Alterations of substrate specificity and steroselectivity of lipases and esterases:

Enzymes are biocatalysts with numerous potential applications in industry and medicine. Compared to chemical catalysts, one of the most important advantages of a biocatalyst is its high selectivity, namely stereoselectivity, regioselectivity, and chemoselectivity. Such high selectivity is desirable in chemical synthesis as it may reduce or eliminate the use of protecting groups, minimize side reactions, simplify separation, and reduce environmental problems. Other advantages of a biocatalyst include high catalytic efficiency and mild operational conditions (pH, temperature, and pressure). Unfortunately, many naturally occurring enzymes are not suitable for industrial applications because of their poor selectivity, low stability in organic solvents, slow reaction rates, and substrate or product inhibition. To overcome these limitations, two complementary enzyme engineering approaches, rational design and directed evolution, have been developed over the past decades. Rational design involves site-specific alterations of selected residues in a protein to cause predicted changes in function, whereas directed evolution mimics the natural evolution process in the laboratory and involves repeated cycles of generating a library of protein variants and selecting the variants with desired properties. Largely due to its high success rate and general applicability, directed evolution has become the preferred engineering approach to generate tailor-made enzymes. The two key steps of directed evolution are generating molecular diversity and identifying improved variants through screening or selection. In order to generate molecular diversity, numerous methods of random mutagenesis and in vitro gene recombination have been developed. For random mutagenesis, error-prone PCR (epPCR) is the most convenient method, whereas DNA shuffling and staggered extension process (StEP) are the two most widely used in vitro gene recombination methods. In parallel to the development of these library creation methods, numerous high throughput screening or selection methods have been reported. Screening involves examining every variant individually for the desired enzyme property and usually relies on colorimetric or fluorogenic substrates or products. In contrast, selection links the survival of the host with the desired enzyme property. Hundreds of articles on enzyme engineering have been published between 2003 and 2009, and cannot be covered here. In this review, we focus on protein engineering of enzymesviz lipases and esterases for selective catalysis, including altering substrate specificity, altering substrate and product selectivity (enantioselectivity and regioselectivity).

LIPASES:

Lipases for applied biocatalysis

Lipases have received great attention as industrial biocatalysts in areas like oils and fats processing, detergents, baking, cheese making, surface cleaning, or fine chemistry. They can catalyse reactions of insoluble substrates at the lipid-water interface, preserving their catalytic activity in organic solvents. This makes of lipases powerful tools for catalysing not only hydrolysis, but also various reverse reactions such as esterification, transesterification, aminolysis, or thiotransesterifications in anhydrous

organic solvents. Moreover, lipases catalyse reactions with high specificity, regio and enantioselectivity, becoming the most used enzymes in synthetic organic chemistry. Therefore, they display important advantages over classical catalysts, as they can catalyse reactions with reduced side products, lowered waste treatment costs, and under mild temperature and pressure conditions. Accordingly, the use of lipases holds a great promise for green and economical process chemistry.

However, performance of a lipase is not always sufficient for an industrial application and most enzymes have sub-optimal properties for processing conditions. In fact, there are still disproportionally few examples of commercial scale applications of such biocatalysts in the manufacture of fine chemicals. In order to improve enzyme-mediated process efficiency, two different pathways can be followed: *i*) fitting the process to the available biocatalyst by medium engineering or modification of the manufacturing system to suit the sensitivities of the biocatalyst or *ii*) obtaining better biocatalysts through different strategies that can be run in parallel. These strategies (Figure 1) include the exploration of biodiversity to expand the sources and number of new biocatalysts, immobilization of existing enzymes, reaction conditions modification or the proper modification of these biocatalysts to get the most suitable variant for a defined industrial process. In this case the use of rational protein design to improve enzymes for which the 3D structure has been elucidated or homology-modelled, or the use of directed evolution can provide optimal biocatalysts.

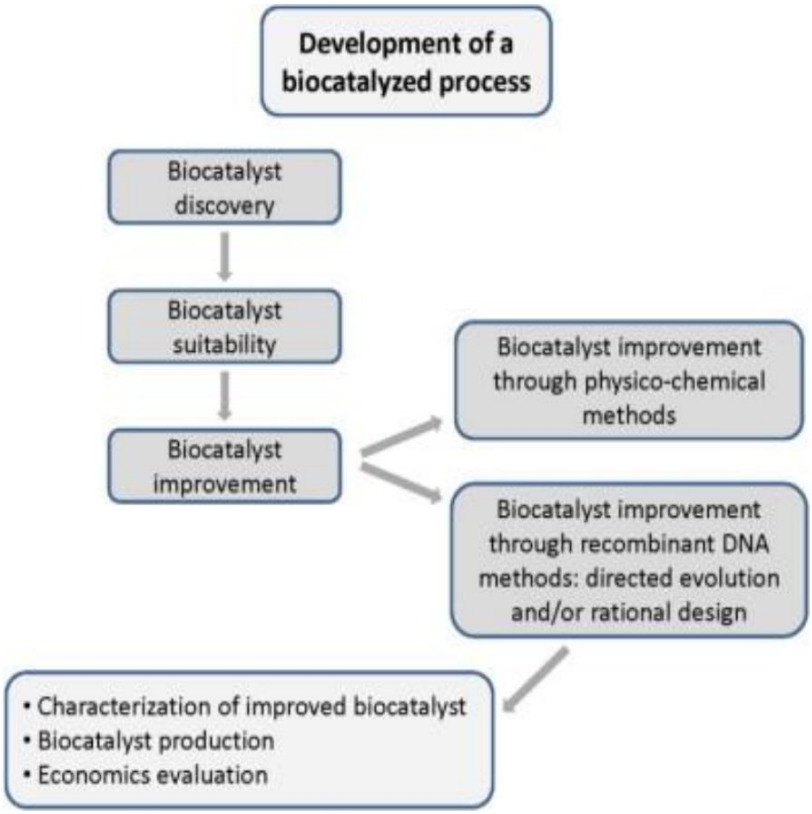


Figure 1. **Strategies that can be run to obtain better biocatalysts for commercial scale application.** Such strategies include the search for novel biocatalysts, and/or their improvement through immobilization or genetic modification to get the most suitable variant for each process using rational design or directed evolution.

Lipase improvement

Rational protein design requires both, the availability of the structure of the enzyme and knowledge about the relationship between sequence, structure and mechanism-function. If the structural data of the enzyme are not available, the structure of a homologous enzyme can be used as a model. All this information can be used to identify specific residues that can be mutated in order to improve a specific property, such as substrate specificity or thermal robustness. The selected residues are then targeted for site-directed mutagenesis, and the variant expressed, purified and analyzed for the desired property.

In contrast to rational protein design, directed evolution does not rely on a detailed understanding of the relationship between enzyme structure and function. It relies on the darwinian principles of mutation and selection. In general terms, directed evolution consists on repeated cycles of random mutagenesis (and/or gene recombination) of a target gene, coupled with selection or high-throughput screening for isolation of the functionally improved variants. In this case, enough diversity can be created in the starting gene so that an improvement in the desired property will be represented in a library of variants. Subsequently, screening or selection methods are used to identify these variants, and then they are used as a template for the next generation of mutagenesis and selection.

