness. Orthogonal to directed evolution is the rational design of proteins. Protein design is based on *ab initio* or empirical calculations and aims to design proteins from scratch. However, in the protein design field, it is widely acknowledged that in order to reach an acceptable level of fitness, for example, in terms of binding affinity and specificity, it is at our current level of knowledge necessary to add directed evolution as a final optimisation step (see for example refs. 78-80).

Summary and outlook

Directed evolution of enzymes has become a highly efficient protocol for development of biocatalysts with high specificity, limited side reactions and tolerance of diverse reaction conditions. Directed evolution is a versatile and efficient path towards optimised enzymes and enzymes with novel functions. A main conclusion emerging from directed evolution research is that enzymes can indeed be tuned to catalyse new reactions, and to reactions very different from the ones catalysed by nature's own enzymes. There is ample room for optimisation and redirection of enzyme function in terms of reactivity, substrate specificity and chemical reactions, as well as tolerance to various reaction conditions. We are probably very far from the limit of which reactions enzymes can catalyse – there is plenty of room for further discovery.

Phage display of peptides and antibodies

Directed evolution of binding proteins is facilitated by a physical coupling between phenotype (high affinity high selectivity binding protein) and genotype (DNA sequence). This is provided by phage display.

Phage display of peptides

Phage display represents a major technology breakthrough and was developed by **George P Smith** (81). The DNA that codes for a specific protein member of the library is packaged inside the phage in such a way that the phage presents the protein on its surface. This simplifies screening based on binding to receptors. This also simplifies the identification and amplification of phages displaying the best binding proteins. Infection by the phages of *E. coli* and multiplication in the host produces an enriched library after each round of selection. It is also possible to re-diversify the phage library between the selection rounds.

In a seminal paper, Smith showed that a peptide could be inserted in a loop of protein III, a minor coat protein on the surface of fusion phage, and that the displayed peptide retained interaction with its target (81). In this work, the displayed peptide was a 57-residue fragment of a restriction endonuclease. Using a single round of affinity purification versus antiserum to the endonuclease, Smith showed that phages presenting the endonuclease peptide inserted into coat protein III of their surface could be enriched 1,000-fold over other phages. He outlined a way forward towards much greater enrichment, which might be achieved by using a sequence of affinity purification rounds. He also predicted that it would be possible to isolate clones from a library of random inserts in a phage vector using affinity purification with an antibody as the bait (81).

In another publication, from 1988 (82), Parmley and Smith introduced several technical improvements to the phage display technology, for example, by moving the location of the displayed peptide within protein III. This was motivated by the fact that protein III mediates infection of *E. coli* by binding to its F pilus. Fully functional protein III is thus essential for propagation of the phage. Smith introduced the term “biopanning” for affinity purification of phages displaying high affinity peptides from a background of lower affinity ones, using streptavidin-coated Petri dishes to which biotinylated target was coupled. This offered an easy device for capture, washing and elution in the affinity-purification procedure. Smith showed the feasibility of his improved approach by achieving 10^8-fold enrichment of phages displaying a peptide representing the epitope of an anti-β-galactosidase antibody using affinity purification towards its target (82).

Smith also proposed that phage display of peptides might aid in vaccine development (81). This inspired studies where peptides from the malaria parasite *Plasmodium falciparum* were displayed on the surface of filamentous phage and found to be active as antigens (83,84).

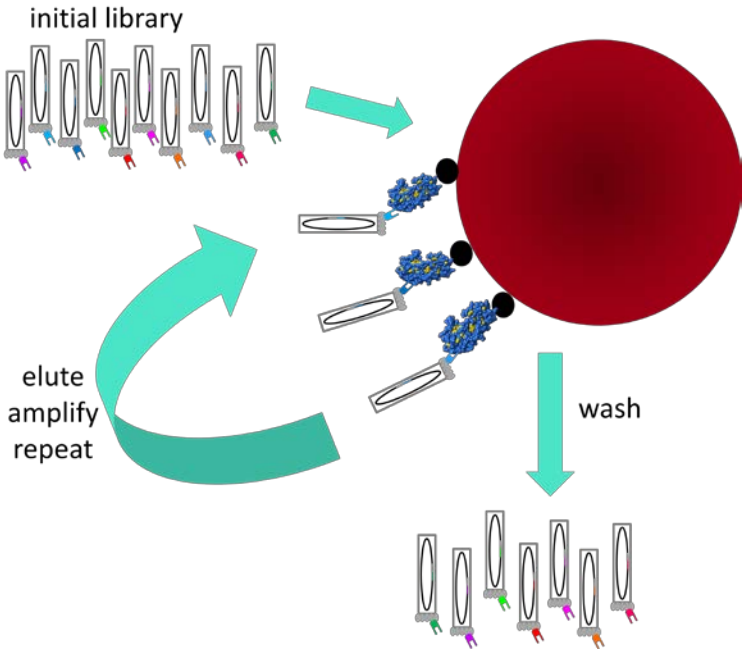

Figure 6. Selection of high-affinity binding proteins from phage display libraries. An initial library of peptide or protein variants displayed on protein III at the tip of the phage (upper left) is added to a target protein (blue-yellow) immobilized on a solid support, here magnetic beads (red). After extensive washing to remove weakly bound phages (lower right), the best binding variants are eluted (e.g. using acid) and used to infect E coli to produce an amplified library enrich in high-affinity members for a next round of affinity purification. The procedure can be repeated as many times as needed to obtain the desired level of selection.

