CHEMBIOCHEM

DOI: 10.1002/cbic.201100412

Enhancing the Thermal Robustness of an Enzyme by Directed Evolution: Least Favorable Starting Points and Inferior Mutants Can Map Superior Evolutionary Pathways

Yosephine Gumulya[a] and Manfred T. Reetz*[a, b]

In a previous directed evolution study, the B-FIT approach to increasing the thermal robustness of proteins was introduced and applied to the lipase from *Bacillus subtilis*. It is based on the general concept of iterative saturation mutagenesis (ISM), according to which sites in an enzyme are subjected to saturation mutagenesis, the best hit of a given library is then used as a template for randomization at other sites, and the process is continued until the desired catalyst improvement has been achieved. The appropriate choice of the ISM sites is crucial; in the B-FIT method the criterion is residues characterized by highest *B* factors available from X-ray crystallography data. In the present study, B-FIT was employed in order to increase the thermal robustness of the epoxide hydrolase from *Aspergillus niger*. Several rounds of ISM resulted in the best variant showing a 21 °C increase in the T_{50}^{60} value, an 80-fold improvement in half-life at 60 °C, and a 44 kcal mol^{-1} improvement in inactivation energy. Seven other variants were also evolved with moderate yet significant improvements; these were characterized by 10–14 °C increases in T_{50}^{60}, 20–30-fold improvement in half-lives at 60 °C and 15–20 kcal mol^{-1} elevations in activation energy. Unexpectedly, in the ISM process the best variants were obtained from essentially neutral or even inferior mutant parents, that is, when a given library contains no improved mutants. This constitutes a practical way to escape from what appear to be local minima ("dead ends") in the fitness landscape—a finding of notable significance in directed evolution.

Introduction

Sufficiently high thermostability of enzymes is a prerequisite for their application in organic chemistry and biotechnology.[1] For this reason protein engineering techniques utilizing site-specific mutagenesis based on rational design[2] or directed evolution[3] by using error-prone polymerase chain reaction (epPCR), DNA shuffling and/or other mutagenesis methods have been applied successfully in numerous cases. Increases in melting temperature (T_m) or T_{50}^{60} value (the temperature at which half of the activity is maintained after a 60 min heat treatment)[4] in the range of 5–15 °C proved to be typical in this important endeavor.[2–5] It should be noted that T_{50}^{60} values are generally used to describe thermostability, but this might in fact include thermoresistance to aggregation accompanied by undesired precipitation (in extreme cases only this factor).

In order to enhance the efficacy of directed evolution in general by producing smaller yet higher-quality mutant libraries requiring less screening effort[6] (the bottleneck of laboratory evolution),[7] we have developed the concept of iterative saturation mutagenesis (ISM).[8,9] It is a straightforward process that begins by identifying sites in an enzyme appropriate for saturation mutagenesis (amino acid randomization), a site being defined as comprising one, two or more amino acid positions in the protein. Following the formation of initial mutant libraries and a screening procedure specific for a given catalytic property, the best variants ("hits") are identified. The gene of a respective hit is then utilized as a template for further saturation mutagenesis at the other sites, and the iterative process is continued until the desired degree of catalyst improvement has been achieved. Different evolutionary pathways are possible and the number of chosen sites, *N*, determine the number of pathways (*N!*). When applying ISM, it is not necessary to explore all pathways, although this can be of theoretical interest. The reason for the pronounced efficacy of ISM, which has been observed in numerous cases involving different enzymes, is due to strong cooperative effects occurring between mutations within a randomized site and between two or more sites as the evolutionary process proceeds (more than additivity).[9]

The choice of the randomization sites is crucial, which in turn depends on the protein property to be evolved. For each catalytic property (e.g., thermostability, rate, stereoselectivity, etc.) a different criterion for choosing appropriate randomization sites needs to be considered. In the case of thermostabilization, thermoresistance to undesired aggregation, or robustness towards hostile organic solvents, we have developed the B-FIT method with the following protocol.[10] Atomic displacement parameters obtained from X-ray crystallography data are utilized as a guide, namely *B* factors, which reflect smearing of atomic electron densities with respect to their equilibrium po-

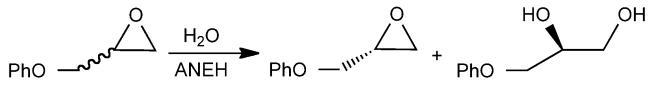
[a] *Dr. Y. Gumulya, Prof. Dr. M. T. Reetz*
Department of Synthetic Organic Chemistry
Max-Planck-Institut für Kohlenforschung
Kaiser-Wilhelm-Platz 1, 45470 Mülheim an der Ruhr (Germany)
E-mail: reetz@mpi-muelheim.mpg.de

[b] *Prof. Dr. M. T. Reetz*
Fachbereich Chemie, Philipps-Universität Marburg
Hans-Meerwein-Strasse, 35032 Marburg (Germany)

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/cbic.201100412.

sitions as a result of thermal motion and positional disorder. Rather than considering only the *B* factor of the Cα atom of a given amino acid in the protein, all atoms of a particular residue are utilized (except the H atoms). The B-FITTER computer aid, which is available free of charge on the homepage of the corresponding author (http://www.kofo.mpg.de/manfred-reetz.html), then provides the average B values of all residues. Those with high average B values indicate the highest flexibility; such data then serve as the criterion for choosing appropriate sites for saturation mutagenesis.[10] The expectation is that proper mutations thereat will increase rigidity, and therefore, enhance thermostability or thermal robustness in the more general sense. Since these residues generally occur on the protein surface, other influences, such as solvent effects and/or protein–protein interactions, might also play a role, among others. In an initial study we showed the success of B-FIT by subjecting the lipase from *Bacillus subtilis*, which is a small 181 amino acid protein, to this method.[10] Preliminary biophysical results show that the apparent increase in robustness is due to the suppression of undesired aggregation, which causes detrimental precipitation of the enzyme upon heating,[11] an unusual effect, which Rao et al. had previously discovered in their approach to stabilization of the same lipase using epPCR as the gene mutagenesis method.[12] Recently, Visser and co-workers have applied B-FIT successfully in the quest to enhance the thermostability of *E. coli* uridine phosphorylase,[13a] and in the case of a lipase the Bornscheuer group has likewise reported positive results using this method.[13b] In these studies extensive exploration of the respective ISM scheme was not necessary.