Recent Strategies

More recent developments have focused on making smaller libraries by using a combination of rational protein design and directed evolution procedures. These methods use information based on enzyme structure, and target specific residues or regions on the protein in each evolution cycle (Figure 2). Based on this approach, saturation mutagenesis has been extensively used in recent enzyme improvement protocols. This method refers to the randomization of all amino acids at a defined position or to the simultaneous randomization of two or more positions in an enzyme. In this case the sequence space modified is smaller and therefore, faster to screen. Following this concept, Iterative Saturation Mutagenesis (ISM) was introduced as a new and more efficient method for directed evolution of functional enzymes. Based on a cartesian view of the protein structure, with iterative cycles of saturation mutagenesis applied at rationally chosen sites of an enzyme, this approach drastically reduces the necessary molecular biological work and the screening effort. Both, rational protein design and directed evolution can be repeated or combined until the biocatalyst with the desired property is generated. Therefore, these protein engineering strategies have established as efficient tools to successfully improve biotechnologically relevant properties of enzymes. More sophisticated approaches of such procedures have been developed by Verenium Co. that provide high throughput screening methods for Gene Site Saturation Mutagenesis (GSSM) or for reassembly of the best properties of different genes through Tunable Gene Reassembly (TGR) technology. Moreover, in recent years, computational protein design is getting more and more attention as a novel strategy to predict the effects of mutations on protein structure, function or stability of libraries of enzyme variants generated by means of in silico approaches.

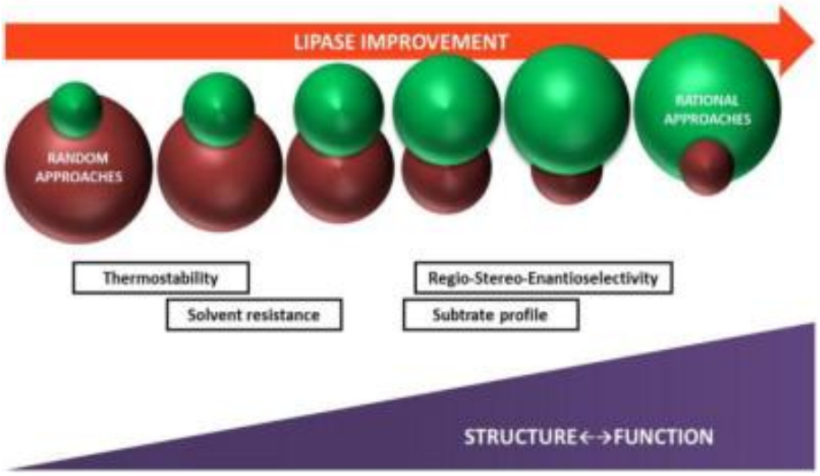


Figure 2. Correlation between structure-function knowledge and current protein engineering strategies to produce lipase variants with improved properties.

The basis for choosing the sites for modification depends on the nature of the catalytic property to be improved. During the so called third wave of biocatalysts engineering, a drastic reduction of work and effort can be achieved, as more focused libraries reduce the amount of screening and can allow more complex assays that more closely reflect the final application conditions. In this context, for many chemical manufacturing processes the target functions for directed evolution are multidimensional: substrate specificity, activity, stabilities (thermal, pH, hydrophobicity, additives), and other parameters can be improved in a concerted effort to meet the process requirements. Furthermore, when a library fails to contain any hits, non-improved or even inferior mutants could be used as templates in the continuation of the evolutionary pathway, thereby escaping from the local minimum.

To make a faster progress in protein engineering, a better understanding of how protein structure influences protein properties and a more critical evaluation of the many protein engineering approaches is needed. For this purpose, a set of tools have been designed for systematic analysis of the relationship between sequence, structure and function of lipases and related proteins: the Lipase Engineering Database (LED), the Data Warehouse System (DWARF) or the alpha/beta-Hydrolase Fold 3DM Database (ABHDB). These databases integrate sequence, structure and annotation information available in public databases like GenBank, PDB and others, being useful tools to identify functionally relevant residues apart from the active site amino acids, and for the design of variants

Regio-, stereo- and enantioselectivity

Lipase stereoselectivity is due to the various diastereomeric interactions that occur between the stereoisomers and the active site of the enzyme. It depends largely on the substrate structure, interaction at the active site, and on the reaction conditions. Although all lipases show the same mechanisms for ester hydrolysis, the structures of their binding sites and their stereopreferences are different, and their activity on defined substrates may vary. Taking advantage of the increasing structural data, molecular modelling strategies allow to study lipase-substrate interactions in each individual case at atomic level, proposing rational changes to promote the activity of one stereoisomer. The structure-based modification strategies

have revealed that subtle changes in the protein can lead to great improvements in the regio- stereo- or enantioselectivity. A very interesting strategy was followed to modify *Candida albicans* lipase A (CAL-A) stereoselectivity and generate a variant with improved activity against *trans* fatty acid esters. Since the identification of residues involved in *cis*/*trans* double bond configuration is challenging, the amino acids forming the long acyl-binding tunnel were randomly mutated by site-directed saturation mutagenesis to generate twelve NDT-codon-based libraries. This strategy produced two CAL-A variants active only against *trans* fatty acid esters. In another example, the molecular model of *Burkholderiacepacia* lipase docking the substrate (a bulky secondary alcohol) was used to detect two key residues interacting with the alcohol. Once detected, the residues were replaced by two sterically favourable amino acids, generating a mutant with excellent enantioselectivity for the kinetic resolution of bulky secondary alcohols.

On the other hand, randomness can also be introduced to selected residues. For instance, modelling a α-substituted *p*-nitrophenyl ester into the structure of CAL-A allowed identification of four residues whose side chains were pointing towards the substrate. Randomization of these residues led to generation of a CAL-A variant with optimum enantioselectivity for α-substituted *p*-nitrophenyl esters. Two related works performed using *Yarrowialipolytica* Lip2p for the hydrolysis of *2*-bromo-arylacetic acid esters showed that the enantioselectivity and enantiopreference of the enzyme is governed by two single residue positions. In a first work, a variant with improved *S*-selectivity was generated containing small amino acid residues at a defined position . This was enough to obtain a mutant with excellent enantioselectivity. On the other hand, bulky residues at that position and variants with single mutations in a second position led to a mutant with slightly reversed enantiopreference. Later on, single mutations at the two positions producing reverse enantiopreference were combined rendering a totally *R*-selective variant with great activity and excellent enantioselectivity.

If introducing changes in the active site is an excellent strategy for enantioselectivity improvement, modification of the amino acids surrounding the active site can also lead to enzyme improvement. Schließmann and collaborators generated a mutant variant of *P. fluorescens* esterase with enhanced enantioselectivity in the kinetic resolution of bulky primary and secondary alcohols by applying combinatorial mutagenesis to the bulky residues forming a bottleneck at the entrance of the active site. Substitution of such bulky residues by less bulky amino acids resulted not only in a variant with increased enantioselectivity but also with higher conversion rates.

