Directed Evolution of Enzymes

Uwe T. Bornscheuer*

Enzymes find increasing applications in organic synthesis.[1] Their high chemo-, regio-, and stereoselectivity under mild reaction conditions, combined with a steady increase in their commercial availability at acceptable prices, make biocatalysts an attractive alternative to chemical catalysts.

In nature, evolution has ensured a steady adaption of the organism and of the proteins involved in their biological function over billions of years. However, enzymes isolated from natural sources do not always fulfill the requirements of an efficient organic synthesis. For the targeted improvement of enzyme properties—such as activity, thermo- and solvent stability, and regio- or stereoselectivity—protein design is increasingly used. Several groups recently described directed evolution as a new and very elegant combinatorial approach for this purpose (Figure 1).[2]

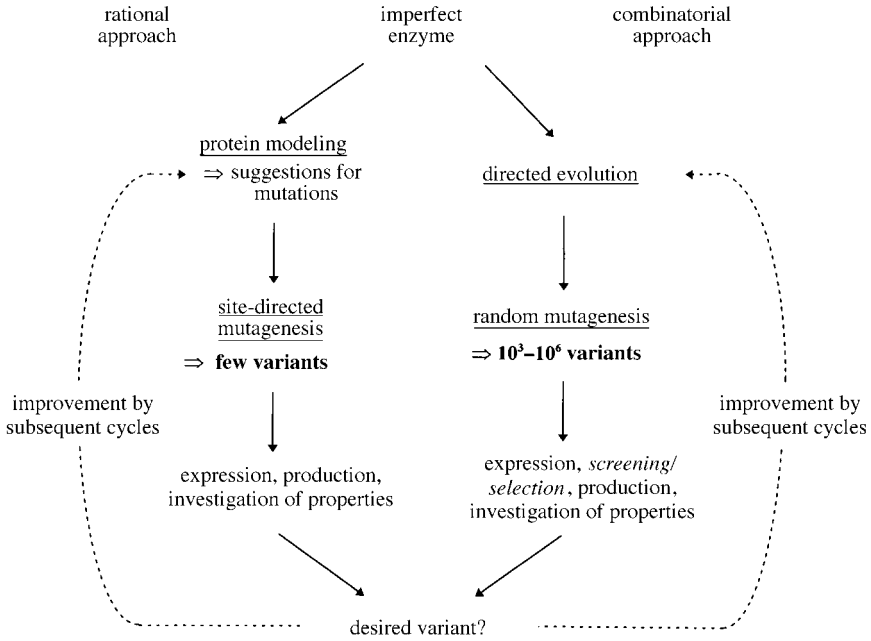


Figure 1. Comparison of the rational and the combinatorial approach for the mutagenesis of enzymes.

This method is based on the creation of enzyme libraries using modern techniques of molecular biology, which in combination with an efficient assay system allow the identification of improved biocatalysts. In contrast to the already employed rational approach of site-directed mutagenesis (Figure 1), no structural data of the enzyme and no knowledge of the relationship between amino acid sequence, structure, and mechanism of catalysis is required. Furthermore, site-directed mutagenesis is a time- and information-intensive method which allows the production of relatively few enzyme variants within reasonable lengths of time.

Prerequisites for a successful directed evolution are an effective mutation strategy for the improvement of the enzymes, the functional expression of the protein in a suitable microbial host, and a fast and reliable assay system for the identification of enzyme variants with desired properties out of a pool of 10^4 to more than 10^6 variants (Figure 1).[2c]

Two different strategies for the generation of improved enzyme variants, which are also named asexual and sexual evolution,[2d] have been described in the literature (Figure 2). In an asexual evolution, random mutagenesis is targeted preferentially towards the gene encoding one parent protein. An enzyme library is generated and then screened for improved properties. The best enzymes identified in the first generation can than be further optimized in subsequent cycles. However, this implies a step-by-step improvement of variants from the first generation by introduction of new and deletion of negative mutations (Figure 3).