Here, we report the result of applying B-FIT to the epoxide hydrolase from *Aspergillus niger* (ANEH). In a previous study this enzyme was subjected to ISM in the form of the combinatorial active site-saturation test (CAST) as part of an effort to increase the enantioselectivity of the test reaction involving the hydrolytic kinetic resolution of racemic phenyl glycidyl ether, the selectivity factor increased from $E = 4.6$ (WT ANEH) to $E = 115$ (best mutant in favor of (*S*)-**2**).[8a]



rac-**1** (*R*)-**1** (*S*)-**2**

The purpose of the present study was not only to obtain notably more robust ANEH mutants for practical purposes, an important goal was also methodology development for more efficient directed evolution. This entailed the exploration of several evolutionary pathways in the ISM scheme—not just one—as well as the development of a strategy for escaping from what can be termed, from a practical point of view, as a "local minimum" or "dead end"; that is, what to do if a given mutant library contains no improved variants (hits). In order to measure thermal robustness of the generated mutants in a high-throughput manner, we employed T_{50}^{15} values, the temperature required to reduce the initial enzymatic activity at room temperature following a 15 min heat treatment. As a pretest for identifying active hits, Reymond's adrenaline-based assay[14] was used, followed by measurement of T_{50}^{15} values, the test reaction being the hydrolysis of *rac*-**1**. The improved variants were then isolated, purified and re-examined by subjecting them to heat treatment for 1 h, resulting in more precise T_{50}^{60} values. For this purpose medium-throughput HPLC analysis was utilized in order to monitor the hydrolysis of *rac*-**1**. On the practical side, we succeeded in evolving a significant increase in robustness, the ΔT_{50}^{60} value of the best ANEH mutant amounted to 21 °C.

Results and Discussion

The X-ray structure of wild-type (WT) ANEH (PDB ID: 1Q07), a 396 residue enzyme, has been determined at 1.8 Å resolution, and has revealed a typical α/β-hydrolase fold.[15] It is a dimer composed of two 44 kDa subunits in the asymmetric unit. The data provide a means to identify residues with the highest average *B* factors as calculated by the B-FITTER computer aid.[10b] The top 50 were ranked as summarized in Table S1 in the Supporting Information. Not all of them were considered in an ISM scheme, because such extensive exploration of the respective protein sequence space was not the goal of this study. From these 50 residues, ten sites each composed of two amino acid positions, which are scattered over the surface of the protein, were initially considered: (Glu26/Gln27), (Ser70/Glu71), (Arg105/Glu138), (Ala217/Cys350), (Met218/Leu330), (Ala220/Ser226), (Glu223/Gly224), (Ser229/Leu230), (Met329/Lys332) and (Asp368/Glu371). The distance between two residues within a given site is substantially less than 10 Å, some of them are direct neighbors. Cooperative effects acting between two such residues can be expected as the evolutionary process of B-FIT proceeds, in addition to cooperativity between the sites themselves.[8b,9,10] It should be noted that a different grouping of the 20 residues would also be possible.

Two additional sites in an unresolved region in the crystal structure of WT ANEH were also chosen, for which no *B* factors are accessible, but which can also be expected to constitute flexible regions: (Ala321/Ser322) and (Gly326/Ala327). The conventional form of saturation mutagenesis as defined by the QuikChange protocol,[16a] which is based on earlier studies regarding saturation mutagenesis,[16b–f] utilizes NNK codon degeneracy (N: adenine/cytosine/guanine/thymine; K: guanine/thymine) encoding all 20 canonical amino acids. In contrast, we discovered in previous studies that reduced amino acid alphabets lead to high-quality libraries, which entail drastically lower screening efforts.[8b,9,10b] For example, for 95% library coverage, randomization of a two-residue site by using NNK requires the screening of about 3000 transformants (assuming the absence of amino acid bias), while the use of, for example, NDT codon degeneracy (D: adenine/guanine/thymine; T: thymine) encoding only 12 amino acids (Phe, Leu, Ile, Val, Tyr, His, Asn, Asp, Cys, Arg, Ser, Gly) needs only 430 transformants.[8b,9,10b] Due to the expected amino acid bias present in most mutagenesis methods in directed evolution, including saturation mutagenesis, such calculations constitute rough but useful guides (see also the Experimental Section).

Therefore, in the present study we generally chose NDT codon degeneracy, the encoded 12 amino acids forming a well-balanced set of polar/nonpolar, aromatic/nonaromatic and charged/uncharged amino acids as building blocks. In one case NNG codon degeneracy (G: guanine) was also tested; this encodes the eight amino acids that are not covered by NDT, namely Ala, Pro, Trp, Thr, Met, Gln, Glu and Lys. In the initial round of saturation mutagenesis at all 12 sites by using NDT codon degeneracy, the most positive results originated from six sites, namely A, B, C, D, E and F (Figure 1; the letter designations A–F are arbitrary). A quality check of the libraries was performed, and revealed, not surprisingly, amino acid bias in some cases (Table S2 in the Supporting Information). Nevertheless, the evolved increases in ΔT_{50}^{60} values proved to be substantial, despite the possibility that even better hits were perhaps missed. The other six libraries contained mutants displaying essentially no improvements in the T_{50}^{15} screen, and they were not utilized in further mutagenesis experiments. The respective sites were likewise not considered in iterative saturation mutagenesis (although past experience has shown that this might in fact be successful).[9] Such decisions need not cause unease, because past and present experience shows that a given ISM study can lead to many different hits, depending on how much time is invested in exploring different sites and pathways. Since the libraries are small, requiring a minimum of screening, such mini-explorations of protein sequence space generally prove to be rewarding.[9]

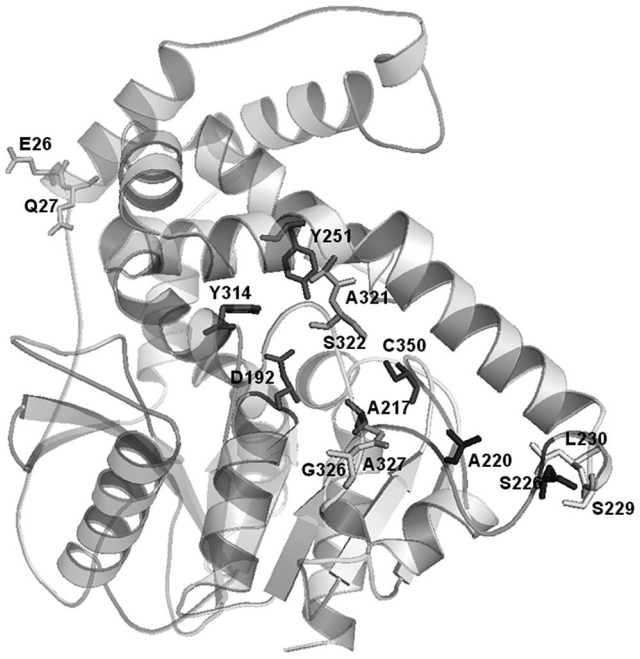


Figure 1. Location of saturation mutagenesis sites A–F in the X-ray structure of WT ANEH[15] by using NDT codon degeneracy, leading to improved mutants with enhanced thermostability. Sites displaying high average *B* factors are, C (residues 220/226), D (residues 229/230), E (residues 26/27), and F (residues 217/350). Sites A (residues 321/322) and B (residues 326/327) are in an unresolved region of the crystal structure. The catalytically active residues, Tyr251 and Tyr314, bind and activate the epoxide substrate by forming H bonds to the epoxy O atom, while Asp192 acts as the nucleophile in the rate-determining step.