Structural knowledge of target enzymes has allowed development of more radical strategies resulting in rearrangements of the active site structure. Thus, a mutant variant of CAL-A with high enantioselectivity in the kinetic resolution of ibuprofen esters was generated by applying a structure-based combinatorial mutagenesis strategy. Docking the substrate into CAL-A model structure allowed detection of nine residues surrounding the substrate pocket. A simultaneous combinatorial mutagenesis employing binary amino acid sets was applied on these residues, obtaining a CAL-A mutant which combines 5 mutations that reshaped the substrate pocket thus displaying an increased activity and enantioselectivity. A more drastic strategy was performed by Boersma and coworkers, who exchanged a loop near the active site entrance of *B. subtilis*LipA by loops from *F. solani* cutinase and *P. purpurogenum* acetylxylan esterase, obtaining LipA hybrids displaying inverted enantioselectivity in the kinetic resolution of 1,2-*O*-isopropylidene-*sn*-glycerol (IPG) esters. Moreover, the enantioselectivity of the most promising variant

(cutinase hybrid) was further evolved by directed evolution. Although the new mutant variants did not show synthetically useful enantioselectivity, the work provided novel perspectives on the evolution of lipases for increased properties.

When assays for enzyme improvement are coupled with extensive structure-function knowledge (Figure 2), new rational approaches for enantioselectivity optimization can arise. This was the case for several esterases aimed at the kinetic resolution of tertiary alcohols. Structure-function studies revealed that only esterases containing a GGG(A)X-type oxyanion hole can hydrolyse tertiary alcohol esters and how the amino acid composition of the oxyanion hole can greatly influence enantioselectivity. Thanks to the extensive structure-function data existing on this matter, it was possible to rationally transform esterase EstA from *Paenibacillusbarcinonensis*, an enzyme showing very low conversion rates towards tertiary alcohols, into a synthetically useful biocatalyst with excellent enantioselectivity, without the need for experimental structural data. A structure-based protein alignment guided the site-directed-mutagenesis of EstA oxyanion hole by introducing a motif previously associated with enantioselectivity, thus generating an EstA variant with excellent enantioselectivity.

Substrate scope

As for enantioselectivity improvement, lipase substrate specificity can be drastically modified with few changes in the protein sequence, a very clear example of such strategy being the work performed with the metagenomic esterase Enzyme R.34. Despite applying random mutagenesis (*ep*-PCR), only one amino acid change converted the esterase into a triacylglycerol lipase. Based on a homology model, the authors hypothesized that the mutated amino acid forms a salt bridge with another residue that causes a distortion of the enzyme structure, thus exposing the catalytic site to larger substrates.

Knowledge of protein structures has provided crucial information for substrate specificity modification in order to adapt lipases to industrial process conditions. Based on *C. rugosa* LIP2 crystal structure, two residues located in the substrate binding pocket were identified and considered for saturation mutagenesis to investigate the impact of these amino acids on substrate specificity. Two mutant variants of the same position showed a shifted specificity from short- to medium/long-chain length triglycerides, revealing that such position has a big impact on substrate specificity. Replacement of the former small side chain group allowed accommodation of medium- to long-chain triglycerides in the substrate binding site. A more accurate approach was used to rationally engineer*Candidarugosa* lipase (CRL) and CAL-A to obtain mutant variants with an altered substrate profile. In both cases a careful study of the structural conformation of the scissile fatty acid binding site guided the rational strategy. CRL acyl-binding tunnel was blocked at different points by introducing bulky amino acids, thus generating mutant variants with different chain length specificity. In a similar way, the main acyl-binding tunnel of CAL-A was blocked in order to produce a small alternative pocket showing increased specificity for medium chain length fatty acids. The generation of CAL-A and CRL variants with altered substrate specificity confirmed the usefulness of the rational approaches.

A remarkable work by Juhl and co-workers took full advantage of bioinformatics to generate a *C. antarctica* lipase B (CAL-B) variant accepting esters with branched and sterically demanding acids. An *in silico* library of 2400 CAL-B variants was built and screened *in silico* by substrate-imprinted docking. From

the virtual screening, nine variants with single amino acid exchanges and increased activity were predicted and generated by site directed mutagenesis. Among the nine predicted variants only one displayed higher activity than the wild type against branched acids but not against sterically demanding acids. This work shows the potential of *in silico* approaches to predict mutant variants, thus reducing the time-consuming high throughput screening assays.

Standing drawbacks

Despite the success of a large number of enzyme improvement studies, some designed modifications have failed to produce the lipase variant with the desired property. That was the case for *P. aeruginosa* LipA, a foldase-dependent enzyme for which obtaining an enzymatically active form in the absence of its chaperone would provide an easy and effective method for expression of such industrially interesting enzyme. Although it was described that a single amino acid substitution (P112Q) in *Pseudomonas* sp. KFCC LipK allowed the lipase to spontaneously overcome the energy barrier of folding without participation of its cognate foldase, attempts to overcome this dependence in *P. aeruginosa* LipA failed to produce the desired variant. Substitution of amino acid 112 by ISM plus random mutagenesis was applied to LipA and the released mutants screened for activity in the absence of foldase. Unfortunately, no foldase-independent LipA variant was obtained, thus remaining yet a challenge for future strategy developments. To make a faster progress in protein engineering, a better understanding of how protein structure influences protein properties and a more critical evaluation of the many protein engineering approaches would be required.

Co-factors/substrate engineering for metabolic engineering

Co-factor engineering

Cofactor engineering, a subset of metabolic engineering, is defined as the manipulation of the use of cofactors in an organism's metabolic pathways. In cofactor engineering, the concentrations of cofactors are changed in order to maximize or minimize metabolic fluxes. This type of engineering can be used to optimize the production of a metaboliteproduct or to increase the efficiency of a metabolic network. The use of engineering single celled organisms to create lucrative chemicals from cheap raw materials is growing, and cofactor engineering can play a crucial role in maximizing production. The field has gained more popularity in the past decade and has several practical applications in chemical manufacturing, bioengineering and pharmaceutical industries.

Cofactors are non-protein compounds that bind to proteins and are required for the proteins normal catalytic functionality. Cofactors can be considered "helper molecules" in biological activity, and often affect the functionality of enzymes. Cofactors can be both organic and inorganic compounds. Some examples of inorganic cofactors are iron or magnesium, and some examples of organic cofactors include ATP or coenzyme A. Organic cofactors are more specifically known as coenzymes, and many enzymes require the addition of coenzymes to assume normal catalytic function in a metabolic reaction. The coenzymes bind to the active site of an enzyme to promote catalysis. By engineering cofactors and

coenzymes, a naturally occurring metabolic reaction can be manipulated to optimize the output of a metabolic network.

Background

Cofactors were discovered by Arthur Harden and William Young in 1906, when they found that the rate of alcoholic fermentation in unboiledyeast extracts increased when boiled yeast extract was added. A few years after, Hans von Euler-Chelpin identified the cofactor in the boiled extract as NAD^+. Other cofactors, such as ATP and coenzyme A, were discovered later in the 1900s. The mechanism of cofactor activity was discovered when, Otto Heinrich Warburg determined in 1936 that NAD^+ functioned as an electron acceptor. Well after these initial discoveries, scientists began to realize that the manipulation of cofactor concentrations could be used as tools for the improvement of metabolic pathways.