In contrast, sexual evolution starts from a pool of homologous parent genes. These can originate from an asexual evolution, from related natural sequences, or from enzyme variants generated by rational design. The principle of this DNA or gene shuffling[3] is based on a partial DNase-I digestion of the gene followed by a recombination of fragments by a polymerase chain reaction (PCR). Again, an accumulation of positive and an elimination of negative

[*] Dr. U. T. Bornscheuer
Institut für Technische Biochemie der Universität
Allmandring 31, D-70569 Stuttgart (Germany)
Fax: (+49) 711-6853196
E-mail: itbubo@po.uni-stuttgart.de

  1433-7851/98/3722-3105 $ 17.50+.50/0

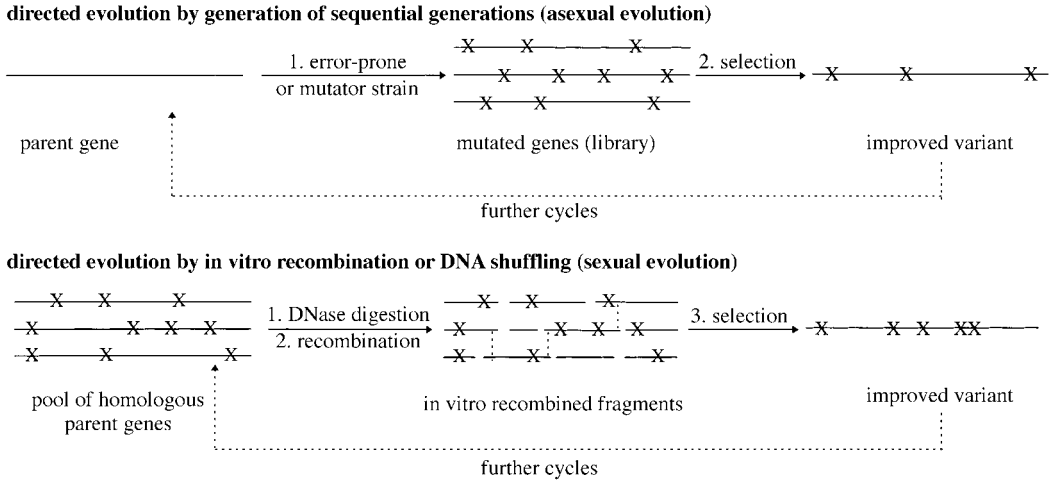


Figure 2. Methods for the directed evolution of enzymes. The asexual evolution of a starting gene through, for example, error-prone PCR or application of mutator strains introduces point mutations. The resulting enzyme library is then screened for improved enzyme variants. For sexual evolution, a pool of homologous genes is partially digested by DNAse-I and recombined by PCR in vitro to a library of enzyme variants differing in their mutation pattern. (Taken from ref. [2d, 3a] with modifications.)

mutations is desired (Figures 2 and 3). Disadvantages are that the homologous genes are not always available, the DNase digestion for the generation of fragments of suitable size is difficult to control, and there are problems with ligation efficiency during PCR. A further, recently described strategy for in vitro mutagenesis and recombination is based on a modified PCR protocol only. This staggered extension process (StEP) method allows the production of full-length genes which carry different sequence information. Beside the use of homologous parent genes, particularly short reaction times and the use of different templates during PCR are the basis of this StEP method.[4]

Probably the most important method for asexual evolution is the error-prone PCR.[5] When nonoptimal reaction conditions are used during PCR, the error rate, for example, for *Thermus aquaticus* (*Taq*) polymerase can be increased from 0.001 – 0.02 % under standard conditions to more than 1 % (Figure 2).[5b] This in vitro method is easy to perform, can be limited to the nucleotide sequence encoding the protein of interest and even shorter parts, and allows adjustment of the mutation rate. Disadvantages are the nonstatistical exchange of nucleotides and an often-observed low ligation efficiency during PCR. For the latter problem, modified PCR protocols have been developed.[5c, 6]

