doi:10.2533/chimia.2007.100 CHIMIA **2007**, *61*, No. 3

Chimia 61 (2007) 100–103

ISSN 0009–4293

Evolution in the Test-Tube as a Means to Create Selective Biocatalysts

Manfred T. Reetz*
Recipient of the Prelog Medal 2006

Abstract: The development of chiral catalysts for use in synthetic organic chemistry is traditionally associated with progress in asymmetric transition metal catalysis or organocatalysis. In addition to these options chemists have employed enzymes for a number of asymmetric transformations for a long time. However, limited substrate acceptance and insufficient enantioselectivity are the crucial factors which prevent the general exploitation of biocatalysis in organic chemistry. In order to solve this long-standing problem, we proposed some time ago the concept of directed evolution of enantioselective enzymes. This report constitutes a short overview of the principle involved, together with a description of the first examples and the illumination of the present and future challenges.

Keywords: Asymmetric catalysis · Directed evolution · Enantioselectivity · Enzymes · Gene mutagenesis · High-throughput screening

In the 1990s we proposed a fundamentally new approach to asymmetric catalysis, namely the directed evolution of enantioselective enzymes for application in synthetic organic chemistry.[1] It makes use of the known molecular biological methods that had previously been developed for improving thermostability and stability of enzymes toward hostile organic solvents.[2] Our approach is therefore based on the appropriate combination of random gene mutagenesis, expression and high-throughput screening to assess thousands of samples for enantiopurity (Fig 1).[1][3]

Whenever the natural (wild-type, WT) enzyme shows an unacceptably low enantiomeric excess (*ee*) or selectivity factor (*E*-value) for a given transformation of interest, a library of mutants is created from which the most enantioselective mutant is identified. Following isolation of the corresponding mutant gene, the process is repeated which exerts evolutionary pressure on the system. Typically, 1000 to 10,000 clones make up a given library, but this can

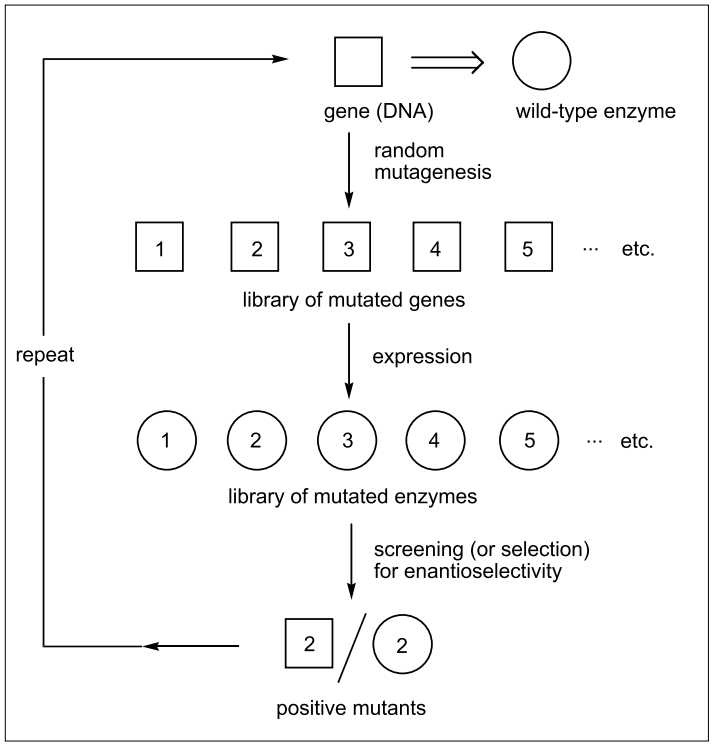


Fig. 1. Strategy for directed evolution of an enantioselective enzyme[1]

be more. Experimentally, in each cycle the library of mutant genes is first inserted in a standard bacterial host such as *Escherichia coli*. Then bacterial colonies are plated out on agar plates and harvested individually by a colony picker. Each colony is placed robotically in a separate well of a microtiter plate containing nutrient broth, so that the bacteria grow and produce the protein of interest. Because each colony originates from a single cell, mixtures of mutant enzymes are avoided (of course, other proteins are present which do not disturb the evolutionary process). A portion of each mutant enzyme occurring in the harvested bacterial colony is then placed on a different micro-