Phage display of peptide libraries

In his 1985 paper (81), Smith predicted that it would be possible to isolate clones from a library of random inserts in a phage vector. In 1988 (82), he wrote that the expression of a large number of peptides with random sequences on the phage surface might be useful for finding the epitopes recognised by antibodies. He reasoned that such technology might reveal the epitope of almost any antibody, since many antibodies recognise short linear segments of five-six amino acid residues. The framework for producing phage display libraries for selection of binding peptides was thus developed in the Smith laboratory, but also immediately picked up by other groups. A number of publications appeared almost simultaneously in the early 1990s, describing the expression and display of combinatorial sequence libraries of short peptides on the surface of filamentous phage, and the selection of peptide binders for various targets.

The first two papers appeared in the same issue of *Science* (85,86). Both libraries were expressed in fusion with coat protein III. Devlin and co-workers reported a library of 20 million different 15-mer peptides, which was used to deduce a peptide binding motif, HPQ, that confers high affinity for the biotin-binding site of streptavidin (85).

Scott and Smith reported a library of 40 million 6-mer peptides used to determine the epitopes of two different monoclonal antibodies that had been raised against a six-residue fragment of myohaemerythrin (86). Antibody-binding peptides were affinity-purified from the libraries through three successive rounds of biopanning, followed by infection of *E. coli* to produce an enriched phage-display library for the next round, and so on. After three rounds of affinity purification, combined with affinity measurements for selected clones, Smith and co-worker showed that only the very highest affinity peptides were retained. A consensus binding motif for the antibody epitope could be deduced based on DNA sequencing of individual clones (86). By comparing a large number of amino acid sequences, they found that the first three amino acids of the epitope are most important for conferring high affinity binding to the antibody (86).

A random library of hexa-peptides was used to select high-affinity binding peptides for an anti-β-endorphin antibody; after three rounds of affinity purification and sequencing of 51 clones, a consensus motif could be deduced and the two most N-terminal residues were found to be the most important in this epitope for determining the antibody affinity (87).

Phage display of peptide libraries made it possible to determine epitopes for antibodies. Through the phage display of antibodies, researchers could reverse the selection procedure, which was an important step towards developing therapeutic antibodies.

Phage display of antibody fragments

In 1990, **Sir Gregory P Winter** reported the display of a folded and fully functional antibody fragment on filamentous phage (88). Intact IgG are large bivalent molecules composed of more than one amino acid chain and therefore are challenging to express on the phage surface. Winter and co-workers instead displayed a single-chain variable fragment, scFv, a construct in which the variable part of the heavy chain was covalently linked via a flexible polypeptide linker to the light chain (89). As such, an scFv carries a single antigen-binding site composed of six variable loops, so-called complementary-defining regions (CDRs).

In the first work (88), Winter and co-workers displayed an scFv derived in immunized mice against hen egg-white lysozyme in fusion with the filamentous phage protein III. They showed that the scFv-phages bound to hen egg-white lysozyme, but not human lysozyme or turkey egg-white lysozyme. The very high specificity implied that the scFv was displayed in functional folded form on the phage surface. A millionfold enrichment of phage carrying the antigen-binding fragment over other phages was achieved using two rounds of affinity purification with intervening amplification of the retained phages (88).

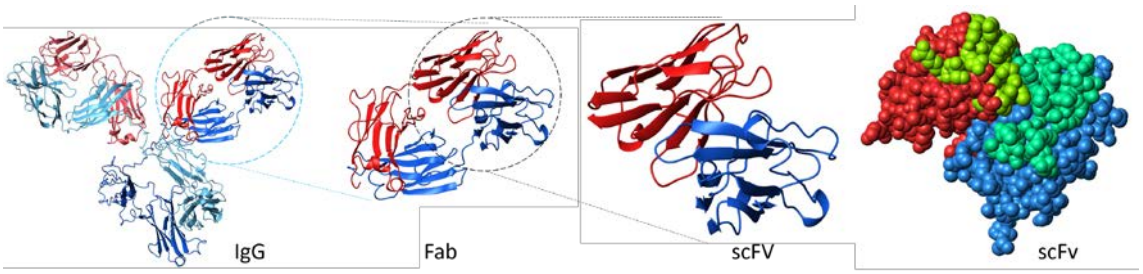

Figure 7. Definition of antibody fragments. Immunoglobulin G (IgG) structure (left) with the two heavy chains in blue and light blue and the two light chains in red and pink, with the location of Fab and scFv (89) indicated. The space-filling model of an scFv (right) shows the variable loops in green.

This achievement was the starting point of a pharmaceutical revolution: the proven efficacy of using phage display of functional antibody fragments for the selection of high-affinity binders for a defined antigen. The technology now makes it possible to bypass immunisation for the generation of monoclonal antibodies with fully human scaffolds, and the possibility to display combinatorial libraries of human antibodies facilitates the development of therapeutic antibodies. Moreover, the phage-display technology makes it possible to derive antibodies recognising toxic antigens, or antigens that would otherwise compromise the host.

Inspired by the phage-display of peptide libraries (81,82,85,86), Winter outlined several future directions and applications of antibody phage display (88). He proposed construction and screening of large combinatorial antibody libraries created by random combination of genes for the heavy- and light-chain variable domains for display on filamentous phage. He also proposed the production and screening of fully synthetic antibody libraries on phage.