Alternatively, mutator strains with defects in their DNA-repair mechanisms can be used, which leads to the introduction of mutations during replication. The sequence which is subjected to mutations can be limited by locating the gene encoding the protein of interest on plasmids. The application of mutator strains which are, for example, commercially available as *Escherichia coli* variant *Epicurian coli* XL1-Red[7] (Stratagene, La Jolla, USA) is simpler than error-prone PCR, but the mutation rate cannot be adjusted. Furthermore, owing to a mutation of the entire plasmid—for example, in the promotor region—defects but also improvements can occur. Thus, while researchers can chose from a variety of methods to generate enzyme variants, it is still unclear which method is the best.

The number of enzyme variants which can be generated by directed evolution grows exponentially with the size of the enzymes and the number of simultaneously exchanged amino acids. Even for a small protein of 200 amino acids more than nine billion possible variants can theoretically be generated simply by introducing three substitutions at the same time.[2c] These cannot be analyzed with common tedious and time-consuming methods such as HPLC or gas chromatography. Therefore, according to the first law for directed mutagenesis proposed by F. Arnold, "you only get what you screen for", an efficient screening system is of extreme importance.[8] A selection can be based, for instance, on altered antibiotic resistance[9] or growth on media lacking components essential for growth.[10] Problems related to organic synthesis require assay systems which provide direct information about the enzyme properties. For example, the identification of a lipase variant with improved stereoselectivity is possible by measuring the ratio of the rates of hydrolysis of enantiomerically pure (*R*)- or (*S*)-*p*-nitrophenol esters.[11] Also fluorescence correlation spectroscopy[12] might allow the rapid identification of active enzyme variants in a library, because the change in fluorescence during the reaction can be followed for single molecules.

Often the enzyme library producing colonies cannot be assayed directly because the enzymes are, for example, produced intracellulary and a cell lysis is necessary. Additional prerequisites for the detection of all variants are correct processing and folding to ensure the production of active enzymes. Furthermore, all enzyme variants should exhibit high stability and activity under assay conditions.

The directed evolution of enzymes is a new area which is especially dominated by biochemistry. This might explain why so far only a few examples deal with the solution of problems related to organic chemistry. Some of them are presented at the end of this article. However, also for other problems, especially for the technical application of biocatalysts, directed evolution represents an attractive alternative to existing methods for protein design. For instance, enzyme often do not