*Correspondence: Prof. Dr. M. T. Reetz
Max-Planck-Institut für Kohlenforschung
Kaiser-Wilhelm-Platz 1
D-45470 Mülheim/Ruhr, Germany
Tel.: +49 208 306 2000
Fax: +49 208 306 2985
E-Mail: reetz@mpi-muelheim.mpg.de

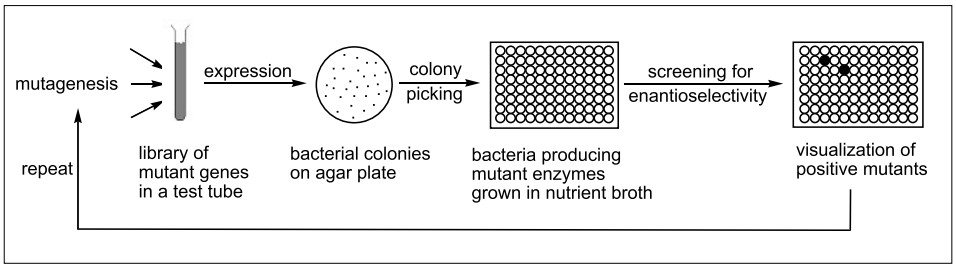


Fig. 2. The experimental stages of directed evolution of enantioselective enzymes[1]

titer plate where the reaction of interest occurs (Fig. 2).

In the extreme case knowledge of the structure of the enzyme or of its mechanism is not necessary, yet enhanced enantioselectivity can be expected! Thus, in contrast to so-called *de novo* design of enzymes using site-specific mutagenesis,[4] our approach ignores any thoughts regarding steric or electronic effects, possible hydrogen bonds, *etc*. Nevertheless, it is also logical, specifically in a different sense because it relies on the Darwinistic character of the overall process. Where are therefore the intellectual challenges in putting Fig. 2 into practice? On the one hand the development of efficient high-throughput *ee*-assays poses a formidable problem. We and subsequently other groups have invested a great deal of efforts in designing and experimentally implementing such screening systems. This crucial aspect of directed evolution of enantioselective enzymes has been reviewed elsewhere and will not be detailed here.[3] Suffice it to say that today's medium- and high-throughput *ee*-screens allow between 400 and 8,000 samples to be analyzed for enantiopurity per day.

Crucial for success is also the choice of the mutagenesis method, specifically the design of a strategy to probe protein sequence space most efficiently. Actually, any mutagenesis method can be expected to generate more or less improved mutants, and in fact in our early work we strived simply for proof-of-concept, which was achieved.[1a] Shortly thereafter it became clear that the most important challenge revolves around the question of how to probe protein sequence space so efficiently that the experimental molecular biological work and the screening effort can be minimized while maximizing the benefits (especially enantioselectivity).[1e,5]