In the first round of saturation mutagenesis at 12 sites with the creation of 12 mutant libraries, a total of 5495 transformants were assayed, which in each case guaranteed 95% library coverage (again, due to the bias phenomenon, this is only a rough guide). Table 1 lists the respective hits. The best mutants, GUY-003 and GUY-004, both showing the greatest improvement ($\Delta T_{50}^{60} = 8\,^{\circ}C$), originate from the library obtained by randomization at site B, which is in an unresolved part of the X-ray structure. Since B-FIT targets residues with the highest degree of flexibility as measured by the respective *B* factor, it is not surprising that most of the randomization sites are located on the surface of the protein, generally in loops. Sites D and E are located between the end of a helix and the beginning of a loop, respectively. We also note that another excellent mutant, GUY-005, with an increase in T_{50}^{60} of 7 °C, originates from site C, which has the highest average *B* factor. Thus, the hypothesis that mutagenesis should be focused on positions of highest flexibility constitutes a practical guide.

In the case of saturation mutagenesis at site D, no improved mutants were detected, although several showed thermal robustness comparable to that of WT ANEH. This could have several reasons; for example, all combinations of residues as encoded by NDT fail to provide superior mutants. Therefore, NNG codon degeneracy, encoding the eight amino acids not included in NDT, was tested. In this case, screening of 765 transformants, which ensured more than 95% coverage, led to the identification of mutant GUY-007 with a slightly improved T_{50}^{15} value of 1 °C, but upon isolation, the same T_{50}^{60} value as WT ANEH was found (Table 1). NNK codon degeneracy was not tested in this case.

In protein engineering of thermostabilization, an increase in T_{50}^{60} of 5–8 °C is generally considered to be respectable.[2–5] Nevertheless, we wanted to continue optimization by testing ISM using some, but for practical reasons, not all of the six sites A–F. In curiosity-driven experiments, we chose two extreme mutants as points for continuation, the best one and the worst one: GUY-003 originating from randomization at site B showing an improvement of 8 °C (same as mutant GUY004, which could also have been chosen), and mutant GUY-007 (site D) lacking any improvement. Two different strategies were followed. In

Table 1. Best mutants resulting from the first round of saturation mutagenesis by using NDT codon degeneracy in the randomization at sites A, B, C, E and F, and NNG codon degeneracy in the case of site D.

Site	Mutant	Mutations	T_{50}^{15} [°C]	ΔT_{50}^{15} [°C]	T_{50}^{60} [°C]	ΔT_{50}^{60} [°C]
	WT ANEH	–	50	–	46	–
A	GUY-001	A321H/S322C	56	6	51	6
	GUY-002	S321C/S322G	56	6	51	6
B	GUY-003	G326F/A327Y	58	8	54	8
	GUY-004	G326Y/A327Y	58	8	54	8
C	GUY-005	A220C	58	8	52	7
	GUY-006	A220V/S226H	56	6	50	5
D	GUY-007	S229E	51	1	46	0
E	GUY-008	E26N/Q27N	53	3	50	4
F	GUY-009	A217V/C350F	55	5	51	6
	GUY-010	A217F	54	4	50	5

the case of the mutant derived from site B as the template for further saturation mutagenesis, the conventional procedure was adhered to by considering only improved mutants, that is, "hits" in the traditional meaning. If no such mutant was found in a given library, the respective pathway was viewed as ending there and was no longer considered. In the case of the mutant identified by saturation mutagenesis at site D, improved mutants were again chosen for a subsequent mutagenesis round, but if none was identified, then an unimproved or inferior variant was chosen as a starting point for the next ISM round. Complete systematization was maintained in the build-up phase of the two options by testing five branches in each case, namely steps WT→B→A, WT→B→C, WT→B→D, WT→B→E and WT→B→F on the one hand, and WT→D→A, WT→D→B, WT→D→C, WT→D→E and WT→D→F on the other hand. Thereafter, only a few pathways were continued, the additional total screening effort amounting to 4300 transformants.

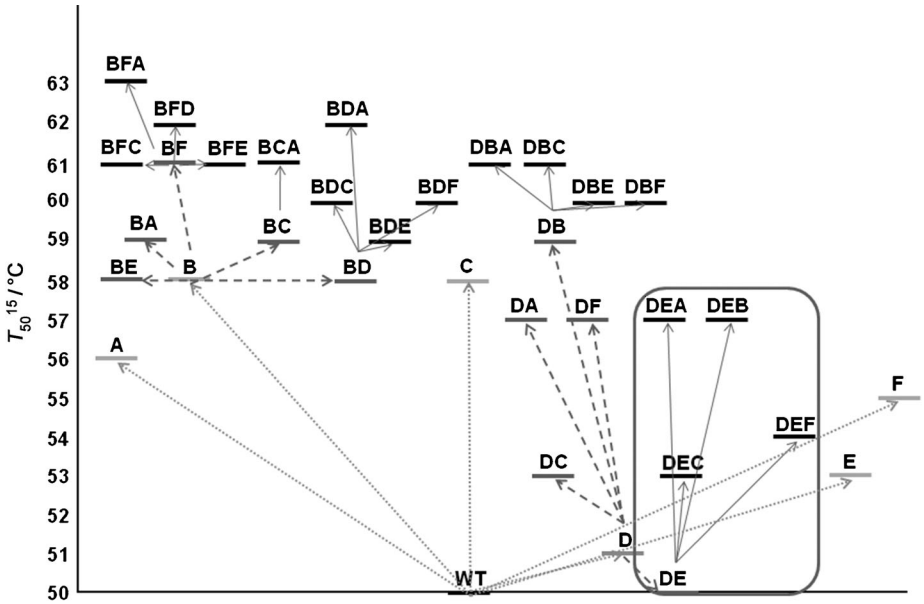


Figure 2. T_{50}^{15} values of variants found in the first, second and third rounds of ISM. Light gray bars: mutants of the first generation; dark gray bars: mutants of the second generation; black bars: mutants of the third generation.