  1433-7851/98/3722-3106 $ 17.50+.50/0

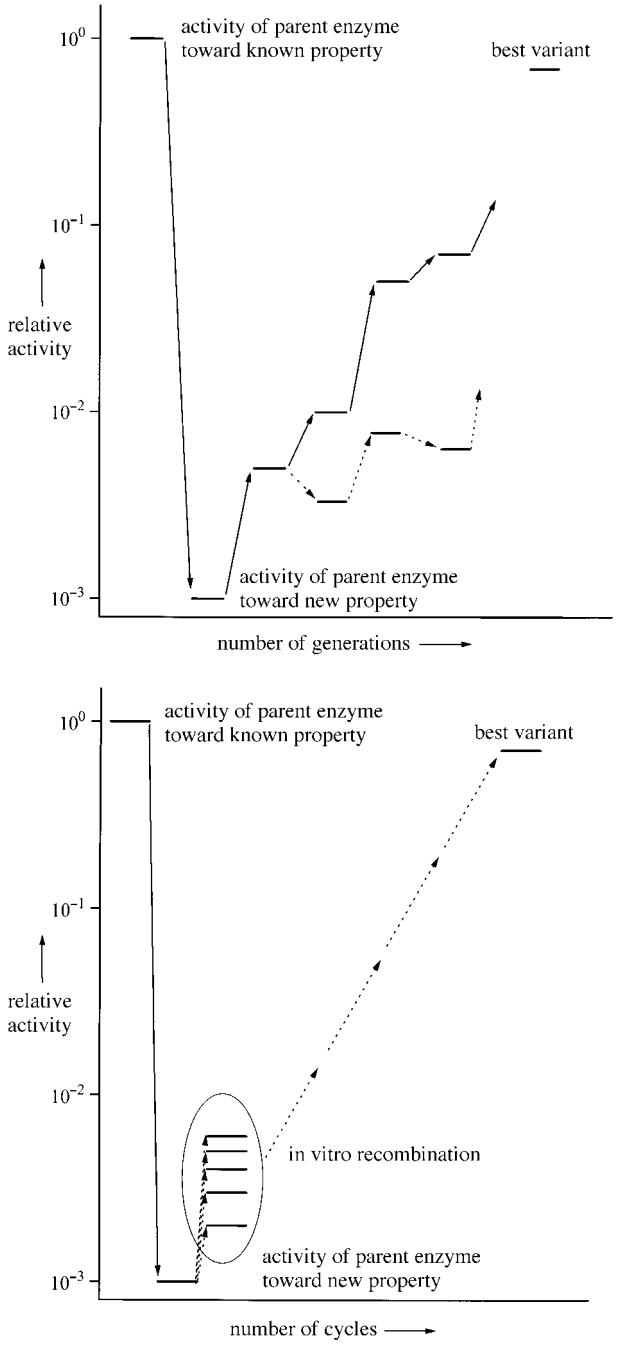


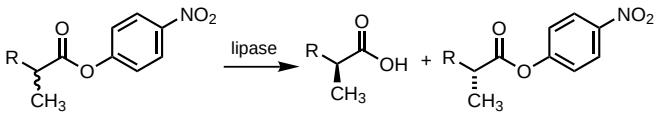
Figure 3. Strategies for improving a wild-type enzyme by production of sequential generations from one parent gene (asexual evolution, top) or by in-vitro recombination through DNA shuffling (sexual evolution, bottom). During asexual evolution a successive improvement (solid arrows) as well as a worsening (dashed arrows) can occur in higher generations. For sexual evolution, a pool of homologous variants from several parent genes (circle) is necessary before an in vitro recombination can be performed. (Taken from ref. [8] with modifications.)

exhibit sufficient stability under process conditions with respect to pH profile, temperature-dependent activity, and stability, and are not active in all solvents. These deficits are extremely difficult to overcome by site-directed mutagenesis. Here, directed evolution can contribute to the fast generation and identification of improved enzymes.

DNA shuffling as a method for sexual evolution led to the generation of a variant of the green-fluorescent protein (GFP) with a 45-fold increase in fluorescence.[13] The recombination of almost identical cephalosporinase genes by DNA shuffling increased the moxalactam resistance by a factor of 540.[9] In the case of cefotaxime resistance even a 16000 times greater resistance of a β-lactamase gene in *E. coli* could be achieved after three cycles of sexual evolution.[3b] The half-life of the protease subtilisin E at 65 °C was increased by 50 times through recombination of two subtilisin E variants by the StEP method.[4] A growth assay was used in the generation of a monomeric chorismate mutase (CM).[10b] This naturally dimeric enzyme catalyzes the conversion of chorismate into prephenate in the biosynthesis of L-tyrosine (L-Tyr) and L-phenylalanine (L-Phe). A combination of directed evolution with selection of variants on a media lacking L-Tyr and L-Phe allowed the identification of a 100-fold more active monomeric CM variant. This provided insight into the influence of the topology on structure, stability, and function of this protein. The generation of new antibiotics and other pharmacologically active substances should be possible through directed evolution of polyketide synthases (PKS). A number of PKS genes has been cloned already, and application of the methods described here for direction evolution has been proposed.[14] Further examples have been described in more detail in a number of recent reviews.[2]

For organic synthesis, the alteration of stereoselectivity and substrate specificity can be considered as being most important. Directed evolution might give access to a large number of optically pure compounds which are not or only insufficiently converted by wild-type enzymes.