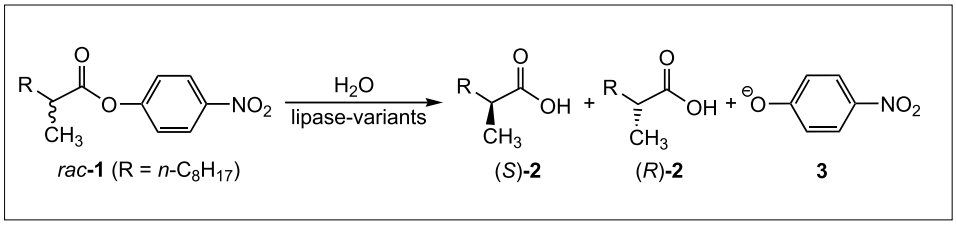
The most important random gene mutagenesis methods at the time of our first directed evolution project were error-prone polymerase chain reaction (epPCR), saturation mutagenesis and DNA shuffling.[2] The most often used method, epPCR, introduces mutations more or less randomly over the whole enzyme (although it is not truly random due to the degeneracy of the genetic code). Saturation mutagenesis induces the randomization of amino acids at a given position in the enzyme or simultaneously at two, three or more amino acid positions. These were the methods that we employed in our initial project which concerned the *Pseudomonas aeruginosa* lipase(PAL)-catalyzed hydrolytic kinetic resolution of *rac*-**1** (Scheme 1). The WT shows a slight preference for the (*S*)-enantiomer ($E = 1.2$).

Following four rounds of epPCR at low mutation rate averaging one amino acid exchange per enzyme molecule, the *E*-value climbed stepwise to 11.3,[1a] which constitutes the first example of directed evolution of an enantioselective enzyme. However, it was obvious at the time that this strategy is not optimal. We therefore suggested the idea that the identified mutations are sensitive positions (hot spots), but that the particular amino acids found there are not optimal, leading to the conclusion that saturation mutagenesis at the hot spots could generate even better mutants. This turned out to be so, but not in all cases. Moreover, DNA shuffling of mutants obtained using epPCR at high mutation rate (three amino acid exchanges per enzyme molecule) also proved to be effective. This research not only culminated in the best mutant ($E = 51$),[6] it also allowed for certain conclusions regarding improved strategies for exploring protein sequence space. These efforts are summarized in Fig. 3. Moreover, the best mutant with six mutations, five of which are remote, was analyzed theoretically (MM/QM) which led to the proposal of an intriguing relay mechanism.[7]

The lessons learned from these studies were then used as a guide in the directed evolution of enantioselective monooxygenases such as cyclohexanone monooxygenase as catalysts in Baeyer-Villiger reactions (*ee* = 90–99%)[8] and sulfoxidation of thio-ethers (*ee* = 95–99%).[9] Reversal of enantioselectivity was achieved in several cases.[8–10] Moreover, other academic and industrial groups have picked up the general concept (Figs 1 and 2) and evolved a number of highly enantioselective enzymes for their particular purpose, including esterases, aldolases, hydantoinases, nitrilases, phosphotriesterases, and monoamine oxidases.[10]

Despite these success stories, we were not fully content with the strategies that were being used to evolve enantioselectivity, because in most cases a time-consuming overall process was necessary,[10] as in most other directed evolution studies.[2] Fast directed evolution seemed necessary, especially from an industrial viewpoint. We therefore developed a method called Iterative Saturation Mutagenesis (ISM),[11,12] which indeed appears to be exceptionally efficient because it reduces the experimental molecular biological work and the screening effort drastically while providing even better results. This type of 'accelerated' directed evolution is based on iterative cycles of saturation mutagenesis at predetermined sites in an enzyme, a given site being composed of one, two or three (or more) amino acid positions. The concept is illustrated here for the case involving four sites A, B, C and D of an enzyme, each site being considered only once in a given upward pathway (Fig. 4).

Experimentally, the respective four initial libraries prepared by saturation mutagenesis at sites A, B, C and D are screened, and the best hit in each case is isolated and sequenced. At this stage a decision has to be made as to which branch in the fitness landscape is chosen for further improvements. It may be hierarchical (take the best mutant), but in fact it is arbitrary. If one pathway fails to provide improved mutants, a different one can be explored. In any case, the gene corresponding to the best hit of a given library is used as a template to perform another round of saturation mutagenesis at a different site (or sites if moving to another branch), and the process is continued (Fig. 4). In the case of four sites as shown in Fig. 4, there are 64 saturation mutagenesis libraries to be constructed if all of the confined protein sequence space is to be explored. At this point, another statistical aspect comes into play, namely oversampling.[12b, 13] For example, if a site is composed of two amino acid positions, 400



Scheme 1.