Another example of phage display of antibody fragments, in this case along the entire length of the phage surface, in fusion with the major coat protein VIII, was reported by Richard Lerner and co-workers in 1991 (90). This work displayed somewhat larger antibody fragments, so-called Fab fragments, derived in mice against tetanus toxin. Two domains of the antibody heavy chain were displayed on the phage and two domains of the antibody light chain were expressed in secreted form. The fragments from the two chains reconstituted the Fab on the phage surface and a 2,700-fold enrichment over non-specific phage was achieved using biopanning; the researchers proposed that higher enrichment would arise using sequential rounds of biopanning (90).

Also in 1991, the phage display of heterodimeric Fab fragments from immunized mice on protein III was reported by Winter and co-workers (91). Two domains of one antibody chain (heavy or light) were expressed in fusion with phage protein III and two domains of the opposite antibody chain (light or heavy, respectively) in secreted form. Both orientations were found to yield

functional Fab; the two fragments reconstituted the Fab on the surface of the filamentous phage. It was reasoned that this provided a format for exploring a much wider variety of sequences (10^{14}) than achievable in single libraries (10^{7}) through the combinatorial advantage of co-expressing two libraries (91). Another study from 1991 displayed a mouse Fab fragment directed against the extracellular domain of the HER2 receptor, and showed that it could be enriched millionfold from a background of non-binding phages (92), or enriched relative to mutated Fab fragment with lower affinity for the same antigen.

Phage display of antibody libraries

A number of antibody libraries displayed on filamentous phage were reported during 1991. This work was pursued in several laboratories in parallel, including the Winter and Lerner laboratories.

The first report described the phage display of a library of scFv antibody fragments (8). This library was derived from mice immunized with the small-molecule antigen 2-phenyloxazol-5-one (phOx) and was displayed at the N-terminus of protein III of the filamentous phage. Using phOx antigen immobilised on a column resin, Winter and co-workers showed that it was possible to select phOx-binding antibodies from the initial library and that the fraction of high-affinity antibodies increased upon sequential rounds of selection. Moreover, they found that higher affinity antibodies could be derived if the variable heavy-chain (VH) and variable light-chain (VL) genes were reshuffled between the affinity selection rounds to produce new combinations not present in the immunized mice (8). The original library of contained 200,000 scFv sequences, but after reshuffling of the VH and VL genes, this was increased to 40 million members, and a larger number of high-affinity phOx binders was retrieved.

Soon after, the display of a combinatorial library of Fab fragments on phage surface protein III was reported (93). This library contained a low fraction of Fab fragments derived from mice immunized with tetanus toxin. The library was used to select binders to tetanus toxin from a background of non-binding phage using biopanning (93). Another publication reported phage-displayed Fab-libraries of 10 million sequence variants, which were derived from the bone marrow of HIV-positive humans (94). These libraries were used to derive high-affinity Fab fragments against the surface glycoprotein gp120 of the virus.

Two scFv libraries of human origin were reported in 1991, displayed on protein III of the filamentous phage (95). These libraries were constructed as combinations of the variable parts of the heavy and light chains of IgM (27 million members) or of the variable parts of the heavy and light chains of IgG (160 million members) derived from non-immunised humans. The ability to derive high-affinity antibodies from these libraries using four rounds of biopanning was demonstrated using as antigens the small molecule phOx, as well as the proteins lysozyme from turkey egg-white and bovine serum albumin.

Early antibody libraries

Prior to antibody phage display (88), a few other formats were tested for screening or selection of high-affinity binders. In 1989, Winter and co-workers reported antibody fragment libraries that could be handled outside cells for selection *in vitro* (96). Native repertoires of antibody VH fragments were derived from mice immunized with lysozyme or haemocyanine and cloned in a format allowing secretion from *E. coli*, alone or in combination with VL (Vκ) chains. Antibodies could be selected from the secreted libraries using immobilised antigen. In a second example from the same year, Lerner and co-workers expressed a large combinatorial library of Fab fragments representing the mouse antibody repertoire in *E. coli* using a lambda vector (97). Screening of this library for antigen-binding members was performed using a plaque lift assay. The technologies used in these earlier attempts for selection and screening of antibody libraries were soon abandoned in favour of the more efficient phage display technology.

A note on humanization

The first monoclonal antibodies produced by hybridoma were not easily adapted to become human medicines, despite their high affinity and high selectivity. These mouse proteins were not well tolerated by humans. Winter invented a “humanization protocol”, in which the most antigenic portions of the mouse monoclonal antibody are replaced by an invariant human sequence. With the introduction of phage displayed human antibodies, the need for “humanization” is avoided.

Role of phage protein copy number and multivalent vs. monovalent display

Some studies have investigated the display on the major coat protein VIII, of which there are about 2,700 copies along the entire length of the filamentous phage (90,98). However, in most phage display libraries, the varied peptides or proteins are fused protein III, of which there are up to five copies at the tip of the phage. The rationale behind the choice of protein VIII has been to maximise avidity effects and to avoid possible interference with phage infection of *E. coli* when protein III is used (90). More commonly nowadays, multi-valency is avoided to permit sorting of phages based on affinity rather than avidity. Early protein III library constructs may have had as many as four to five copies of peptide displayed on the four to five copies of protein III, a feature that may lead to selection based on avidity rather than affinity (86,87,99). Monovalent display, i.e., a maximum of one protein III copy per phage, made selection based on affinity possible, and the use of helper phages eliminated any infection bias between members of the library. Wells and co-workers introduced monovalent phage display with the use of helper phages in the case of human growth hormone (hGH; H100). The methodology was further improved for the display of hGH libraries (J101,102). Other early examples of monovalent phage display facilitated the affinity-based selection of members in antibody libraries (92,93).