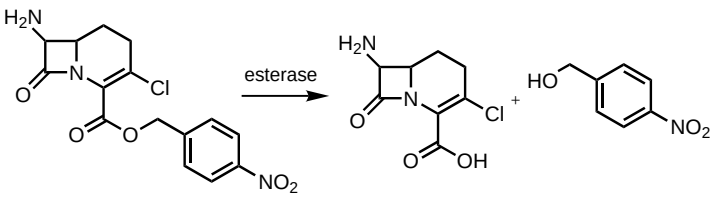
Based on the stereoselective photometrical assay with *p*-nitrophenol esters, the enantioselectivity of a lipase from *Pseudomonas aeruginosa* PAO1 for hydrolysis of racemic 2-methyl decanoate could be increased from 2% *ee* (wild-type) to 81% *ee* (variant) by error-prone PCR (Scheme 1).[11] A directed evolution of an esterase from



Scheme 1. Improvement of the stereoselectivity of a lipase from *Pseudomonas aeroginosa* PAO1 from 2% *ee* to 81% *ee* for hydrolysis of racemic 2-methyl decanoate by error-prone PCR.[11]

Pseudomonas fluorescens (PFE) with the mutator strain *Epicurian coli* XL1-Red produced a double mutant of this enzyme, which catalyzed the hydrolysis of a sterically hindered 3-hydroxy ester at a moderate rate and with moderate stereoselectivity.[10a] The 3-hydroxy ester can be used as a precursor in the synthesis of epothilones, and was not cleaved by 20 hydrolases including wild-type PFE. The screening of an esterase library was based on a combination of an indicator assay and a growth assay.[10a] A combination of error-prone PCR and DNA shuffling led to the generation of a more stable and active variant of an esterase from *Bacillus subtilis*.[15] This enzyme hydrolyzes the *p*-nitrobenzyl ester of loracarbef, a cephalosporin antibiotic, with 150 times higher activity compared to the wild-type enzyme in 15% dimethylformamide (Scheme 2).[8]

  1433-7851/98/3722-3107 $ 17.50+.50/0



Scheme 2. A combination of error-prone PCR and DNA shuffling led to a variant of an esterase from *Bacillus subtilis* which exhibits 150-fold higher activity in 15% DMF compared to the wild-type enzyme for cleavage of the *p*-nitrobenzyl ester of loracarbef. For practical reasons, the assay was based on the hydrolysis of *p*-nitrophenol esters.[8, 15]

Despite the methods developed so far and the successful examples given above, several problems have to be solved to allow a broader application of directed evolution. This includes the further optimization of methods for the generation of mutations and enzyme libraries, and the development of molecular techniques which allow the use of directed evolution for other microbial expression systems. Of upmost importance are highly efficient assay systems, without which the generation of enzyme libraries does not make much sense. Here, equipment and protocols already developed for high-throughput screening in combinatorial chemistry might be used and optimized to allow the quick screening of large enzyme libraries. This might be achieved, for example, by the use of fusions proteins based on the easy detectable green-fluorescent protein.[16]

In general, directed evolution represents a promising extension to methods existing so far for protein design. It can be expected that it will largely contribute to the development of new or improved biocatalysts, but might also allow a better understanding of structure–activity relationships. Directed evolution might even become a biotechnological equivalent of combinatorial chemistry.