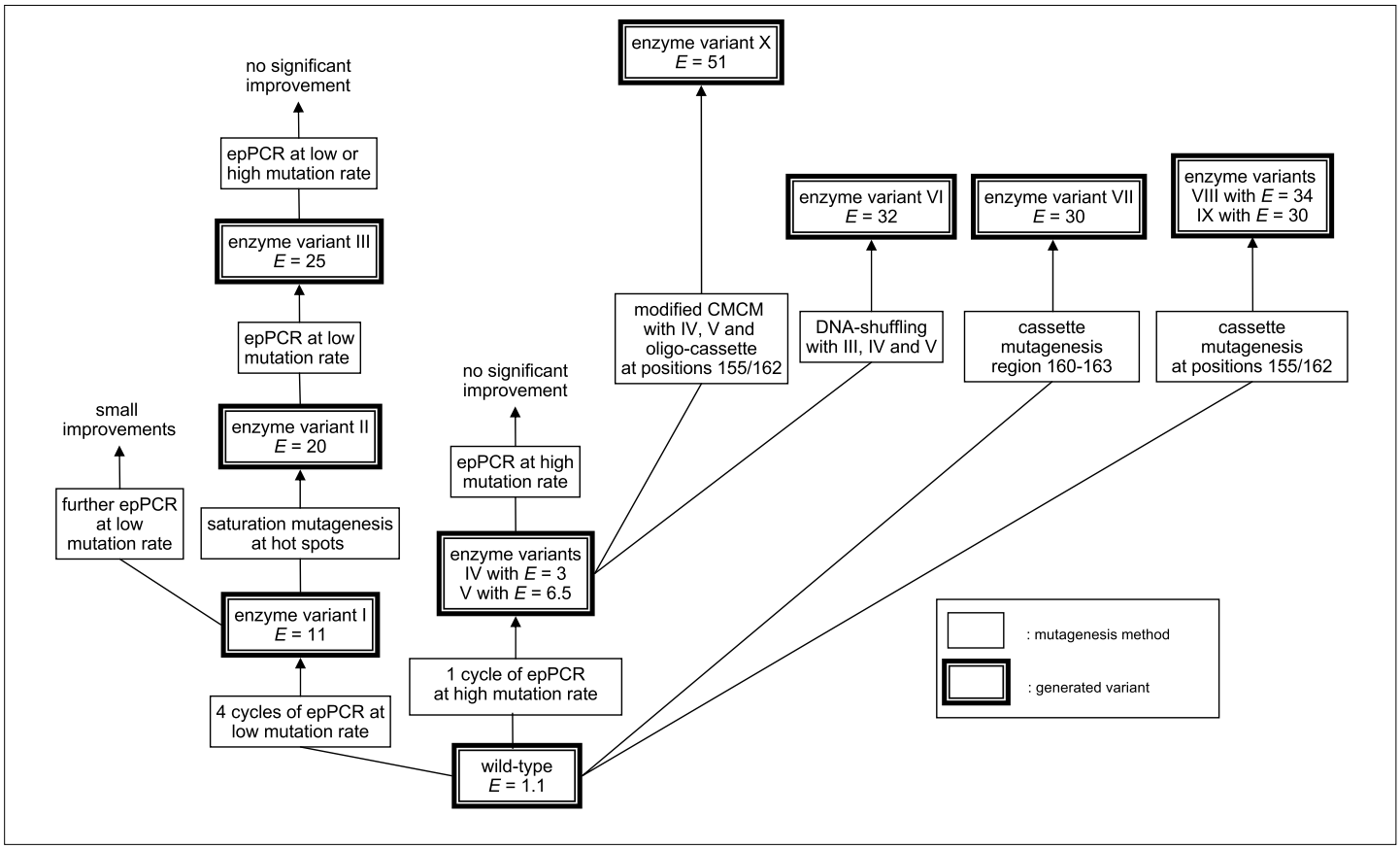


Fig. 3. Schematic summary of the directed evolution of enantioselective lipase-variants originating from the WT PAL used as catalysts in the hydrolytic kinetic resolution of ester *rac*-**1**. CMCM = Combinatorial multiple-cassette mutagenesis.[5a,6]