An important group of organic cofactors is the family of molecules referred to as vitamins. Vitamin B12 (cobalamin), for example, plays a crucial role in the human body, whilecoenzyme B12, its derivative, is found in the metabolisms of every type of cell in our bodies. Its presence affects the synthesis and regulation of cellular DNA as well as taking part in fatty acid synthesis and energy production. Cofactors are required by many important metabolic pathways, and it is possible for the concentrations of a single type of cofactor to affect the fluxes of many different pathways

Minerals and metallic ions that organisms uptake through their diet provide prime examples of inorganic cofactors. For instance Zn^{2+} is needed to assist the enzyme carbonic anhydrase as it converts carbon dioxide and water to bicarbonate and protons. A widely recognized mineral that acts as a cofactor is iron, which is essential for the proper function of hemoglobin, the oxygen transporting protein found in red blood cells. This example in particular highlights the importance of cofactors in animal metabolism.

Significance

Cofactor engineering is significant in the manipulation of metabolic pathways. A metabolic pathway is a series of chemical reactions that occur in an organism. Metabolic engineering is the subject of altering the fluxes within a metabolic pathway. In metabolic engineering, a metabolic pathway can be directly altered by changing the functionality of the enzymes involved in the pathway. Cofactor engineering, offers a distinct approach, and some advantages, to altering a metabolic pathway. Instead of changing the enzymes used in a pathway, the cofactors can be changed. This may give metabolic engineers an advantage due to certain properties of cofactors and how they can be modified.

Metabolic pathways can be used by metabolic engineers to create a desired product. By modifying the types of cofactors used and the times at which they are used, the outcome of the metabolic network can change. To create a greater production of a product, metabolic engineers have the ability to supply the network with whichever cofactor is best suited for that specific process. This leads to the optimization of networks to give a higher production of desired products. Also, changing the cofactors used in a network may be the an ingenious solution to a complicated problem. A network that is present in the cell, but is

often unused, may have a desirable product. Instead of engineering a completely new set of pathways to produce the product, cofactor engineering can be applied. By replacing enzymes to use cofactors readily available in a cell, the typically unused network is no longer cofactor-limited, and production may be increased.

In addition to modifying the yield of metabolic networks, changing the cofactors used in a network can reduce operation costs when trying to form a desired product. NADH and NADPH are two extremely common cellular cofactors, differing only by the presence of a phosphate group. However, this phosphate group makes NADPH much less stable than NADH, and therefore more expensive to synthesize. Thus, it is advantageous to try and use NADH is some cellular networks because it is often cheaper, more readily available, and accomplishes the same task as NADPH.

Tools and processes

Cofactor engineering most often deals with the manipulation of microorganisms such as *Saccharomyces cerevisiae* and *Escherichia coli*, and as such requires the use ofrecombinant DNA techniques. These techniques utilize small circular segments of DNA called plasmids, which can be introduced and incorporated by microorganisms such as*Escherichia coli*. These plasmids are specifically designed in labs to be easily incorporated, and affect the expression of various protein, metabolites and enzymes. For instance, a particular plasmid may cause a change in an enzyme's amino acid sequence, which could increase its affinity for a particular substrate.

Microorganisms require a medium to grow in, and one commonly used for cultures of *Escherichia coli* is Luria-Bertani (LB) broth. This medium is often supplemented with glucoseand will often contain additional molecules designed to facilitate optimal culture growth. Pre-cultures may then be grown in shake flasks. These are simply plugged Erlenmeyer flasks which are left on an orbital shaker machine, which revolves at very high RPM. This process aerates the culture, which is necessary for optimal growth. Once the pre-cultures are ready, the plasmids needed by specific experiments are added to each culture separately, and then each culture is transferred to a bio-reactor. Bio-reactors are systems which allow cultures to grow in a controlled environment. This leaves the introduced plasmids as the only independent variable. The required temperature, pH, metabolite concentrations, and various other environmental factors can be maintained by the bio-reactor ensuring identical growth conditions for each culture.

Once samples are allowed to grow in the reactor for a specified period, they can be removed and studied to determine whether the intended alterations to the organism are evident. Since cofactor engineering most often deals with metabolic pathways, these organisms are often studied but introducing specific tagged fluorescent metabolites and documenting their progression through various pathways. In other cases results are more obvious and easily observable, such as with the decreased ethanol production of yeastreferred to below.

Applications

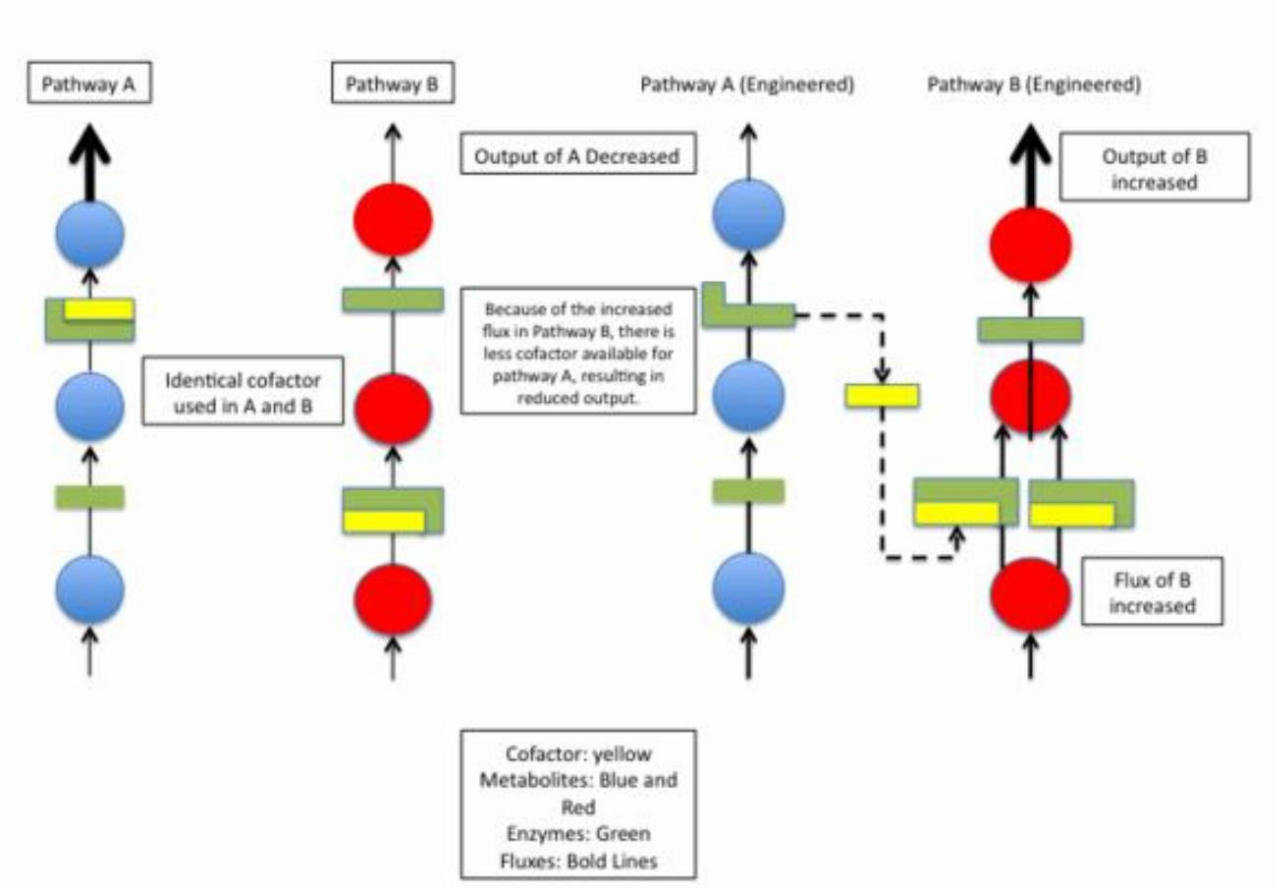
Changing an Enzyme's Cofactor From NADPH to NADH