This limited exploration led to some remarkable results as summarized in Figure 2 (Table S3 in the Supporting Information). At this point we note that many, but not all, pathways provide improved variants. Surprisingly, positive results also originate in the pathway that begins with the least thermostabilized mutant originally obtained in the library formed by randomization at site D. While the step from WT to D indicates a very slight improvement (Figure 2), isolation of the mutant and characterization by T_{50}^{60} shows that no improvement was in fact achieved (Table 1). Nevertheless, pathways WT→D→A, WT→D→B, WT→D→C and WT→D→F provided improved mutants. Interestingly, upon exploring pathway WT→D→E, protein robustness (or "thermoresistance" as a referee has suggested) failed to be improved in the early stages of the evolutionary pathway. This means that the respective epistatic effects are either antagonistic or non-additive. Normally, one would abandon such a pathway, that is, as soon as a library lacking improved mutants is encountered. It is a "dead end" or "local minimum" as judged from a practical point of view, whatever the reason(s) might be. These could include insufficient screening, inherent amino acid bias, imperfect PCR conditions or an actual local minimum on the theoretical fitness landscape. Contrary to conventional wisdom in the field of directed evolution,[3,6,9] application of the second strategy based on the use of non-improved or even inferior mutants in ISM led to superior mutants in the third generation (Figure 2). This interesting phenomenon might be related to the theory of neutral drift[17] and/or to Eigen's concept of quasi-species in natural evolution,[18] which Mannervik has invoked previously in several important directed evolution studies.[19] Likewise, our previous study regarding the enhancement of activity and enantioselectivity of an enoate reductase, in which it was discovered that it is not always optimal to pick the very best variant for the next ISM round, becomes relevant, specifically when two different catalytic parameters are being optimized (rate and stereoselectivity).[20]

Inspired by these results, we decided to explore additional pathways defined by mutants obtained by saturation mutagenesis at sites B and D, this time following the second strategy in both cases, that is, just continuing with an inferior mutant if a given library failed to contain hits in the conventional sense. Again, unusual observations were made. As already noted, pathway WT→B→C→A provides a mutant with a respectable T_{50}^{60} value of 59 °C, but further ISM failed to improve thermal robustness (Figure 3 and the Supporting Information). In the case of WT→B→F→A→C→E→D, the situation is different, because it was possible to escape from two local minima simply by continuing with non-improved mutants, and the final randomization at site D led to a better mutant (T_{50}^{60} = 60 °C).

Importantly, when starting from the least improved original mutant GUY-007 (site D), pathway WT→D→E→C failed to achieve any positive results as already pointed out, but continuing on this seemingly hopeless pathway according to WT→D→E→C→F→A→B, proved to be highly successful, and provided the best mutant of the present study. An increase in thermal robustness amounting to ΔT_{50}^{60} = 21 °C was achieved, which is a high score in directed evolution of thermally robust proteins.

An ISM scheme based on six sites, A, B, C, D, E, and F, as in the present study, has so many different pathways (720) that exploring all of them, requires excessive experimental work. In

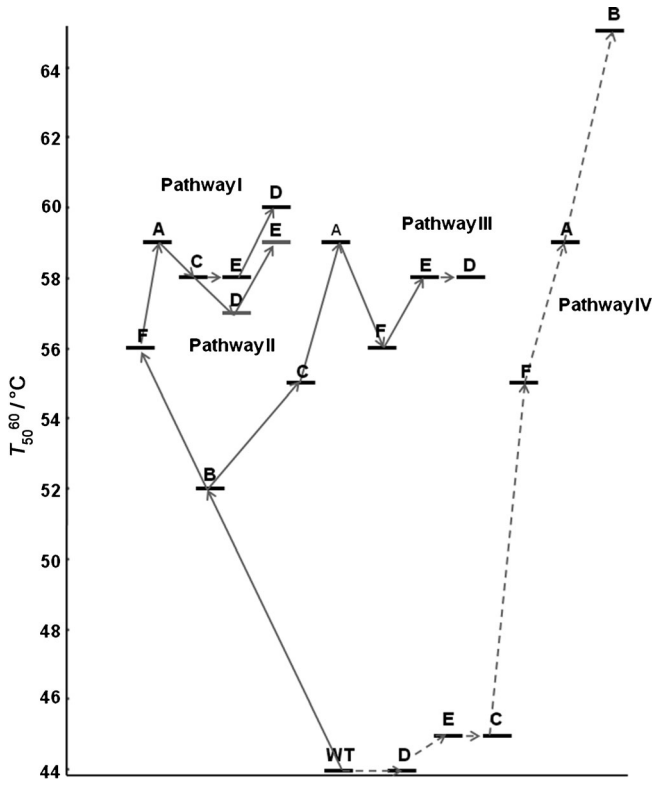


Figure 3. Results of limited ISM exploration starting from the best mutant, GUY-003 (site B), and the worst mutant, GUY-007, in the initial round of saturation mutagenesis at sites A–F. In all cases NDT codon degeneracy was used except when performing saturation mutagenesis at site D, in which case NNG codon degeneracy was applied.

the present study only four complete upward pathways were studied, namely WT→B→F→A→C→E→D, WT→B→F→A→C→D→E, WT→B→C→A→F→E→D and WT→D→E→C→F→A→B (Figure 3), the final mutants were GUY-055, GUY-057, GUY-060, and GUY-066, and showed T_{50}^{60} values of 60, 59, 58 and 65 °C, respectively. The sequences and thermostability data are summarized in Table 2. It can be seen that all four pathways lead to notably improved mutants as measured by ΔT_{50}^{60} values ranging between 14 and 21 °C. It is also evident that the point mutations are quite different in nature, which means that the effects on a molecular level must also be diverse.

Analysis of thermal robustness of WT ANEH and the four selected mutants at temperatures 55, 60, 65 and 70 °C enable the formal estimation of half-lives and the determination of inactivation energies (E_a). These data were collected for WT and two mutants, GUY-060 and GUY-066. The improvements are clearly evident (Table 3). For example, the thermal E_a improves from 39.7 kcal mol^{-1} (WT) to 71 kcal mol^{-1} (best mutant GUY-066). Further data corroborating the increase in thermal robustness are collected in the Supporting Information, namely specific activity of WT ANEH and of mutants GUY-060 and GUY-066 at different temperatures (Figure S1 in the Supporting Information) as well as residual activity (%) of these two mutants as a function of temperature (Figure S2 in the Supporting Information).

In the absence of detailed biophysical studies, it is difficult to interpret the results on a molecular level. One of several possibilities, although not proven by current data, concerns the introduction of cysteine residues in all of the improved mutants, these perhaps lead to stabilizing disulfide bridges.

Other strategies for improving thermostability

Numerous studies of structural comparisons between mesophilic and thermophilic proteins have revealed the importance of salt bridges, sometimes in networks, in promoting thermostability.[3–5,21] Therefore, in addition to directed evolution based on iterative B-FIT libraries, as described above, we created several libraries for introducing salt bridges or removing unfavorable electrostatic interactions. The residues chosen for library creation were Arg36, Ala41, Gln49, Phe54, Ser58, Leu61, Asn82, Glu104, Gln129, Gln175, and Glu292. The codon degeneracy used for these libraries was RRK, which encodes for acidic (Asp, Glu) and basic (Arg, Lys) amino acids. However, no improved mutants were found in these libraries[22] (data not shown).

The N and C termini of protein chains are normally known to adopt flexible conformations, and some examples in the literature show that shortening the termini and/or anchoring

Table 3. Estimated half-life ($t_{1/2}$) and thermal inactivation energy (E_a) of WT ANEH and mutants at different temperatures.