theoretically different mutants are possible, but in order to reach 95% coverage of this defined section of protein sequence space, an excess of about 3000 clones (bacterial colonies) need to be harvested and screened. Of course, such high coverage is usually not necessary. Moreover, a different codon usage can be employed, *e.g.* one which uses only twelve amino acids as building blocks. This reduces the necessary number clones to be screened drastically. [12b]

Crucial for the success of ISM is the criterion for choosing the sites A, B, C, D, *etc*. The concept of ISM constitutes a combination of rational design and combinatorial randomization, because it is based on a Cartesian view of the enzyme to be genetically manipulated. This means that the enzyme is dissected into defined regions on the basis of structural information. In the case of expanding substrate acceptance (rate) or enantioselectivity, the choice for defining the sites is made by applying the Combinatorial Active-Site Saturation Test (CAST).[12a] Accordingly, an X-ray structure or a homology model is used to identify appropriate sites around the complete binding pocket. Such a systematization distinguishes CASTing from previous focused libraries that we and other groups had reported earlier.[2,6] CASTing was originally developed and used to expand the range of substrate acceptance of enzymes, but only the initial saturation mutagenesis libraries were considered. For further evolutionary improvement two strategies can be considered: 1) Simply combining the mutations of the hits originating from two different libraries, or 2) Performing ISM as illustrated in Fig. 4.

We first applied iterative CASTing, which is an embodiment of ISM, in order to evolve the enantioselectivity of an epoxide hydrolase as a catalyst in the hydrolytic kinetic resolution of a racemic epoxide.[11] The selectivity factor E increased from 4.6 to 115 in just five cycles of iterative saturation mutagenesis. This is dramatically better than the result of our earlier study based on epPCR at different mutation rates which resulted in $E = 10.8$.[14] In both cases the same number of clones were involved (20,000), which is strong evidence that ISM is more efficient than the traditional approach.

ISM can also be used to enhance the thermostability of enzymes in a process which is also rapid.[15] This concerns a completely different catalytic parameter, and therefore the criterion for choosing the sensitive sites A, B, C, D, *etc*. for saturation mutagenesis is expected to be different. Indeed, in a study directed towards enhancing the thermostability of a lipase, we made use of the so-called B-factors which are available from X-ray data. They reflect electron smearing

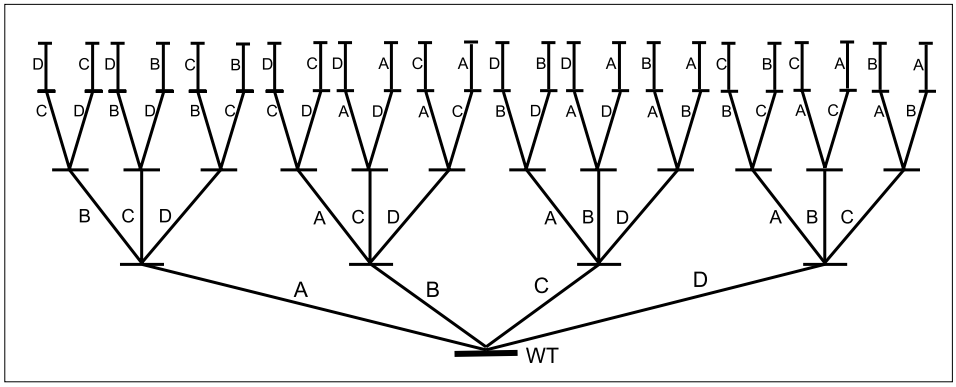


Fig. 4. Iterative Saturation Mutagenesis (ISM) employing four sites A, B, C and D, each site in a given upward pathway being visited only once[11]