Biocatalysts are required for the production of chiral building blocks needed in pharmaceuticals and other chemicals used by society. Many such biocatalysts require NADPH as acofactor. NADPH, a cofactor quite similar to NADH, is both more expensive and less stable than its counterpart NADH. For these reasons, manufacturers would prefer that the biocatalysts they use in their production lines accept NADH over NADPH. Cofactor engineering has recently been successful in altering enzymes to prefer NADH as a cofactor instead of NADPH. In 2010, a group of scientists performed cofactor engineering on the enzyme Gre2p, an NADPH-preferring dehydrogenase found in *Saccharomyces cerevisiae*. Gre2p reduces the compound diketone 2,5-hexanedione into the chiral building blocks (5S)-hydroxy-2-hexanone and (2S,5S)-hexanediol. The scientists determined that Asn9 (Asparagine, position 9) was an important amino acid the active site of Gre2p. Specifically, Asn9 binds to the 3'-hydroxyl group and the 2'-oxygen atom of adenylribosemoiety. Through direct mutagenesis, the scientists exchanged the Asn9 for both Asp (Aspartic Acid) and Glu (Glutamic Acid). This change caused Gre2p to have a decreased dependency on NADPH, and an increased affinity for NADH. This resulted in increased Gre2p activity when using NADH. It was observed that substituting Asn9 with Glu produced a greater effect than changing Asn9 to Asp. Asn contains a polar uncharged side chain, while both Asp and Glu contain a polar charged side chain. The increased effect of Glu is caused by the extra carbon in its side chain that brings it closer to adenyl ribose moiety. This allows for stronger hydrogen bonding between the 2'- and 3'-ribosehydroxyl groups and the side chain carboxyl group. The maximum velocity of the reaction doubled, while using NADH, when Asn9 was substituted with Glu. With these results, the scientists successfully engineered Gre2p to prefer NADH over NADPH and increased the speed of 2,5-hexanedione reduction. This will allow chemical companies to decrease their manufacturing costs by using NADH instead of NADPH at least for this particular reduction.

Changing a Network's Cofactor Preference

An alternative example of changing an enzyme's preference for cofactors is to change NADH dependent reaction to NADPH dependent reactions. In this example, the enzymes themselves are not changed, but instead different enzymes are selected that accomplish the same reaction with the use of a different cofactor. An engineered pathway was created to make 1-butanol from Acetyl-CoA by changing enzymes in the metabolic pathway of S. elongatus. The Clostridium genus is known to produce 1-butanol, providing a pathway that could be inserted in S. elongatus. This pathway synthesizes 1-butanol using the reverse β-oxidation pathway. The enzymes involved in this newly engineered pathway were NADH specific, which was problematic for replicating the pathway in S. elongatus as cyanobacteria produce much more NADPH than NADH.

The research group then identified enzymes that utilize NADPH or both NADPH and NADH by bioprospecting. Acetoacetyl-CoA reductase (PhaB) was found to be a suitable replacement for hydroxybutyric dehydrogenase (Hbd). To replace AdhE2, the researchers found that NADP-dependent alcohol dehydrogenase (YqhD) from E. coli to be effective for the pathway. Furthermore, the researchers needed a dehydrogenase to replace the aldehyde dehydrogenase capacity of AdhE2. CoA-acylatingbutyraldehyde dehydrogenase (Bldh) from C. saccharoperbutylacetonicum was found to be a good suit. Together, PhaB, Bldh, YqhD can replace Hbd and AdhE2, respectively, to change the cofactor preference of 3-ketobutyryl-CoA reduction from using NADH to using NADPH. The authors then constructed various combinations of the different enzymes (of those found in the reverse oxidation pathway and the NADPH utilizing enzymes) by overexpressing different genes in cultures of S. elongatus PCC 7942. In order to do so, they constructed plasmids containing the genes corresponding to the enzymes and combined them into the genome of S. elongates. After enzyme assays, the strain of cyanobacteria expressing the NADPH utilizing enzymes produced the greatest amount of 1-butanol (29.9 mg/L), exceeding that of strains that did not consist of the NADPH utilizing enzymes by four times. Overall, 1-butanol was produced in S. elongatus using a pathway from another organism. This pathway was modified in order to match the preferred reducing cofactor for the cyanobacteria



In cofactor engineering, a metabolic pathway is altered by changing the concentrations of specific cofactors that are produced either in that particular pathway or in a separate pathway. For example, an hypothetical organism could have two arbitrary pathways called A and B where some enzymes in both A and B utilize the same cofactors. If scientists wanted to decrease the output of pathway A, they may first consider directly engineering the enzymes involved in A, perhaps to decrease a particular active site's affinity for its substrate. In some cases however, the enzymes in A may be difficult to engineer for various reasons, or it may be impossible to engineer them without dangerously affecting some third metabolic pathway C, which utilizes the same enzymes. As a separate option, scientists could increase the flux of B, which may be easier to engineer. This in turn could "tie up" the cofactors needed by A, which would slow enzymatic activity, decreasing output in A. This is one hypothetical example of how cofactor engineering can be used, but there are many other unique cases where scientists use cofactors as a way of altering metabolic pathways. A major advantage to cofactor engineering is that scientists can use it to successfully alter metabolic pathways that are difficult to engineer by means of ordinary metabolic engineering. This is achieved by targeting more easily engineered enzymes in separate pathways, which use the same cofactors. Since many cofactors are used by different enzymes in multiple pathways, cofactor engineering may be an efficient, cost effective alternative to current methods of metabolic engineering

Yeast are commonly used in the beer and wine industry because they are capable of efficiently producing ethanol through the metabolic pathway fermentation in the absence of oxygen. Fermentation requires the enzyme glycerol-3-phosphate dehydrogenase (GPDH) which depends on the cofactor NADH. This pathway involves the conversion of glucoseto both ethanol and glycerol, both of which use NADH as a cofactor. Scientists engineered *Saccharomyces cerevisiae* to overproduce GPDH, which shifted the cells metabolic fluxaway from ethanol and toward glycerol, by limiting NADH availability in the ethanol production portion of the pathway. The opposite effect was achieved by influencing a separate pathway in the cell, the Glutamate Synthesis pathway. Inactivating the expression of the enzyme glutamate dehydrogenase, which is NADPH dependent, and over expressing the enzymes glutamine synthetase and glutamate synthetase, which rely on NADH as a cofactor shifted the cofactor balance in glutamate synthesis pathway. The pathway is now dependent on NADH rather than NADPH, which decreases NADH availability in the fermentation pathway. This in turn causes increased ethanol production and decreasedglycerol production. This method of manipulating metabolic fluxes could be visualized much like global fuel markets, where the increased production of ethanol for use in the automotive industry would decrease its availability in the food industry. Essentially, producing more engines which run on ethanol could result in decreased consumption of processed sweets, which contain high fructose corn syrup. This engineering of cofactors is applicable to the beer and wine industry since it allows for the regulation of ethanol levels in alcoholic beverages. Advancements in the wine industry have caused a steady increase in ethanol content, so winemakers in particular would be interested in the possibility of reducing the ethanol levels of some of their wines.