Other display formats

Phage display is the first example of a method for evolution of binding proteins, which offers a physical coupling of phenotype (protein sequence) and genotype (DNA sequence). Several other library display methods that couple phenotype to genotype have since been developed, for example, ribosome display (103,104), bacterial display (105) and yeast display (106). Still, phage display remains the most widely used platform for derivation of antibodies and antibody fragments, with huge medical impact.

Methods for selection of high-affinity binding peptides and proteins

The standard method used for selection of binding peptides or proteins from phage display libraries is in principle the same as classical affinity purification; those individuals in the library with highest affinity for a target are captured on a solid support, followed by multiple and extensive washing steps before the best binding peptides with their connected phages are eluted with acid, neutralised and used to infect *E. coli* to produce an enriched library. The enriched library thus contains many copies of several higher-affinity peptides and fewer copies of lower-affinity peptides. The whole experiment is repeated several times with intervening amplification in *E. coli* and successive increase of the stringency of the selection by lowering the amount of target used. The selection of phage-displayed peptides and proteins is typically performed using targets immobilised on chromatography resins, Petri dishes, magnetic beads or other supports. However, in cases of cellular display, it is more common to use flow cytometry, in which case the binding event may be coupled to a fluorescence readout.

Selection based on kinetics, stability or protease sensitivity

It is possible to base the selection strategy on association or dissociation rates between the target and library members, rather than on affinity, by careful choice of incubation and washing times (J3, J4). By invoking some kind of selection pressure such as elevated temperature or the presence of proteases, phage display can be used to improve protein stability (J5, J6). In one of the first examples where selection was based on other properties than affinity, a linker was placed between phage protein III and hGH. The linker sequence was randomised and challenged by the presence of specific proteases when hGH receptor was used as a bait; this allowed the sequence determinants of protease sensitivity versus protease resistance to be determined (J9). Phage display can provide useful molecular information about the sequence determinants of binding affinity and specificity (J10), as well as the sequence determinants of protein stability and folding specificity (J10, J11).

Applications and outlook

Directed evolution of binding proteins has become a highly efficient protocol for development of therapeutic antibodies. Phage display of peptides and antibodies is used to derive variants with high affinity for a given target, and to provide valuable information regarding the sequence

requirements and molecular driving forces for affinity and specificity in binding reactions, protein folding and stability.

Directed evolution to improve the binding affinity while retaining the high selectivity of antibodies has been used in the past 25 years. Such “affinity maturation” procedures have derived antibodies with affinities for their antigens in the 10^{12}-10^{15} M^{-1} range representing a thousand-to-millionfold improvement over antibodies resulting from immunization, and later from phage display libraries. Such very high affinity enables their use as subcutaneous, self-administered therapeutic agents, instead of requiring larger quantities injected intravenously at a doctor’s office.

The first approved fully human therapeutic antibody derived using phage display was Adalimumab (approved 2002). Adalimumab binds with very high affinity to the secreted protein TNF-α, a pro-inflammatory cytokine, and is used in treatment of rheumatoid arthritis, psoriasis and inflammatory bowel disease, among other conditions. A range of human antibodies have been and will be identified using phage display technology and are in clinical use against for instance inflammatory diseases and cancer.

Sara Snogerup Linse
Professor of Physical Chemistry and Molecular Protein Science, Lund University
Member of the Nobel Committee for Chemistry

References

1. Eigen M, Gardiner W (1984) Evolutionary molecular engineering based on RNA replication. Pure Appl. Chem. 56, 967-978.
2. Guterman and Ladner (1991) EP0436597 B, published 17 Juli 1991, priority date 2 sep 1988.
3. Chen KQ, Arnold FH (1991) Enzyme engineering for nonaqueous solvents: random mutagenesis to enhance activity of subtilisin E in polar organic media. Biotechnology (N Y). 9:1073-1077.
4. Chen K, Arnold FH (1993) Tuning the activity of an enzyme for unusual environments: sequential random mutagenesis of subtilisin E for catalysis in dimethylformamide. Proc Natl Acad Sci U S A. 90:5618-5622.
5. Arnold FH (1993) Engineering proteins for unusual environments. FASEB J 7:744-749.
6. Stemmer WP (1994) Rapid evolution of a protein in vitro by DNA shuffling. Nature 370:389-391.
7. Stemmer WP (1994) DNA shuffling by random fragmentation and reassembly: in vitro recombination for molecular evolution. Proc Natl Acad Sci U S A. 91:10747-10751.
8. Clackson T, Hoogenboom HR, Bonnert TP, McCafferty J, Griffiths AD, Winter G (1991) Making antibody fragments using phage display libraries. Nature 352:624-628.
9. Marks JD, Griffiths AD, Malmqvist M, Clackson TP, Bye JM, Winter G. (1992) By-passing immunization: Building high affinity human antibodies by chain shuffling. Nat. Biotech. 10, 779-783.