around nuclei as a result of thermal motion or disorder, which correlates with flexibility (high B-factors indicate flexibility, low B-factors point to rigidity). Since it was known that hyperthermophilic enzymes are more rigid than the mesophilic analogs, the goal was to increase rigidity at the appropriate sites of a given enzyme, specifically at those sites showing the highest average B-factors. Using this criterion, we developed a method for enhancing thermostability rapidly, called B-FIT, which is another embodiment of ISM. We applied B-FIT to the lipase from *Bacillus subtilis*, and in just five cycles of iterative saturation mutagenesis along a hierarchically chosen upward pathway an increase in thermostability of 45 °C was readily achieved.[15] This constitutes a dramatic increase in thermostability.

In summary, the concept of directed evolution of enantioselective enzymes has emerged as a principally new and viable approach to asymmetric catalysis. The traditional uses of several rounds of epPCR, saturation mutagenesis and/or DNA shuffling are successful in this endeavour, but more recent strategies for probing protein sequence space are much more efficient, specifically Iterative Saturation Mutagenesis (ISM). This form of rapid directed evolution can be employed in the quest to enhance both enantioselectivity and thermostability of enzymes. It has also been used successfully in the directed evolution of enantioselective hybrid catalysts, *i.e.* of proteins to which synthetic achiral transition metal catalysts have been anchored covalently or non-covalently in order to obtain asymmetric catalysts.[16] ISM thus opens up new perspectives for White Biotechnology, including possible applications in metabolic engineering.

Acknowledgements

Generous support from the Fonds der Chemischen Industrie and the EU (IBAAC MCRTN-CT-2003-505020) is gratefully acknowledged.

Received: January 18, 2007

[1] a) 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; b) M. T. Reetz, A. Zonta, K. Schimossek, K. Liebeton, K.-E. Jaeger, DE-A 197 31 990.4, **1997**; c) M. T. Reetz, *Pure Appl. Chem*. **1999**, *71*, 1503–1509; d) K. Schimossek, Ph. D. Thesis, Ruhr-Universität Bochum, **1998**; e) M. T. Reetz, in 'Methods in Enzymology', Eds. D. E. Robertson, J. P. Noel, Elsevier Academic Press, San Diego, California, **2004**, Vol. 388, p 238–256.

[2] a) F. H. Arnold, G. Georgiou, Eds., 'Directed Enzyme Evolution: Screening and Selection Methods', Humana Press, Totowa, New Jersey, Vol. 230, **2003**; b) S. Brakmann, K. Johnsson, Eds., 'Directed Molecular Evolution of Proteins or How to Improve Enzymes for Biocatalysis', Wiley-VCH, Weinheim, **2002**; c) S. Brakmann, A. Schwienhorst, Eds., 'Evolutionary Methods in Biotechnology (Clever Tricks for Directed Evolution)', Wiley-VCH, Weinheim, **2004**; d) S. V. Taylor, P. Kast, D. Hilvert, *Angew. Chem*. **2001**, *113*, 3408–3436; *Angew. Chem. Int. Ed.* **2001**, *40*, 3310–3335; e) K. A. Powell, S. W. Ramer, S. B. del Cardayré, W. P. C. Stemmer, M. B. Tobin, P. F. Longchamp, G. W. Huisman, *Angew. Chem*. **2001**, *113*, 4068–4080; *Angew. Chem. Int. Ed.* **2001**, *40*, 3948–3959.

[3] a) M. T. Reetz, *Angew. Chem*. **2001**, *113*, 292–320; *Angew. Chem. Int. Ed*. **2001**, *40*, 284–310; b) J.-L. Reymond, Ed., 'Enzyme Assays – High-throughput Screening, Genetic Selection and Fingerprinting', Wiley-VCH, Weinheim, **2005**.

