Review

Enzyme engineering for enantioselectivity: from trial-and-error to rational design?

Linda G. Otten, Frank Hollmann and Isabel W.C.E. Arends

Delft University of Technology, Department of Biotechnology, Biocatalysis and Organic Chemistry Group, Julianalaan 136, 2628 BL Delft, The Netherlands

The availability of tailored enzymes is crucial for the implementation of biocatalysis in organic chemistry. Enantioselectivity is one key parameter defining the usefulness of an enzyme and, therefore, the competitiveness of the corresponding industrial process. Hence, identification of enzymes with high enantioselectivity in the desired transformation is important. Currently, this is achieved by screening collections and libraries comprising natural or man-made diversity for the desired trait. Recently, a variety of improved methods have been developed to generate and screen this diversity more efficiently. Here, we present and discuss the most important advances in both library generation and screening. We also evaluate future trends, such as moving from random evolution to more rational.

Enzyme engineering towards enantioselective enzymes

Biocatalysis is enjoying increasing popularity in chemical synthesis. As well as the promise of enabling processes that are more benign environmentally and also economically attractive, the intrinsic chirality of enzymes is their mostly appreciated property. During the past decade, enzyme engineering technologies to tailor enzyme enantioselectivity have virtually exploded, resulting in a range of successful examples of re-designing natural enzymes to the requirements of a desired chemical synthesis.

Enzyme engineering comprises two crucial key steps. First, variants of the template enzyme are generated from which, in a second step, the most suited candidates are identified. Bottlenecks and limitations in either step result in longer timelines before discovering the best enzymes and this has been the reason for the increased research efforts. Here, we highlight developments made since 2004 in tailoring enzyme enantioselectivity and screening of natural or artificial diversity. Currently, enzyme engineering is in a transition from purely random approaches, which necessitate immense screening efforts, to (semi)rational design, which will simplify the process of tailoring enzyme enantioselectivity.

Corresponding authors: Otten, L.G. (l.g.otten@tudelft.nl); Hollmann, F. (f.hollmann@tudelft.nl).

Tailoring enantioselective enzymes: from random to rational design

Biocatalysts exhibiting the desired stereochemical properties are generally not readily available. Principally, two popular approaches can be distinguished to obtain enantioselective enzymes (Figure 1). The first aims to identify new enzymes by screening natural diversity. This can be achieved by screening microbial or metagenomic collections, which require at least basic microbiological know-how and facilities. Also an increasing number of enzymes are becoming commercially available that represent a diversity of varying substrate- and enantioselectivities. If the desired enantioselectivity cannot be readily found in nature, a given catalyst can be engineered using molecular biology. The potential of this approach has been demonstrated many times by several examples of improved and even inversed enantioselectivity of enzymes towards 'unnatural' substrates [1,2]. For example, Reetz *et al*. evolved a non-selective esterase into a highly enantioselective enzyme via directed evolution and semi-rational design [3]. Another example is the improvement of an epoxide hydrolase from E=12 to E>200 on *cis*-2,3-epoxybutane [4].

There are two different approaches to enzyme engineering: random approaches or (semi)rational design. Although these began as two opposing approaches, both methods are currently converging to produce small high-quality libraries, which are easily screened, without ignoring non-obvious amino acid changes [5,6].

Directed evolution

Directed evolution encompasses stochastic methods (*i.e.* the random changing of amino acids without prior knowledge of their exact function or position in the protein; Box 1). In the early days of enzyme engineering, focus was on designing practical mutagenesis protocols and was therefore applied to enzymes showing easily selectable traits, such as antibiotic resistance and thermostability. Since then, the scope of enzyme engineering has increased with tailoring of industrially relevant enzymes for improved stability, accepted substrates and stereospecificity. Consequently, directed evolution has emerged as a popular tool for enzyme engineering, as demonstrated by the many scientific and patent publications over the past decade [7–9].

However, the stochastic nature of directed evolution constitutes a mixed blessing. On the one hand, no structural

0167-7799/$ – see front matter doi:10.1016/j.tibtech.2009.10.001 Available online 11 November 2009