10. Collet TA, Roben P, O'Kennedy R Barbas III, CF, Burton DR, Lerner RA (1992) A binary plasmid system for shuffling combinatorial antibody libraries. Proc Natl Acad Sci U S A. 89:10026-10030.
11. Crameri A, Stemmer WP (1995) Combinatorial multiple cassette mutagenesis creates all the permutations of mutant and wild-type sequences. Biotechniques 18:194-196.
12. Crameri A, Whitehorn EA, Tate E, Stemmer WP (1996) Improved green fluorescent protein by molecular evolution using DNA shuffling. Nat Biotechnol 14:315-319.
13. Miyazaki K, Arnold FH (1996) Exploring nonnatural evolutionary pathways by saturation mutagenesis: rapid improvement of protein function. J Mol Evol. 49:716-720.
14. Crameri A, Dawes G, Rodriguez E Jr, Silver S, Stemmer WP (1997) Molecular evolution of an arsenate detoxification pathway by DNA shuffling. Nat Biotechnol. 15:436-438.
15. Zhang JH, Dawes G, Stemmer WP (1997) Directed evolution of a fucosidase from a galactosidase by DNA shuffling and screening. Proc Natl Acad Sci U S A. 94:4504-4509.
16. Moore JC, Jin HM, Kuchner O, Arnold FH (1997) Strategies for the in vitro evolution of protein function: enzyme evolution by random recombination of improved sequences. J Mol Biol. 272:336-347.
17. Zhao H, Arnold FH (1997) Functional and nonfunctional mutations distinguished by random recombination of homologous genes. Proc Natl Acad Sci U S A. 94:7997-8000.
18. Zhao H, Arnold FH (1997) Optimization of DNA shuffling for high fidelity recombination. Nucleic Acids Res. 25:1307-1308.
19. Zhao H, Giver L, Shao Z, Affholter JA, Arnold FH (1998) Molecular evolution by staggered extension process (StEP) in vitro recombination. Nat Biotechnol 16:258-61.
20. Crameri A, Raillard SA, Bermudez E, Stemmer WP (1998) DNA shuffling of a family of genes from diverse species accelerates directed evolution. Nature 391:288-291.
21. Shao Z, Zhao H, Giver L, Arnold FH (1998) Random-priming in vitro recombination: an effective tool for directed evolution. Nucleic Acids Res 26:681-683.
22. Miyazaki K, Arnold FH (1999) Exploring non-natural evolutionary pathways by saturation mutagenesis: rapid improvement of protein function. J Mol Evol 49:716-720.
23. Christians FC, Scapozza L, Crameri A, Folkers G, Stemmer WP (1999) Directed evolution of thymidine kinase for AZT phosphorylation using DNA family shuffling. Nat Biotechnol 17:259-264.
24. Ostermeier M, Sinn JH, Benkovic SJ (1999) A combinatorial approach to hybrid enzymes independent of DNA homology. Nat Bioctech 17:1205-1209.
25. Sieber V, Martinez CA, Arnold FH (2001) Libraries of hybrid proteins from distantly related sequences. Nat biotechnol 19:456-460.
26. Voigt CA, Martinez C, Wang ZG, Mayo SL, Arnold FH (2002) Protein building blocks preserved by recombination. Nat Struct Biol 9:553-558.
27. Rice GC, Goeddel DV, Cachianes G, Woronicz J, Chen EY, Williams SR, Leung DW (1992) Random PCR mutagenesis screening of secreted proteins by direct expression in mammalian cells. Proc Natl Acad Sci USA 89:5467-5471.
28. Fu AY, Spence C, Scherer A, Arnold FH, Quake SR (1999) A microfabricated fluorescence-activated cell sorter. Nat Biotechnol 17:1109-1111.
29. Tawfik DS, Griffiths AD (1998) Man-made cell-like compartments for molecular evolution. Nat Biotechnol 16:652-656.
30. You L, Arnold FH (1994) Directed evolution of subtilisin E in *Bacillus subtilis* to enhance total activity in aqueous dimethylformamide. Protein Eng 1996;9:77–83.
31. Moore JC, Arnold FH (1996) Directed evolution of a para-nitrobenzyl esterase for aqueous-organic solvents. Nat Biotechnol 14:458-467.