Citric acid cycle

Coenzyme A (CoA) and acetyl-CoA are two intermediate metabolites, most notably found in the Citric Acid Cycle, which participate in over 100 different reactions in themetabolism of microorganisms. Recent experiments have shown that over expression of the enzyme pantothenate kinase and supplementation of pantothenic acid in the CoA biosynthesis pathway have allowed adjustments of both CoA and acetyl-CoA fluxes. This increased concentration of cofactors resulted in an increased carbon flux in the isoamyl acetate synthesis pathway, increase the production efficiency of isoamyl acetate. Isoamyl acetate is used industrially for artificial flavoring and for testing the effectiveness ofrespirators. In addition to the production of isoamyl acetate, the manipulation of CoA biosynthesis during the pyruvate hydrogenase reaction also causes an increase in the production of both succinate and lycopene, each of which have beneficial effects on the human body. An increase in succinate concentration, which is used as a catalyst, may lead to an increase in the speed of the Citric Acid Cycle, and consequently an individual's metabolism. Increasing lycopene concentrations, has been shown to decrease the risk of prostate cancer. The potential rewards of repeating such feat of cofactor engineering and successfully incorporating them into industry practices are innumerable.

Paper manufacturing

Many important industrial enzymes use cofactors to catalyze reactions. By using cofactors to manipulate metabolic pathways, it is possible to reduce material cost, eliminate steps in production, reduce production time, decrease pollution, and increase overall production efficiency. One case that demonstrates several of these manufacturing benefits involves the genetic engineering of aspen trees. In the paper production process, manufacturing plants must break down lignin, a biochemical compound that gives a tree trunk its stiffness, in order to form the pulp used throughout the rest of production. The chemical pulping process requires the manufacturing plant to use a significant amount of energy, as well as many expensive and toxic chemicals. A group of genetic engineers, through **cofactor engineering**, engineered a genetically superior aspen tree that produced less lignin. These genetically engineered trees have allowed for paper mills to reduce their costs, pollution, and manufacturing time.

Co-factors/substrate engineering for gene therapy

Gene therapy isdefined as the introduction of genetic modifications into target cells in order to cure a disease or to treat its symptoms. Potentially, such as procedure should have a significantinfluence on human health, and indeed new treatments for many inherited and acquired diseases are promised. Redesigned homing endonucleases with defined DNA sequence specifi city represent potential agents for gene therapy . The most promising means of creating a highly specifi c endonuclease would be to use homing endonucleases as scaffolds and then to alter their specifi city, either by introducing mutations at the active site or through a chimeric protein approach. An alternative here is zinc - fi nger nucleases ; these are constructed by linking a zinc - fi nger protein, which is capable of recognizing a target DNA sequence, and a nonspecifi c DNA cleavage domain of Fok I, a Type IIS restriction enzyme . The use of engineered

zinc - fi nger proteins with novel sequence specifi city may permit the creation of a wide variety of custom - made zinc fi nger nucleases to target different diseases. Small - molecule - regulated gene expression systems (so - called ‘ gene switches ’) may also be useful for gene therapy, as they may be used to control target gene expression by regulating when, how, and to what extent, a gene will become expressed . Recently, an orthogonal ligand – receptor pair was created for use in gene switches by combining rational design and directed evolution. A similar method could be used to create other receptors that are inducible by small- molecule drugs. Further in vivo testing would be required in order to monitor the efficacy and toxicity associated with homing endonucleases and gene switches. However, both methods may be effective in treating single -g ene diseases such as sickle cell anemia, cystic fi brosis, hemophilia and muscular dystrophy.

***Combinational manipulation of polyketides and non ribosomal peptides**

Non-ribosomal peptides are small peptides synthesised mainly by bacteria and fungi. Despite their small size, they are highly diverse in terms of the monomers that can be incorporated. According to the most recent published reports there are 1,164 different non-ribosomal peptides known , which collectively contain over 500 unique monomers, including both proteinogenic and non-proteinogenic L- and D-amino acids, as well as carboxylic acids and amines. Non-ribosomal peptides also exhibit high structural diversity with only 27 % being linear; the remainder having cyclic, branched or other complex primary structures

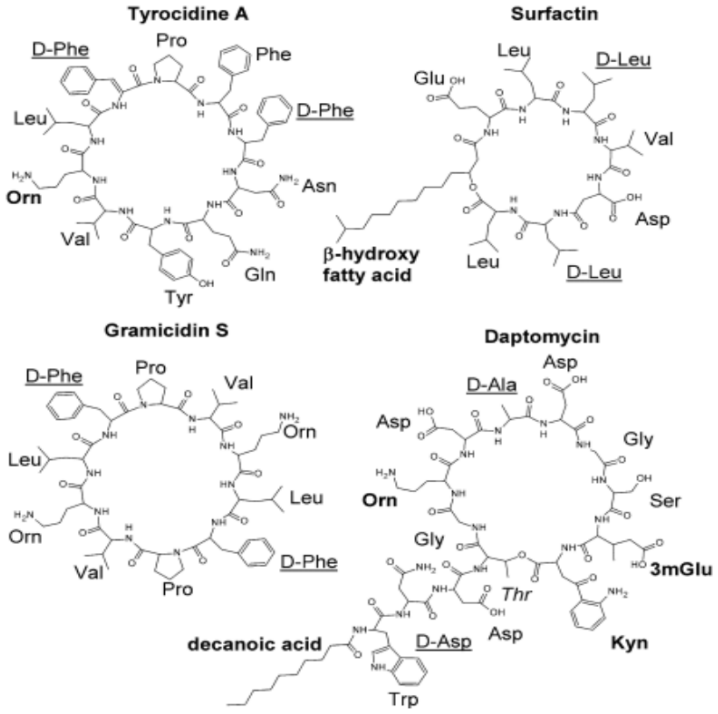


Fig. 1 Structures of some non-ribosomal peptides relevant to biotechnology that are highlighted in this review. Non-proteinogenic amino acids or substituents are labelled in bold, D-isomers are underlined, and a threonine residue in daptomycin that is involved in an atypical ester bond via the side-chain hydroxyl is labelled in italics. Orn, ornithine; 3mGlu, 3-methylglutamate; Kyn, kynurenine; standard three letter abbreviations are used for proteinogenic amino acids

The diversity of non-ribosomal peptides imparts to them many properties of relevance to biotechnology; for example, peptides have been identified with antibiotic, antiviral, anti-cancer, anti-inflammatory, immunosuppressant and surfactant qualities. Importantly for medicine, natural products often need to be modified to improve clinical properties and/or bypass resistance mechanisms. Due to their typically complex structures, most clinical natural product derivatives are created by means of semisynthesis; a process whereby the natural product is chemically modified post-isolation from biological sources.

The multiple template model of non-ribosomal peptide synthesis

Non-ribosomal peptide synthesis generally follows the multiple template model. According to this model, peptides are synthesised in a modular assembly line-like manner by NRPS enzymes (''the template''). The modules that comprise an NRPS template may be clustered on a single enzyme or located within multiple distinct enzymes that associate post-translation; and are classified as either initiation, elongation, or termination modules depending on their location in the assembly line (Fig. 2a). Modules act in a concerted but semiautonomous fashion, and are defined by their ability to recognise, activate and incorporate a specific monomer into the final peptide product.