Mutant	$t_{1/2}$ [min]					E_a
	60 °C	62 °C	65 °C	68 °C	70 °C	[kcal mol^{-1}]
WT	≈1	≈1	–	–	–	39.7
GUY-060	24	21	7	3	2	62.9
GUY-066	80	77	16	6	4	71.0

Table 2. Characterization of the final isolated and purified ANEH mutants arising from the four fully transversed pathways.

Pathway	Mutant	Mutations						ΔT_{50}^{60} [°C]
WT→B→F→A→C→D→E	GUY-055	A321G	G326F	A220C	S229M	E26L	A217V	16
		S322C	A327Y	S226N		Q27R	C350F	
WT→B→F→A→C→E→D	GUY-057	A321G	G326F	A220C	S229E	E26N	A217V	15
		S322C	A327Y	S226N		Q27H	C350F	
WT→B→C→A→F→E→D	GUY-060	A321N	G326F	A220C	S229E	E26G	A217V	14
		S322C	A327Y	S226L		Q27F	C350F	
WT→D→E→C→F→A→B	GUY-066	A321L	G326C	A220C	S229E	E26Y	A217C	21
		S322C	A327G	S226G		Q27L		

them increases thermostability.[23] We, therefore, performed saturation mutagenesis at residues Lys3 and Glu138 (N termini) as well as Asp107, Trp396, Val109 and Val392 (C termini) using NDT codon degeneracy. However, again positive results were not achieved[22] (data not shown).

Conclusion

Iterative saturation mutagenesis (ISM) has emerged as a powerful method in protein engineering, and makes directed evolution more efficient, and therefore, faster than in the past.[9] We, and others, have applied this concept to a number of different proteins.[9,24] Efficacy is increased even more when using reduced amino alphabets.[8b,9] Smaller, and at the same time, higher-quality libraries result from these techniques, thereby drastically reducing the screening effort, which is traditionally the bottleneck of directed evolution. So far, the two embodiments of ISM are B-FIT for enhancing thermal robustness,[10] and CASTing[8,9] for manipulating substrate scope (rate) and stereoselectivity. However, it is not always clear which pathway in an ISM scheme should be chosen, and what to do if a local minimum in the respective fitness landscape is encountered, that is, the absence of any improved variants in a given library. In the case of CASTing, several previous studies have shown that many pathways lead to improved mutants.[8,9] Nevertheless, a complete ISM scheme has yet to be explored in the quest to answer the question as to how many of the theoretically possible upward pathways do in fact provide notably improved mutants in terms of activity and/stereoselectivity.

Relative to the widely used iterative CASTing,[9] thus far B-FIT has not been applied as often.[10,13] In the present study, we resorted to ISM in the form of B-FIT in an effort to increase the thermal robustness (thermoresistance) of the epoxide hydrolase from *Aspergillus niger* (ANEH), a 396 amino acid protein. Ten residues characterized by the highest average *B* factors and two unresolved residues were grouped into six sites for saturation mutagenesis, each comprising two amino acid positions. The initial six libraries were created by using conventional NNK codon degeneracy encoding all 20 canonical amino acids. In order to reduce the screening effort while still maintaining 95% library coverage, reduced amino alphabets were used in subsequent randomization experiments, in most cases NDT encoding 12 amino acids, and in one instance NNG encoding the eight amino acids not covered by NDT.

A limited amount of ISM exploration was performed, and led to several notably improved mutants characterized by ΔT_{50}^{60} values in the range of 14 to 21 °C. This underscores the viability of B-FIT. In four cases the complete six-step ISM pathways were constructed and evaluated. In several instances local minima were encountered, meaning that in a given library no improved hits were detected. Normally, one would terminate the exploration of such a pathway, the only possibilities being backtracking or probing a different pathway.[9] In the present study we developed an unconventional strategy by testing unimproved and even inferior mutants in such "dead-end" libraries as templates in subsequent mutagenesis steps. Surprisingly, continuation on a pathway of this type did in fact deliver improved mutants. Thus, such a strategy constitutes the simplest technique for escaping from a "local minimum", whatever the reasons for the occurrence of such a "dead-end" might be. It opens up new perspectives for directed evolution based on ISM, but also for protein engineering in general by using other techniques, such as error-prone PCR or DNA shuffling. Indeed, in preliminary work we have observed that this strategy is also successful when optimizing the stereoselectivity of enzymes. These unusual observations could be related to the phenomenon of neutral drift[17] and/or to Eigen's hypothesis of quasispecies in natural evolution,[18] as recently invoked by Mannervik in a series of seminal papers regarding the directed evolution of other enzymes (glutathione transferases).[19]

While several arbitrarily chosen pathways as studied in the present investigation proved to be successful, more systematic work is necessary in order to fully illuminate the nature of ISM as a general method in directed evolution. The exploration of a complete ISM scheme would shed additional light on this important issue, be it for increasing thermostabilization, substrate acceptance or stereoselectivity. Nevertheless, for practical purposes when attempting to solve a specific biocatalyst problem, this is not necessary. We also note that many of the mutations in the evolved variants described herein differ considerably in nature, which means that future theoretical and biophysical work needs to focus on the source of increased thermal robustness of the ANEH mutants on a molecular level. The effect of amino acid bias and of imperfect PCR conditions in general as well as the question of how residues are optimally grouped into sites when performing ISM, likewise constitute current issues of interest in our laboratory. Success is likely even if such parameters are ignored, but optimization can be expected to increase efficacy even more.

Experimental Section

Materials: Glycidyl phenyl ether (*rac*-**1**) was purchased from Acros (Geel, Belgium).

Calculation of oversampling and % coverage of libraries: All numbers regarding the degree of oversampling as a function of % coverage of a mutant library and codon usage were calculated by using the computer aids B-FIT or CASTER, available free of charge on the homepage of the corresponding author (http://www.kofo.mpg.de/manfred-reetz.html). These programs are based on statistical algorithms published previously by Patrick and Firth assuming the absence of amino acid bias;[25a] for related algorithms, see the papers by Ostermeier[25b] and by Pelletier.[25c]

Mutant library generation: ANEH mutant libraries were created by QuikChange™ saturation mutagenesis method[16a] (Stratagene). Each PCR reaction contained (50 μL final volume): 10×KOD buffer (Novagen, 5 μL), dNTPs (2 mM each, 5 μL), $MgCl_2$ (25 mM, 2 μL), degenerate primers (forward and reverse 2.5 μM, 2.5 μL), template plasmid (20–50 ng μL^{-1}, 0.5 μL), KOD Hot start polymerase (0.5 U). The thermal cycler program consisted of 1×95 °C 3 min, 20×95 °C 30 s, gradient 1 min, 72 °C 7 min, and 1×72 °C 16 min. The PCR product was transformed into *E. coli* DH5α after DpnI treatment (New England Biolabs, 1 U, 37 °C, 2 h, 2×digestion in buffer provided by the supplier). The qualities of the libraries generated and the ANEH mutants were analyzed by sequencing with primers pQE-for

(GTATC ACGAG GCCCT TTCGT CT) or pQE-rev (CATTA CTGGA TCTAT CAACA GGAG; Eurofins MWG Operon, Ebersberg, Germany).