32. Giver L, Gershenson A, Freskgard PO, Arnold FH (1998) Directed evolution of a thermostable esterase. Proc Natl Acad Sci U S A. 95:12809-12813.
33. Zhao HM, Arnold FH (1999) Directed evolution coverts subtilisin E into a functional equivalent of thermitase. Protein Eng. 12:47-53.
34. Wintrode PL, Miyasaki K, Arnold FH (2000) Cold adaption of a mesophilic subtilisin-like protease by laboratory evolution. J Biol Chem 275:31635-31640.
35. Wintrode PL, Miyasaki K, Arnold FH (2001) Patterns of adaption of a mesophilic subtilisin-like protease by laboratory evolution. BBA Protein Struct Mol Evol 1549:1-8.
36. Reetz MT, Carballeira JD, Vogel A. Iterative saturation mutagenesis on the basis of B factors as a strategy for increasing protein thermostability. Angew Chem Int Ed 2006;45:7745–7751.
37. Bloom JD, Wilke CO, Arnold FH, Adami C (2004) Stability and evolvability of function in model proteins. Biophys J 86:2758-2764.
38. Bloom JD, Labthavikul ST, Otey CR, Arnold FH (2006) Protein stability promotes evolvability. Proc Natl Acad Sci USA 103:5869-5874.
39. Drummond DA, Siberg JJ, Meyer MM, Wilke CO, Arnold FH (2005) On the conservative nature of intragenic recombination. Proc Natl Acad Sci USA 102:5380-5385.
40. Bloom JD, Romero PA, Lu Z, Arnold FH (2007) Neutral drift ca alter promiscuous protein functions, potentially altering functional evolution. Biology Direct 2.
41. Bershtein S, Segal M, Bekerman R, Tokuriki N, Tawfik D (2006) Robustness-epistasis link shapes the fitness landscape of a randomly drifting protein. Nature 444:929-932.
42. Drummond DA, Bloom JD, Adami C, Wilke CO, Arnold FH (2005) Why highly expressed proteins evolve slowly. Proc Natl Acad Sci USA 102:14338-14343.
43. Kazlauskas RJ (2005) Enhancing catalytic promiscuity for biocatalysis. Curr Opin Chem Biol. 9:195-201.
44. Bornscheuer UT, Kazlauskas RJ (2004) Catalytic promiscuity in biocatalysis: using old enzymes to form new bonds and follow new pathways. Angew Chem Int Ed Engl 43:6032-6040.
45. Arnold FH (2015) The nature of chemical innovation: new enzymes by evolution. QRB 48:404-410.
46. Arnold FH (2018) Directed evolution: bringing new chemistry to life. Angew. Chem. Int. Ed. 57:4143-4148.
47. Buller AR, Brinkmann-Chen S, Romney DK, Herger M, Murciano-Calles J, Arnold FH (2015) Directed evolution of the tryptophan synthase β-subunit for stand-alone function recapitulates allosteric activation. Proc Natl Acad Sci U S A 112:14599-14604.
48. Herger M, van Roye P, Romney DK, Brinkmann-Chen S, Buller AR, Arnold FH (2016) Synthesis of β-Branched Tryptophan Analogues Using an Engineered Subunit of Tryptophan Synthase. J Am Chem Soc. 138:8388-8391.
49. Murciano-Calles J, Romney DK, Brinkmann-Chen S, Buller AR, Arnold FH (2016) A Panel of TrpB Biocatalysts Derived from Tryptophan Synthase through the Transfer of Mutations that Mimic Allosteric Activation. Angew Chem Int Ed Engl 55:11577-11581.
50. Boville CE, Romney DK, Almhjell PJ, Sieben M, Arnold FH (2018) Improved Synthesis of 4-Cyanotryptophan and Other Tryptophan Analogues in Aqueous Solvent Using Variants of TrpB from Thermotoga maritima. J Org Chem 83:7447-7452.
51. Buller AR, van Roye P, Cahn JKB, Scheele RA, Herger M, Arnold FH (2018) Directed Evolution Mimics Allosteric Activation by Stepwise Tuning of the Conformational Ensemble. J Am Chem Soc 140:7256-7266.

52. Coelho PS, Wang ZJ, Ener ME, Baril SA, Kannan A, Arnold FH, Brustad EM (2013) A serine-substituted P450 catalyzes highly efficient carbene transfer to olefins in vivo. Nat Chem Biol. 9:485-487.
53. Coelho PS, Brustad EM, Kannan A, Arnold FH (2013) Olefin cyclopropanation via carbene transfer catalyzed by engineered cytochrome P450 enzymes. Science 339:307-310.
54. Wang ZJ, Renata H, Peck NE, Farwell CC, Coelho PS, Arnold FH (2014) Improved cyclopropanation activity of histidine-ligated cytochrome P450 enables the enantioselective formal synthesis of levomilnacipran. Angew Chem Int Ed Engl. 53:6810-6813.
55. Prier CK, Hyster TK, Farwell CC, Huang A, Arnold FH (2016) Asymmetric Enzymatic Synthesis of Allylic Amines: A Sigmatropic Rearrangement Strategy. Angew Chem Int Ed Engl. 55:4711-4715.
56. Brandenberg OF, Fasan R, Arnold FH (2017) Exploiting and engineering hemoproteins for abiological carbine and nitrene transfer reactions. Curr Opin Biotechnol. 47:102-111.
57. Lewis RD, Garcia-Borràs M, Chalkley MJ, Buller AR, Houk KN, Kan SBJ, Arnold FH (2018) Catalytic iron-carbene intermediate revealed in a cytochrome *c* carbene transferase. Proc Natl Acad Sci U S A 115:7308-7313.
58. Prier CK, Zhang RK, Buller AR, Brinkmann-Chen S, Arnold FH (2017) Enantioselective, intermolecular benzylic C-H amination catalysed by an engineered iron-haem enzyme. Nat Chem. 9:629-634.
59. Chen K, Huang X, Kan SBJ, Zhang RK, Arnold FH (2018) Enzymatic construction of highly strained carbocycles. Science 360:71-75.
60. Schmidt-Dannert C, Umeno D, Arnold FH (2000) Molecular breeding of carotenoid biosynthetic pathways. Nat biotechnol 18:750-753.
61. May O, Nguyen PT, Arnold FH (2000) Inverting enantioselectivity by directed evolution of hydantoinase for improved prodcutionof L-methionine. Nat biotechnol 18:317-320.
62. Glieder A, Farinas ET, Arnold FH (2002) Laboratory evolution of a soluble, self-sufficient, highly active alkane hydroxylase. Nature Biotech 20:1135–1139.
63. Bastian S, Liu X, Meyerowitz JT, Snow CD, Chen MMY, Arnold FH (2011) Engineered ketol-acid reductoisomerase and alcoholdehydrogenase enable anaerobic 2-methylpropan-1-ol production a ttheoretical yield in Escherichia coli. Metabolic Engineering 13 (2011) 345–352.
64. Kan SB, Lewis RD, Chen K, Arnold FH (2016) Directed evolution of cytochrome c for carbon-silicon bond formation: Bringing silicon to life. Science 354:1048-1051.
65. Kan SBJ, Huang X, Gumulya Y, Chen K, Arnold FH (2017) Genetically programmed chiral organoborane synthesis. Nature 552:132-136.
66. McIntosh JA, Coelho PS, Farwell CC, Wang ZJ, Lewis JC, Brown TR, Arnold FH (2013) Enantioselective intramolecular C-H amination catalyzed by engineered cytochrome P450 enzymes in vitro and in vivo. Angew Chem Int Ed Engl 52:9309-9312.
67. Matcham GW, Bowen ARS (1996) Biocatalysis for chiral indermediates: meeting commercial and technical challenges. Chimica Oggi (Chemistry Today) 14, 20-24.
68. Reetz, M.T., Zonta, A., Schimossek, K., Liebeton, K. & Jaeger, K.-E. (1997). Creation of enantioselective biocatalysts for organic chemistry by in vitro evolution. Angew. Chem. Int. Ed. Engl. 36, 2830-2832.
69. Liebeton K, Zonta A, Schimossek K, Nardini M, Lang D, Dijkstra BW, Reetz MT, Jaeger KE (2000) Directed evolution of an enantioselective lipase. Chem Biol. 7:709-718.
70. Reetz, M.T., Becker, M.H., Klein, H.-W. & Stoëckigt, D.A. (1999). Method for high-throughput screening of enantioselective catalysts. Angew. Chem. Int. Ed. 38, 1758-1761.