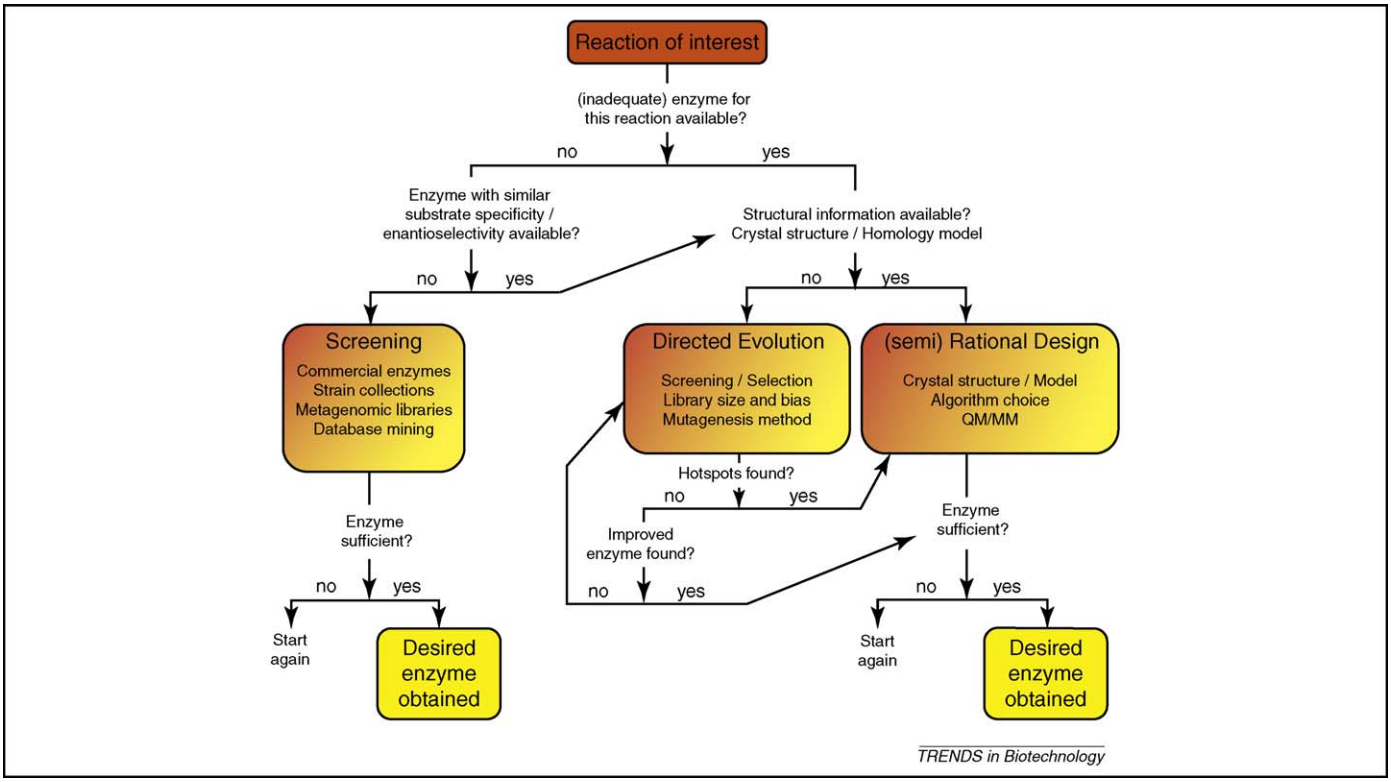

Figure 1. Decision tree for selecting the most appropriate enzyme-engineering strategy.

information of the enzyme of interest is required, making directed evolution applicable to most given enzymes. On the other hand, numerous mutant enzymes need to be screened to cover as much of the so-called sequence space as possible. Statistically, only ~10% of the amino acid residues of an enzyme are within 10 Å of the active site [10] and these residues are expected to exert the most significant influence on chiral discrimination of the enzyme [11]. As a result, 90% of all variants might be irrelevant, leading to an undesirably increased screening effort often exceeding the practical range.

Furthermore, many directed evolution protocols still suffer from technical shortcomings, such as biased amino acid exchanges (Box 1) [12]. As opposed to pursuing optimisation attempts for more efficient screening and selection systems of large libraries, the use of smaller libraries that are of higher quality (*i.e.* percentage of variants exhibiting the desired improvement) are an attractive alternative to reduce screening efforts [13,14]. This can be achieved by confining mutations to amino acids that are most likely to influence the desired trait. However, in the absence of structural information regarding the enzyme of interest,

Box 1. Library preparation in directed evolution

Directed evolution mimics darwinian evolution in a test tube using iterative cycles of mutagenesis, recombination and selection or screening for the enzyme with the desired properties. The main advantage over other enzyme-engineering approaches is that enzyme properties and functions can be engineered easily without any structural knowledge being required.

The creation of mutant libraries is an important step in directed evolution. The quality of a library is assessed by the (uniform) distribution of the mutations over the gene and the (unbiased) mutational spectrum (*i.e.* the possibility of any amino acid changing into any of the other 19 possible amino acids). Various methods to improve these two properties have been described and comprise both random mutagenesis and recombination methods [3,12,14].

In random mutagenesis, one or more mutations are randomly introduced into an enzyme. These methods are typically inexpensive and relatively simple to implement, but the mutational bias is large. First, polymerases tend to insert the same base when introducing an error, which results in mutations that are not entirely random. Engineered polymerases exhibiting less bias, or the use of complementary polymerases, can partly counteract this [67]. A second issue is the biased codon usage in the genetic code. Changing only one of the three bases encoding an amino acid does not lead to the introduction of a random amino acid, but rather to the same amino acid or a chemically similar one [68]. To circumvent this bias, degenerate oligonucleotides can be used to introduce specific (random) mutations at certain positions, or to even change entire codons [69–72]. However, this approach can become expensive when the entire gene is to be mutated.

Recombination mimics the genetic evolution via sexual reproduction. By combining mutations that are present in different enzyme homologues (parents), the diverse variants will exhibit higher frequencies of active enzymes compared to random mutagenesis. The reason for the higher quality of such a library is that the parental DNA encodes fully functional enzymes, thus sampling only productive mutations (selected over millions of years by evolution). Although the first shuffling technique developed by Stemmer is relatively easy to implement [73], its general applicability is impaired by several drawbacks, including the need for high homology between two parental enzymes and the recombination bias disfavouring combination between two nearby sites. All these problems have been tackled by different research groups and have resulted in novel recombination methods, the benefits and drawbacks of which have been extensively reviewed elsewhere [7].