Expression of gene libraries: Colonies were picked with colony picker QPIX (Genetix, New Milton, UK), inoculated into 96-deep-well plates containing LB medium (800 μL per well; supplemented with 100 $\mu g\,mL^{-1}$ carbenicillin), and incubated, overnight, at 37 °C with shaking (800 rpm). An aliquot (100 μL) of this overnight culture was transferred into a glycerol plate (96-well plate filled with 100 μL glycerol solution 70% each well) and stored at −80 °C. Another aliquot (100 μL) was used to inoculate fresh medium and grown for 24 h at 30 °C with shaking at 800 rpm. The cells were harvested by centrifugation (1055 *g*, 10 min), the supernatant was removed and the cells were resuspended by adding lysis buffer (400 μL). The cell lysate was shaken at 37 °C for 1 h before being stored at −20 °C. The cells were thawed the following day by being shaken at 37 °C for 1 h and centrifuged at 1055 *g* for 1 h.

Screening for thermal robustness: The screening for improved thermal robustness was performed in 96-well PCR plates by heating enzyme solutions (cell free lysate) at 50 °C for 15 min in a PCR thermocycler (Whatman Biometra, Göttingen, Germany). After heat treatment, the residual enzyme activity was measured by using glycidyl phenyl ether (*rac*-**1**; GPE) as substrate. Heat treated cell free lysates (22 μL) were transferred to a 96-well UV/Vis plate reader; phosphate buffer saline (145 μL; pH 7.2, 10 mM) and substrate solution (13 μL; 125 mg solubilized in 20 mL acetonitrile) were added. After being shaken for 1 h at 30 °C (800 rpm), the reaction was stopped by adding $NaIO_4$ (20 μL, 25 mM). The reaction was further developed for another 30 min. A blank measurement was first carried out ($A_{490\,cells}$ = absorbance of cell free lysate+buffer+solvent+substrate+$NaIO_4$). The adenochrome absorption was then measured after adding L-adrenaline (20 μL, 27.5 mM; $A_{490\,adren}$ = absorbance of cell free lysate+buffer+solvent+substrate+$NaIO_4$+L-adrenaline). The same reaction was also performed for the untreated samples (without heat treatment). The thermostability was assessed by measuring the residual activity after exposure to high temperatures. For each assay (12 different temperatures), temperature was plotted against residual activity (see the Supporting Information). Using this graph, the temperature, *T*, at which 50% residual activity pertains was derived.

Screening scale reproduction: Positive clones were confirmed by regrowing them in 96-deep-well plates, overnight, at 37 °C, 800 rpm. Then an aliquot (500 μL) of this overnight culture was transferred into a culture tube (12 mL) filled with LB medium (5 mL), supplemented with carbenicillin (100 $\mu g\,mL^{-1}$), and incubated for 24 h at 30 °C at 800 rpm. The cells were harvested by centrifugation at 1055 *g* for 10 min; lysis buffer (2.5 mL) was then added and the cells were shaken for 1 h at 37 °C before being stored at −20 °C. Cells were thawed the following day by being shaken at 37 °C (800 rpm) for 1 h and then centrifuged at 1055 *g* for 1 h. The cell free lysate was heat-shocked in a gradient PCR thermocycler (45–65 or 70 °C). The enzyme residual activity was measured, as described above.

Expression and purification of thermostable mutants: The vector carrying the ANEH gene (pQE-60, Qiagen) has His_6 coding sequence at the gene C terminus. The stop codons of pQEEH, WT ANEH and mutants were deleted to allow transcription to be continued into the polyhistidine tagged coding sequence. Site-directed mutagenesis was performed according to the improved PCR method by using the mutagenic primers: GAGCA GGTGT GGCAG AAGAG ATCTC ATCAC C and CCTCG CTCTG CTAAT CCTGT TACCA GTGGC (Invitrogen). WT ANEH and mutants were over-expressed in *E. coli* DH5α. An aliquot (20 mL) of overnight culture was inoculated into expression culture (200 mL, TB medium supplemented with 100 $\mu g\,mL^{-1}$ carbenicillin). The culture was grown for 24 h at 30 °C (200 rpm). Cells were harvested by centrifugation at 1650 *g* for 10 min at 4 °C. The cell pellet (about 4 g wet mass) was resuspended in lysis buffer (40 mL, with 0.5 M NaCl, 20 mM imidazole). The cell suspension was disrupted by sonication (Bandelin, 60 s, six times, 40% pulse, on ice). The cell debris was removed by centrifugation at 9500 *g* for 1 h at 10 °C. The clarified lysate was firstly filtered (0.22 μm filter) before being loaded onto 5 mL HisTrap FF affinity column (GE Healthcare). His-tagged enzymes were eluted with an imidazole linear gradient by using ÄKTA purifier. The fraction containing the ANEH (10 mL) was concentrated by using ultrafiltration centrifugal filter (10 kDa cut-off membrane, Amicon) to a final volume of 3 mL, and was applied to a PD-10 desalting column in order to remove excess imidazole. The eluted protein was diluted with PBS buffer (9 mL; 10 mM, pH 7.2), filter-sterilized, and stored at 4 °C. Enzyme purity was monitored by SDS-PAGE. The concentration of the ANEH mutants was measured spectrophotometrically (the absorbance of a 1 $mg\,mL^{-1}$ solution at 280 nm is 1.428 based on a molar extinction coefficient of $\varepsilon_{280\,nm} = 64\,400$).

Analysis of the thermal robustness of purified ANEHs: The heat treatment of purified ANEHs was carried out in thin-walled PCR tubes (0.2 mL) by using a PCR thermocycler to allow precise temperature control. The purified ANEHs (70 μL, 0.012 $\mu g\,\mu L^{-1}$ in 20 mM sodium phosphate buffer pH 7.2) was incubated at various temperatures (e.g., 50, 55, 60, 65, and 70 °C) for a defined period of time (e.g., 60 or 120 min). The samples were cooled on ice and equilibrated at room temperature before the residual activity was measured. Purified ANEH (56 μL), with or without heat treatment, was transferred into 96-well plates and reaction mixture (124 μL; 26 μL glycidyl phenyl ether (41.62 mM solubilized in acetonitrile HPLC grade) + 98 μL sodium phosphate buffer, 20 mM, pH 7.2) was added into the protein solution. The reaction plate was shaken at 30 °C for 1 h and then centrifuged at 1055 *g* for 20 min. Then the reaction mixture (150 μL) was transferred into new reaction plates and mixed with the same volume of internal standard solution (*R*)-(+)-1-phenyl-1-butanol (6.66 mM in HPLC-grade methanol). This mixture was further centrifuged for another 30 min at 1055 *g*. This solution (200 μL) was finally transferred to 96-well microtiter plates and the conversion was measured by HPLC. The chiral analysis of the hydrolytic kinetic resolution of *rac*-gylcidyl phenyl ether (*rac*-**1**) were performed by using Kromasil 3-Cellucoat RP 100 chiral column (4.6 mm i.d, Akzo Nobel, Sweden) and Kromasil 5-Cellucoat 10 (4.6 mm i.d.) precolumn. The conditions of analysis were as follows: methanol/H_2O 70:30, 1.0 $mL\,min^{-1}$, UV 220 nm, 298 K. The corresponding retention times were, (*R*)-**2**: 1.9 min; (*S*)-**2**: 2.1 min; (*R*)-**1**: 4.1 min; (*S*)-**1**: 4.7 min.