71. Reetz MT, Jaeger KE (2000) Enantioselective enzymes for organic synthesis created by directed evolution Chemistry 6:407-412.
72. Reetz MT (2000) Evolution in the test tube as a means to create enantioselective enzymes for use in organic synthesis. Sci Prog. 83(Pt2):157-172.
73. Jaeger KE, Reetz MT (2000) Directed evolution of enantioselective enzymes for organic chemistry. Curr Opin Chem Biol 4:68-73.
74. Fasan R, Crook NC, Peters MW, Meinhold P, Buelter T, Landwehr M, Cirino PC, Arnold FH (2011) Improved product-per-glucose yields in P450-dependent propane biotransformations using engineered Escherichia coli. Biotechnol Bioeng. 108:500-10.
75. Fischbach MA, Lai JR, Roche ED, Walsh CT, Liu DR (2007) Directed evolution can rapidly improve the activity of chimeric assembly-line enzymes. Proc Natl Acad Sci USA 104:11951.
76. Bajaj P, Sreenilayam G, Tyagi V, Fasan R (2016) Gram-Scale Synthesis of Chiral Cyclopropane-Containing Drugs and Drug Precursors with Engineered Myoglobin Catalysts Featuring Complementary Stereoselectivity. Angew Chem Int Ed Engl. 55:16110-16114.
77. Hernandez KE, Renata H, Lewis RD, Jennifer Kan SB, Zhang C, Forte J, Rozzell D, McIntosh JA, Arnold FH (2016) Highly Stereoselective Biocatalytic Synthesis of Key Cyclopropane Intermediate to Ticagrelor. ACS Catal. 6:7810-7813.
78. Jäckel C, Kast P, Hilvert D (2008) Protein design by directed evolution. Annu Rev Biophys 37:153-173.
79. Althoff EA, Wamg L, Jiang L, Giger L, Lassila JK, Wang Z, Smith M, Hari S, Kasp P, Hersclag D, Hilvert D, Baker D (2012) Robust design and optimization of retroaldol enzymes. Protein Sci 21:717-726.
80. Obexer R, Godina A, Garrabau X, Mitti PR, Baker D, Griffiths AD, Hilvert D (2017) Emergence of a catalytic tetrad during evolution of a highly active artificial aldolase. Nat Chem 9:50-56.
81. Smith GP (1985) Filamentous fusion phage: novel expression vectors that display cloned antigens on the virion surface. Science 228:1315-1317.
82. Parmley SF, Smith GP (1988) Antibody-selectable filamentous fd phage vectors: affinity purification of target genes. Gene 73:305-318.
83. de la Cruz VF, Lal AA, McCutchan TF (1988) Immunogenicity and epitope mapping of foreign sequences via genetically engineered filamentous phage. J Biol Chem 263:4318-4322.
84. Greenwood J, Willis AE, Perham RN (1991) Multiple display of foreign peptides on a filamentous bacteriophage. Peptides from Plasmodium falciparum circumsporozoite protein as antigens. J Mol Biol. 1991, 220:821-827.
85. Devlin JJ, Panganiban LC, Devlin PE (1990) Random peptide libraries: a source of specific protein-binding molecules. Science 249:404-406.
86. Scott JK, Smith GP (1990) Searching for peptide ligands with an epitope library. Science 249:386-390.
87. Cwirla SE, Peters EA, Barrett RW, Dower WJ (1990) Peptides on phage: a vast library of peptides for identifying ligands. Proc Natl Acad Sci USA 87:6378-6382.
88. McCafferty J, Griffiths AD, Winter G, Chiswell DJ (1990) Phage antibodies: filamentous phage displaying antibody variable domains. Nature 348:552-554.
89. Huston JS, Levinson D, Mudgett-Hunter M, Tai MS, Novotný J, Margolies MN, Ridge RJ, Bruccoleri RE, Haber E, Crea R (1988) Protein engineering of antibody binding sites: recovery of specific activity in an anti-digoxin single-chain Fv analogue produced in *Escherichia coli*. Proc Natl Acad Sci U S A 85:5879-5883.