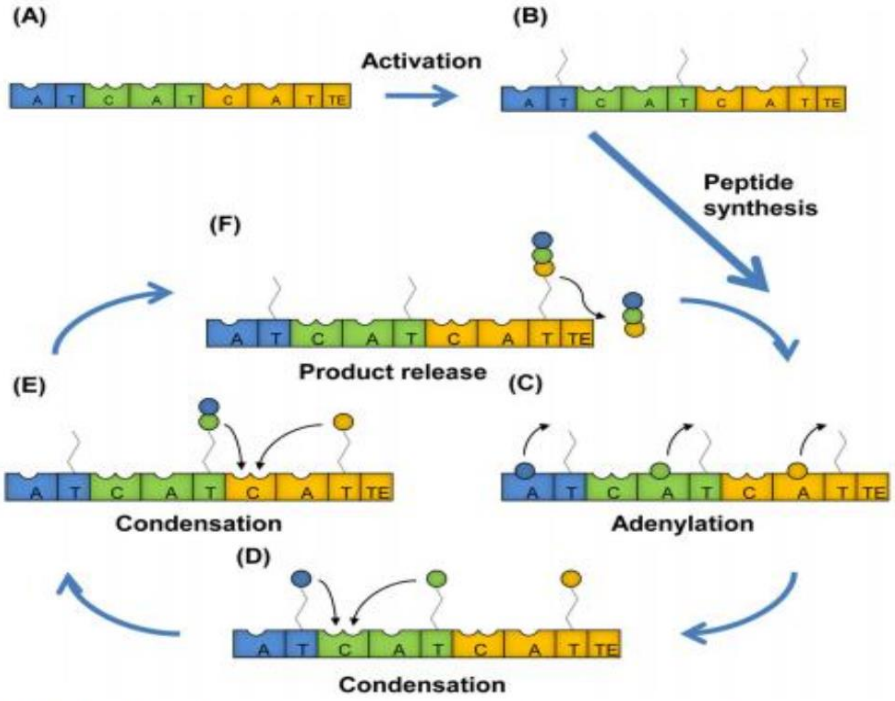


Fig. 2 The multiple template model of non-ribosomal peptide synthesis. **a** A schematic arrangement of domains within a hypothetical three module NRPS that contains an initiation (*blue*), an elongation (*green*) and a termination (*orange*) module. There may be multiple elongation modules present within a single NRPS template. **b** Activation of NRPS modules by post-translational attachment of a 4′-phosphopantetheine (PPT) cofactor to each T domain. **c** Domains within a module act in a semi-autonomous fashion, beginning with the A domain, which activates and tethers a specific monomer substrate to the PPT prosthetic group attached to the T domain immediately downstream. **d** The substrate linked to each T domain is then passed to a C domain, which catalyses peptide bond formation between the donor substrate provided by the upstream module, and the acceptor substrate from the downstream module. C domains are generally located immediately upstream of A domains and C-A-T domain units comprise a basic elongation module. **e** Peptide bond formation hydrolyses the upstream thioester bond, yielding a peptide that is now attached to the downstream T domain, and which goes on to serve as the donor substrate at the C domain of the following module. **f** After addition of the final monomer by the termination module, the peptide product is released by a thioesterase (TE) domain; most commonly by intra-molecular cyclisation to yield a macrocyclic lactone or lactam, or by hydrolysis to yield a linear peptide product. Following product release, the NRPS is returned to step (**c**) and peptide synthesis can repeat in an iterative fashion

Within each module, an adenylation (A) domain recognises and activates a specific substrate by addition of AMP (Fig. 2c). The activated substrate is then tethered to a flexible 40 -phosphopantetheine (PPT) prosthetic group, which is itself covalently attached to a thiolation (T) domain (also known as a peptidyl carrier protein (PCP) domain) (Fig. 2b). The T domain lies at the heart of the biosynthetic process, with its flexible PPT prosthesis effectively the ''swinging arm'' of a biomolecular assembly line that transfers peptide intermediates between different domains and modules. Post-attachment of an activated substrate by its A domain partner, a T domain then passes that substrate to a condensation (C) domain, which catalyses peptide bond formation between the donor substrate provided by the T domain immediately

upstream, and the acceptor substrate provided by the downstream T domain (Fig. 2d). Following the initial condensation event, the process can repeat in an iterative fashion, with the previous peptide intermediate now serving as the donor substrate for the C domain of the next module in an NRPS complex (Fig. 2e). Along the way, certain modules may contain additional tailoring domains that modify individual substrates in a directed fashion (e.g., epimerisation (E) domains, for conversion from L- to D-enantiomers). The growing peptide continues to be passed from the T domain of one module to the T domain of the next until the product is released, typically via a hydrolysis or intramolecular cyclisation reaction catalysed by a thioesterase (TE) domain associated with the final module in an NRPS complex (Fig. 2f).

Strategies to create novel peptide products via genetic manipulation of NRPS templates

The modular structure of the NRPS assembly line suggests that it should be possible to rationally alter one or more residues in a non-ribosomal peptide product by substitution or engineering of the module(s) that specify the target residue(s). In nature, the diversity of non-ribosomal peptides is thought to have arisen from point mutation, substitution of domains or modules for alternatives that specify different substrates, and/or the insertion/deletion of modules. The modular structure of the NRPS assembly line suggests that it should be possible to rationally alter one or more residues in a non-ribosomal peptide product by substitution or engineering of the module(s) that specify the target residue(s). In nature, the diversity of non-ribosomal peptides is thought to have arisen from point mutation, substitution of domains or modules for alternatives that specify different substrates, and/or the insertion/deletion of modules.

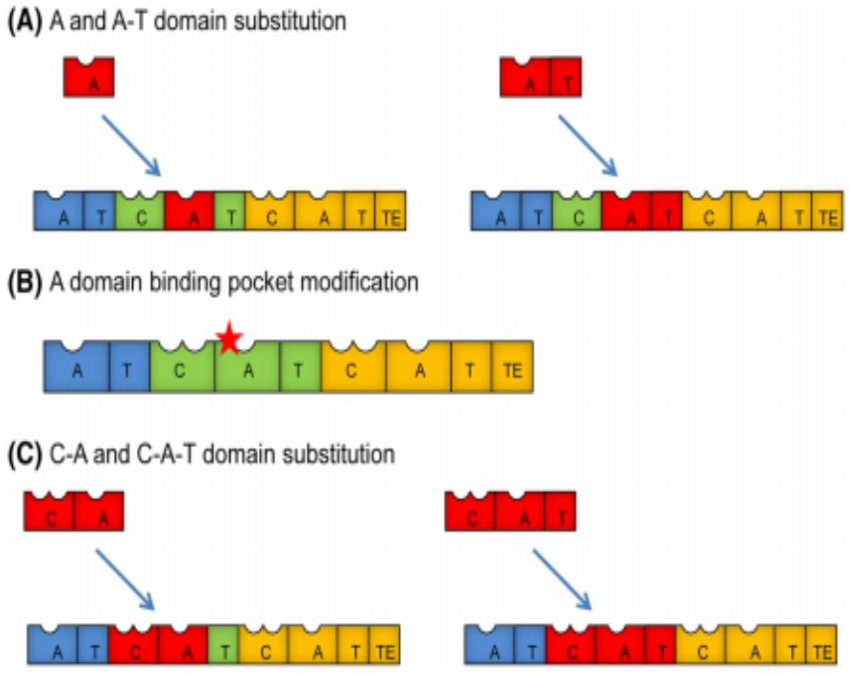


Fig. 3 The three main strategies employed to re-engineer NRPS templates at a genetic level. **a** Substitution of the substrate-specifying A domain, with or without its native T domain partner. **b** Modulation of A domain substrate specificity by site-directed mutagenesis, to avoid major perturbation of tertiary and quaternary structure. **c** Strategies that treat C and A domains as inseparable pairs, to address the phenomenon of strong C domain substrate specificity at the acceptor site