[4] F. Cedrone, A. Ménez, E. Quéméneur, *Curr. Opin. Struct. Biol.* **2000**, *10*, 405–410.

[5] a) M. T. Reetz, *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 5716–5722; b) S. Lutz, W. M. Patrick, *Curr. Opin. Biotechnol.* **2004**, *15*, 291–297.

[6] M. T. Reetz, S. Wilensek, D. Zha, K.-E. Jaeger, *Angew. Chem.* **2001**, *113*, 3701–3703; *Angew. Chem. Int. Ed.* **2001**, *40*, 3589–3591.

[7] a) M. Bocola, N. Otte, K.-E. Jaeger, M. T. Reetz, W. Thiel, *ChemBioChem* **2004**, *5*, 214–223; b) M. T. Reetz, M. Puls, J. D. Carballeira, A. Vogel, K.-E. Jaeger, T. Eggert, W. Thiel, M. Bocola, N. Otte, *ChemBioChem* **2007**, *8*, 106–112.

[8] a) M. T. Reetz, B. Brunner, T. Schneider, F. Schulz, C. M. Clouthier, M. M. Kayser, *Angew. Chem.* **2004**, *116*, 4167–4170; *Angew. Chem. Int. Ed.* **2004**, *43*, 4075–4078; b) M. D. Mihovilovic, F. Rudroff, A. Winninger, T. Schneider, F. Schulz, M. T. Reetz, *Org. Lett.* **2006**, *8,* 1221–1224.

[9] M. T. Reetz, F. Daligault, B. Brunner, H. Hinrichs, A. Deege, *Angew. Chem*. **2004**, *116*, 4170–4173; *Angew. Chem. Int. Ed.* **2004**, *43*, 4078–4081.

[10] Comprehensive review of directed evolution of enantioselective enzymes covering 1997–2005: M. T. Reetz, in 'Advances in Catalysis', Eds. B. C. Gates, H. Knözinger, Elsevier, San Diego, **2006**, Vol. 49, p 1–69.

[11] M. T. Reetz, L.-W. Wang, in part M. Bocola, *Angew. Chem.* **2006**, *118*, 1258–1263; Erratum, 2556; *Angew. Chem. Int. Ed.* **2006**, *45*, 1236–1241; Erratum, 2494.

[12] a) M. T. Reetz, M. Bocola, J. D. Carballeira, D. Zha, A. Vogel, *Angew. Chem.* **2005**, *117*, 4264–4268; *Angew. Chem. Int. Ed.* **2005**, *44*, 4192–4196; b) M. T. Reetz, J. D. Carballeira, *Nat. Protoc.*, in press.

[13] a) A. D. Bosley, M. Ostermeier, *Biomol. Eng.* **2005**, *22*, 57–61; b) W. M. Patrick, A. E. Firth, *Biomol. Eng*. **2005**, *22*, 105–112.

[14] M. T. Reetz, C. Torre, A. Eipper, R. Lohmer, M. Hermes, B. Brunner, A. Maichele, M. Bocola, M. Arand, A. Cronin, Y. Genzel, A. Archelas, R. Furstoss, *Org. Lett.* **2004**, *6*, 177–180.

[15] M. T. Reetz, J. D. Carballeira, A. Vogel, *Angew. Chem.* **2006**, *118*, 7909–7915; *Angew. Chem. Int. Ed.* **2006**, *45*, 7745–7751.

[16] a) M. T. Reetz, M. Rentzsch, A. Pletsch, M. Maywald, *Chimia* **2002**, *56*, 721–723; b) M. T. Reetz, J. J.-P. Peyralans, A. Maichele, Y. Fu, M. Maywald, *Chem. Commun*. **2006**, 4318–4320.