90. Kang AS, Barbas CF, Janda KD, Benkovic SJ, Lerner RA (1991) Linkage of recognition and replication functions by assembling combinatorial antibody Fab libraries along phage surfaces. Proc Natl Acad Sci U S A 88:4363-4366.
91. Hoogenboom HR, Griffiths AD, Johnson KS, Chiswell DJ, Hudson P, Winter G (1991) Multi-subunit proteins on the surface of filamentous phage: methodologies for diplaying antibody (Fab heavy and light chains. Nucl. Acid. Res. 19:4133-4137.
92. Garrard LJ, Yang M, O'Connell MP, Kelley RF, Henner DJ (1991) Fab assembly and enrichment in a monovalent phage display system. Biotechnology 9:1373-1377.
93. Barbas CF, Kang AS, Lerner RA, Benkovic SJ (1991) Assembly of combinatorial antibody libraries on phage surfaces: the Gene III site. Proc Natl Acad Sci U S A. 88:7978-7982.
94. Burton DR, Barbas CF 3rd, Persson MA, Koenig S, Chanock RM, Lerner RA (1991) A large array of human monoclonal antibodies to type 1 human immunodeficiency virus from combinatorial libraries of asymptomatic seropositive individuals. Proc Natl Acad Sci U S A. 88:10134-10137.
95. Marks JD, Hoogenboom HR, Bonnert TP, McCafferty J, Griffiths AD, Winter G. (1991) By-passing immunization: Human antibodies from V-gene libraries displayed on phage. J Mol Biol 222:581-597.
96. Ward ES, Güssow D, Griffiths AD, Jones PT, Winter G (1989) Binding activities of a repertoire of single immunoglobulin variable domains secreted from Escherichia coli. Nature 341:544-546.
97. Huse WD, Sastry L, Iverson SA, Kang AS, Alting MM, Burton DR, Benkovic SJ, Lerner RA (1989) Generation of a large combinatorial library of the immunoglobulin repertoire in phage lambda. Science 246:1275-1281.
98. Smith GP (1988) Filamentous phage assembly: morphogenetically defective mutants that do no ill the host. Virology 157:156-165.
99. Hoogenboom HR, Winter G (1992) By-passing immunisation. Human antibodies from synthetic repertoires of germline VH gene segments rearranged in vitro. J Mol Biol 227:381-388.
100. Bass S, Greene R, Wells JA (1990) Hormone phage: an enrichment method for variant proteins with altered binding properties. Proteins 8:309-314.
101. Lowman HB, Bass SH, Simpson N, Wells JA (1991) Selecting high-affinity binding proteins by monovalent phage display. Biochemistry 30:10832-10838.
102. Lowman HB, Wells JA (1993) Affinity maturation of human growth hormone by monovalent phage display. J Mol Biol 5:564-568.
103. Mattheakis L, Bhatt R, Dower W (1994) An in vitro polysome display system for identifying ligands from very large peptide libraries. 91:9022-9026.
104. Hanes J, Plückthun A (1997) In viro selelction and evolution of functional proteins by using ribosome display. 94:4937-4942.
105. Georgiou G, Staphopolous C, Daugherty P, Nayak AR, Iverson BL, Curtiss R 3rd (1997) Display of heterologous proteins on the surface of microorganisms: rom the screening if combinatorial libraries to live recombinant vaccines. Nature Bioctech. 15:29-34.
106. Boder ET, Wittrup KD (1997) Yeast surface display for screening combinatorial polypeptide libraries. Nature Biotech. 15:553-557.
107. Malmborg AC, Borrebaeck CA (1993) BIAcore as a tool in antibody engineering. J Immunol Methods 183:7-13.
108. Kristensen P, Winter G (1998) Proteolytic selection for protein folding using filamentous bacteriophages. Folding Design 3:321-328.

109. Sieber V, Plückthun A, Schmid FX (1998) Selecting proteins with improved stability by a phage-based methods. Nat. Biotech. 16:955-960.
110. Matthews DJ, Wells JA (1993) Substrate phage: selection of protease substrates by monovalent phage display. Science 260:1113-1117.
111. Smith GP, Schultz DA, Ladbury JE (1993) A ribonuclease S-peptide antagonist discovered with a bacteriophage display library. Gene 128:37-42.
112. Linse S, Voorhies M, Norström E, Schultz DA (2000) An EF-hand display study of calmodulin subdomain pairing. J Mol Biol 296:473-486.