Creation of novel peptide products via A domain substitution

Due to their similarity to A domain substitutions, paired A-T domain substitutions will also be considered in this section (Fig. 3a). The first reported efforts to alter the products of NRPS enzymes by domain substitution targeted a leucine-specifying A-T domain pair, in the termination module of the Bacillus subtilissurfactin NRPS template SrfA-C.

Non-ribosomal peptide synthetases (NRPS) are large modular enzymes that govern the synthesis of numerous biotechnologically relevant products. Their mode of action is frequently compared to an assembly line, in which each module acts in a semi-autonomous but coordinated manner to add a specific monomer to a growing peptide chain, unfettered by ribosomal constraints. The modular nature of these systems offers tantalising prospects for synthetic biology, wherein the assembly line is re-engineered at a genetic level to generate a specific or combinatorial modified product. However, despite some success stories, a "one size fits all" approach to NRPS synthetic biology remains elusive. This review examines both rational and random mutagenesis strategies that have been employed to modify NRPS function, in an

attempt to highlight key points that should be considered when seeking to reengineer an NRPS biosynthetic template.

Structure based engineering of PHA synthase enzyme and monomer supplying enzymes

Three main issues have hindered widespread use of PHAs: (1) the high production cost compared to petroleum-based polymers with similar properties; (2) the inability to produce high-performance PHAs in substantial amounts; and (3) the difficulty in controlling the life cycle of PHAs, i.e., the control of their biodegradability and their effective chemical recycling.

To solve the former two issues, we have focused on the genetic engineering of PHAs metabolism, which will lead to the cost-effective biological production of PHAs and the improvement of their properties, such as molecular mass and monomer composition. In particular, protein engineering of PHA synthase can improve both PHA production efficiency and the properties of the generated polymer because PHA synthase plays a central role in PHA biosynthesis . Here we would like to highlight the current special topic on the biosynthesis of new PHA polymers incorporating unusual monomer units such as LA by PHA synthase engineering. Further, gene cloning and expression in plants has created new possibilities of using photosynthesis to convert atmospheric CO_2 directly into PHA, in hopes of reducing production cost in the future.

In addition, to solve the latter issue, we have also focused on the engineering of PHB depolymerases. PHB is the most common form of PHAs. In natural environment, the microbial and enzymatic degradation of PHB is an important first step in the PHB recycling process. However, PHB degradation depends on the surrounding conditions and proceeds on the order of a few months in anaerobic sewage or a few years in seawater. Such PHB degradation process is undesirable from the standpoint of the efficient use of biomass resources. To overcome this issue, chemical recycling using

spent PHB materials as recyclable monomer-concentrated resources is rapidly gaining importance due to its high degradation rate. In addition, as chemical recycling is cost-efficient and has low CO_2 emissions, it has great potential as a low-cost and environmentally compatible process. PHB monomerization, the first step in chemical recycling, is currently carried out via a thermal decomposition process. However, this chemical recycling method presents some drawbacks, such as racemization of the decomposed products, high reaction temperature, and contamination with residual metal catalysts. As one of the solutions, the development of alternative PHB monomerization methods that use such enzymes as PHB depolymerases is highly awaited because those methods do not produce undesirable byproducts, have high enantio- and regioselectivities, and can be performed at moderate temperatures. Moreover, as the efficient use of biocatalysts requires suitable enzymes with high activity and stability under process conditions, the desired substrate selectivity, and high enantioselectivity, the improvement of PHB depolymerases is expected to result in the construction of an effective PHB chemical recycling system. In this chapter we will also provide some case studies on protein engineering of PHB depolymerase based on domain structure-based and random mutagenesis approaches.

Protein engineering of PHA synthases

BIOCHEMICAL PROPERTIES AND ENGINEERING CONCEPTS OF PHA SYNTHASES

PHA synthases catalyze the polymerization reaction of hydroxyalkanoate (HA) to polymer PHA. The monomer substrates of PHA synthase are mainly 3HA-CoAs with various side-chain lengths, and only*R*-enantiomer HA-CoAs are accepted for polymerization by synthase. Over 60 different PHA synthases have been classified into four types based on their substrate specificities and subunit compositions of enzymes.

Type I and type II PHA synthases consist of single subunits (PhaC). Type I PHA synthases, represented by *Ralstoniaeutropha* enzyme, mainly polymerize SCL-monomers (C3–C5), whereas type II PHA synthases, represented by *Pseudomonas oleovorans* enzyme, polymerize MCL-monomers (C6–C20). Type III PHA synthases, represented by *Allochromatiumvinosum* enzyme, consist of two hetero-subunits (PhaC and PhaE). PhaC subunits of type III synthase are smaller than those of type I and II synthases, but possess catalytic residues. Like the type I synthases, these PHA synthases prefer to polymerize SCL-monomers (C3–C5). Type IV PHA synthases, represented by *Bacillus megaterium*, are similar to the type III PHA synthases with respect to possessing two subunits. However, unlike the PhaE of type III PHA synthases, a smaller protein designated as PhaR is required for full activity expression of type IV PhaC.

The lack of a suitable structural model for any PHA synthase has limited attempts to improve the activity and to alter the substrate specificity of these enzymes in “irrational” manners, such as random mutagenesis and gene shuffling . Generally, natural diversity provides us with attractive starting materials for artificial evolution as it represents functionalized sequence spaces to some extent. A tremendous

population (over 60 species) of randomly screened PHA producing bacteria suggests that attractive prototype enzymes for molecular breeding would exist. Among them, enzyme evolution approach has been applied to the following type I and type II PHA synthases derived from some bacteria.

Reference

1. THE BACKGROUND TO ARTHUR HARDEN'S DISCOVERY OF COZYMASE - ROBERT E. KOHLER JR.

Bulletin of the History of Medicine Vol. 48, No. 1 (SPRING, 1974), pp. 22-40. Published by: The Johns Hopkins University Press

2. https://en.wikipedia.org/wiki/Hans_von_Euler-Chelpin

3. Otto Heinrich Warburg & The Free Radical Theory: An abundant supply of Oxygen & Antioxidants for disease prevention and treatment of Cancer, Heart Attack, Diabetes, and other diseases. - https://www.prnewswire.com/news-releases/otto-heinrich-warburg--the-free-radical-theory-an-abundant-supply-of-oxygen--antioxidants-for-disease-prevention-and-treatment-of-cancer-heart-attack-diabetes-and-other-diseases-259901311.html