The inactivation energy (E_a) of the mutants were measured by using Arrhenius plots of the rate of thermal inactivation ($\ln K_d$, min^{-1}) at different temperatures versus the reciprocal of the absolute temperature (1000/*T*, K^{-1}).

Acknowledgements

Support from the Arthur C. Cope Foundation (USA) is gratefully acknowledged.

Keywords: directed evolution · enzymes · epoxide hydrolases · quasi-species · neutral drift

[1] a) K. Faber, *Biotransformations in Organic Chemistry*, 5th ed., Springer, Berlin, **2004**; b) K. Drauz, H. Waldmann, *Enzyme Catalysis in Organic Synthesis: A Comprehensive Handbook, Vols. I – III*, 2nd ed., Wiley-VCH, Weinheim, **2002**; c) A. Liese, K. Seelbach, C. Wandrey, *Industrial Biotransformations*, 2nd ed., Wiley-VCH, Weinheim, **2006**; d) J. Tao, G.-Q. Lin, A. Liese, *Biocatalysis for the Pharmaceutical Industry*, Wiley-VCH, Weinheim, **2009**; e) V. Gotor, I. Alfonso, E. García-Urdiales, *Asymmetric Organic Synthesis with Enzymes*, Wiley-VCH Weinheim, **2008**.

[2] a) T. Oshima, *Curr. Opin. Struct. Biol.* **1994**, *4*, 623 – 628; b) C. Ó'Fágáin, *Enzyme Microb. Technol.* **2003**, *33*, 137 – 149; c) V. G. H. Eijsink, A. Bjørk, S. Gåseidnes, R. Sirevåg, B. Synstad, B. van den Burg, G. Vriend, *J. Biotechnol.* **2004**, *113*, 105 – 120.

[3] Reviews of directed evolution of thermal robustness of proteins: a) P. L. Wintrode, F. H. Arnold, *Adv. Protein Chem.* **2001**, *55*, 161 – 225; b) V. G. H. Eijsink, S. Gaseidnes, T. V. Borchert, B. van den Burg, *Biomol. Eng.* **2005**, *22*, 21 – 30; c) A. Martin, F. X. Schmid, V. Sieber in *Directed Enzyme Evolution: Screening and Selection Methods, Vol. 230*, (Eds.: F. H. Arnold, G. Georgiou), Humana, Totowa, New Jersey, **2003**, pp. 57 – 70; d) N. Amin, A. D. Liu, S. Ramer, W. Aehle, D. Meijer, M. Metin, S. Wong, P. Gualfetti, V. Schellenberger, *Protein Eng. Des. Sel.* **2004**, *17*, 787 – 793; e) R. A. Chica, N. Doucet, J. N. Pelletier, *Curr. Opin. Biotechnol.* **2005**, *16*, 378 – 384.

[4] Examples of the use of T_{50} values or simply of data pertaining to residual activity after heat treatment at a given temperature: a) L. Sun, I. P. Petrounia, M. Yagasaki, G. Bandara, F. H. Arnold, *Protein Eng.* **2001**, *14*, 699 – 704; b) B. Morawski, S. Quan, F. H. Arnold, *Biotechnol. Bioeng.* **2001**, *76*, 99 – 107; c) J. Hao, A. Berry, *Protein Eng. Des. Sel.* **2004**, *17*, 689 – 697; d) N. Declerck, M. Machius, P. Joyet, G. Wiegand, R. Huber, C. Gaillardin, *Protein Eng.* **2003**, *16*, 287 – 293; e) S. Akanuma, A. Yamagishi, N. Tanaka, T. Oshima, *Protein Sci.* **1998**, *7*, 698 – 705; f) H. Zhao, F. H. Arnold, *Protein Eng.* **1999**, *12*, 47 – 53; g) B. Van den Burg, A. de Kreij, P. Van der Veek, J. Mansfeld, G. Venema, *Biotechnol. Appl. Biochem.* **1999**, *30*, 35 – 40; h) K.-H. Oh, S.-H. Nam, H.-S. Kim, *Protein Eng.* **2002**, *15*, 689 – 695; i) M. G. Pikkemaat, A. B. M. Linssen, H. J. C. Berendsen, D. B. Janssen, *Protein Eng.* **2002**, *15*, 185 – 192; j) L.-J. Wang, X.-D. Kong, H.-Y. Zhang, X.-P. Wang, J. Zhang, *Biochem. Biophys. Res. Commun.* **2000**, *276*, 346 – 349; k) J. Mansfeld, G. Vriend, B. W. Dijkstra, O. R. Veltman, B. Van den Burg, G. Venema, R. Ulbrich-Hofmann, V. G. H. Eijsink, *J. Biol. Chem.* **1997**, *272*, 11152 – 11156; l) E. H. Muslin, S. E. Clark, C. A. Henson, *Protein Eng.* **2002**, *15*, 29 – 33; m) O. Turunen, M. Vuorio, F. Fenel, M. Leisola, *Protein Eng.* **2002**, *15*, 141 – 145; n) T. W. Johannes, R. D. Woodyer, H. Zhao, *Appl. Environ. Microbiol.* **2005**, *71*, 5728 – 5734; o) J.-H. Kim, G.-S. Choi, S.-B. Kim, W.-H. Kim, J.-Y. Lee, Y.-W. Ryu, G.-J. Kim, *J. Mol. Catal. B* **2004**, *27*, 169 – 175; p) D. E. Stephens, K. Rumbold, K. Permaul, B. A. Prior, S. Singh, *J. Biotechnol.* **2007**, *127*, 348 – 354.