German version: *Angew. Chem.* **1998**, *110*, 3284–3288

Keywords: enzyme catalysis · molecular evolution · mutagenesis

[1] a) *Enzyme Catalysis in Organic Synthesis, Vol. 1 and 2* (Eds.: K. Drauz, H. Waldmann), 1. ed., VCH, Weinheim, **1995**; b) K. Faber, *Biotransformations in Organic Chemistry*, 3. ed., Springer, Berlin, **1997**; c) H. G. Davies, R. H. Green, D. R. Kelly, S. M. Roberts, *Biotransformations in Preparative Organic Chemistry: The Use of Isolated Enzymes and Whole Cell Systems in Synthesis*, 1. ed., Academic Press, London, **1989**; d) C.-H. Wong, G. M. Whitesides, *Enzymes in Synthetic Organic Chemistry*, 1. ed., Pergamon, Oxford, **1994**; e) *Biotechnology, Vol. 8: Biotransformations* (Eds.: H. J. Rehm, G. Reed, A. Pühler, P. J. W. Stadler), VCH, Weinheim, **1998**.

[2] a) W. P. C. Stemmer, *Bio/Technology* **1995**, *13*, 549–553; b) G. F. Joyce, *Sci. Am.* **1992**, *267*(6), 48–55; c) O. Kuchner, F. H. Arnold, *Trends Biotechnol.* **1997**, *15*, 523–530; d) F. H. Arnold, *Acc. Chem. Res.* **1998**, *31*, 125–131; e) A. Schwienhorst, *Biospektrum* **1998**, *4*, 44–47; f) S. Harayama, *Trends Biotechnol.* **1998**, *16*, 76–82.

[3] a) W. P. C. Stemmer, *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 10747–10751; b) W. P. C. Stemmer, *Nature* **1994**, *370*, 389–391.

[4] H. Zhao, L. Giver, J. A. Affholter, F. H. Arnold, *Nature Biotechnol.* **1998**, *16*, 258–261.

[5] a) D. W. Leung, E. Chen, D. V. Goeddel, *Technique (Philadelphia)* **1989**, *1*, 11–15; b) R. C. Cadwell, G. F. Joyce, *PCR Meth. Appl.* **1992**, *2*, 28–33; c) J. H. Spee, W. M. de Vos, O. P. Kuipers, *Nucleic Acids Res.* **1993**, *21*, 777–778.

[6] J.-P. Vartanian, M. Henry, S. Wain-Hobson, *Nucleic Acids Res.* **1996**, *24*, 2627–2631.

[7] A. Greener, M. Callahan, B. Jerpseth, *Methods Mol. Biol.* **1996**, *57*, 375–385.

[8] F. H. Arnold, J. C. Moore, *Adv. Biochem. Eng. Biotechnol.* **1997**, *58*, 1–14.

[9] A. Crameri, S.-A. Raillard, E. Bermudez, W. P. C. Stemmer, *Nature* **1998**, *391*, 288–290.

[10] a) U. T. Bornscheuer, J. Altenbuchner, H. H. Meyer, *Biotechnol. Bioeng.* **1998**, *58*, 554–559; b) G. MacBeath, P. Kast, D. Hilvert, *Science* **1998**, *279*, 1958–1961.

[11] M. T. Reetz, A. Zonta, K. Schimossek, K. Liebeton, K.-E. Jaeger, *Angew. Chem.* **1997**, *109*, 2961–2963; *Angew. Chem. Int. Ed. Engl.* **1997**, *36*, 2830–2832.

[12] M. Eigen, R. Rigler, *Proc. Natl. Acad. Sci. USA* **1994**, *91*, 5740–5747.

[13] A. Crameri, E. A. Whitehorn, E. Tate, W. P. C. Stemmer, *Nature Biotechnol.* **1996**, *14*, 315–319.

[14] C. Khosla, R. Caren, C. M. Kao, R. McDaniel, S.-W. Wang, *Biotechnol. Bioeng.* **1996**, *52*, 122–128.

[15] J. C. Moore, F. H. Arnold, *Nature Biotechnol.* **1996**, *14*, 458–467.

[16] M. M. Enzelberger, F. Zocher, C. Schmidt-Dannert, U. T. Bornscheuer, R. D. Schmid, B. Hauer, H. Eipel, *BIOforum* **1998**, *4*, 192–194.

  1433-7851/98/3722-3108 $ 17.50+.50/0