[5] Recent examples of protein engineered enzyme thermostabilization: a) M. J. McLachlan, T. W. Johannes, H. Zhao, *Biotechnol. Bioeng.* **2008**, *99*, 268 – 274; b) A. Nakamura, Y. Takakura, N. Sugimoto, N. Takaya, K. Shiraki, T. Hoshino, *Biosci. Biotechnol. Biochem.* **2008**, *72*, 2467 – 2471; c) K. Hirokawa, A. Ichiyanagi, N. Kajiyama, *Appl. Microbiol. Biotechnol.* **2008**, *78*, 775 – 781; d) L. Foit, G. J. Morgan, M. J. Kern, L. R. Steimer, A. A. von Hacht, J. Titchmarsh, S. L. Warriner, S. E. Radford, J. C. A. Bardwell, *Mol. Cell* **2009**, *36*, 861 – 871; e) H. Jochens, D. Aerts, U. T. Bornscheuer, *Protein Eng. Des. Sel.* **2010**, *23*, 903 – 909; f) A. S. Bommarius, J. K. Blum, M. J. Abrahamson, *Curr. Opin. Chem. Biol.* **2011**, *15*, 194 – 200; g) E. García-Ruiz, D. Maté, A. Ballesteros, A. T. Martinez, M. Alcade, *Microb. Cell Fact.* **2010**, *9*, 17; h) J. C. Joo, S. P. Pack, Y. H. Kim, Y. J. Yoo, *J. Biotechnol.* **2011**, *151*, 56 – 65; i) D. Gessmann, F. Mager, H. Naveed, T. Arnold, S. Weirich, D. Linke, J. Liang, S. Nussberger, *J. Mol. Biol.* **2011**; DOI: 10.1016/j.jmb.2011.07.054.

[6] a) S. Lutz, W. M. Patrick, *Curr. Opin. Biotechnol.* **2004**, *15*, 291 – 297; b) M. R. Parikh, I. Matsumura, *J. Mol. Biol.* **2005**, *352*, 621 – 628; c) J. M. Joern, P. Meinhold, F. H. Arnold, *J. Mol. Biol.* **2002**, *316*, 643 – 656; d) T.-W. Wang, H. Zhu, X.-Y. Ma, T. Zhang, Y.-S. Ma, D.-Z. Wei, *Mol. Biotechnol.* **2006**, *34*, 55 – 68; Recent reviews of directed evolution: e) T. W. Johannes, H. Zhao, *Curr. Opin. Microbiol.* **2006**, *9*, 261 – 267; f) S. Lutz, U. T. Bornscheuer, *Protein Engineering Handbook, Vols. 1 – 2*, Wiley-VCH, Weinheim, **2009**; g) N. J. Turner, *Nat. Chem. Biol.* **2009**, *5*, 567 – 573; h) C. Jäckel, P. Kast, D. Hilvert, *Annu. Rev. Biophys. Biomol. Struct.* **2008**, *37*, 153 – 173; i) S. Bershtein, D. S. Tawfik, *Curr. Opin. Chem. Biol.* **2008**, *12*, 151 – 158; j) P. A. Romero, F. H. Arnold, *Nat. Rev. Mol. Cell Biol.* **2009**, *10*, 866 – 876; k) M. T. Reetz in *Asymmetric Organic Synthesis with Enzymes* (Eds.: V. Gotor, I. Alfonso, E. García-Urdiales), Wiley-VCH, Weinheim, **2008**, pp. 21 – 63; l) L. G. Otten, F. Hollmann, I. W. C. E. Arends, *Trends Biotechnol.* **2010**, *28*, 46 – 54; m) A. V. Shivange, J. Marienhagen, H. Mundhada, A. Schenk, U. Schwaneberg, *Curr. Opin. Chem. Biol.* **2009**, *13*, 19 – 25.

[7] a) J.-L. Reymond, *Enzyme Assays: High-Throughput Screening, Genetic Selection and Fingerprinting*, Wiley-VCH, Weinheim, **2006**; b) Y. L. Boersma, M. J. Dröge, W. J. Quax, *FEBS J.* **2007**, *274*, 2181 – 2195; c) J.-L. Reymond, V. S. Fluxà, N. Maillard, *Chem. Commun.* **2009**, 34 – 46.

[8] a) M. T. Reetz, L.-W. Wang, M. Bocola, *Angew. Chem.* **2006**, *118*, 1258 – 1263; *Angew. Chem. Int. Ed.* **2006**, *45*, 1236 – 1241; b) M. T. Reetz, D. Kahakeaw, J. Sanchis, *Mol. BioSyst.* **2009**, *5*, 115 – 122.

[9] Review of iterative saturation mutagenesis for evolving stereoselective enzymes: M. T. Reetz, *Angew. Chem.* **2011**, *123*, 144 – 182; *Angew. Chem. Int. Ed.* **2011**, *50*, 138 – 174.

[10] a) M. T. Reetz, J. D. Carballeira, A. Vogel, *Angew. Chem.* **2006**, *118*, 7909 – 7915; *Angew. Chem. Int. Ed.* **2006**, *45*, 7745 – 7751; b) M. T. Reetz, J. D. Carballeira, *Nat. Protoc.* **2007**, *2*, 891 – 903.

[11] W. Augustyniak, A. A. Brzezinska, T. Pijning, H. Wienk, R. Boelens, B. Dijkstra, M. T. Reetz, unpublished results.

[12] a) P. Acharya, E. Rajakumara, R. Sankaranarayanan, N. M. Rao, *J. Mol. Biol.* **2004**, *341*, 1271 – 1281; b) S. Ahmad, B. M. Rao, *Protein Sci.* **2009**, *18*, 1183 – 1196.

[13] a) D. F. Visser, F. Hennessy, J. Rashamuse, B. Pletschke, D. Brady, *J. Mol. Catal. B* **2011**, *68*, 279 – 285; b) H. Jochens, D. Aerts, U. T. Bornscheuer, *Protein Eng. Des. Sel.* **2010**, *23*, 903 – 909.

[14] a) D. Wahler, J.-L. Reymond, *Angew. Chem.* **2002**, *114*, 1277 – 1280; *Angew. Chem. Int. Ed.* **2002**, *41*, 1229 – 1232; b) D. Wahler, O. Boujard, F. Lefèvre, J.-L. Reymond, *Tetrahedron* **2004**, *60*, 703 – 710; c) F. Cedrone, T. Bhatnagar, J. C. Baratti, *Biotechnol. Lett.* **2005**, *27*, 1921 – 1927; d) D. Kahakeaw, M. T. Reetz, *Chem. Asian J.* **2008**, *3*, 233 – 238.

[15] a) J. Y. Zou, B. M. Hallberg, T. Bergfors, F. Oesch, M. Arand, S. L. Mowbray, T. A. Jones, *Structure* **2000**, *8*, 111 – 122; b) http://www.pdb.ccdc.cam.ac.uk/pdb.

[16] a) H. H. Hogrefe, J. Cline, G. L. Youngblood, R. M. Allen, *BioTechniques* **2002**, *33*, 1158 – 1165; selected early developments of saturation mutagenesis methods and various applications: b) J. A. Wells, M. Vasser, D. B. Powers, *Gene* **1985**, *34*, 315 – 323; c) M. S. Z. Horwitz, L. A. Loeb, *Proc. Natl. Acad. Sci. USA* **1986**, *83*, 7405 – 7409; d) A. R. Oliphant, A. L. Nussbaum, K. Struhl, *Gene* **1986**, *44*, 177 – 183; e) J. F. Reidhaar-Olson, R. T. Sauer, *Science* **1988**, *241*, 53 – 57; f) L. B. Evnin, J. R. Vasquez, C. S. Craik, *Proc. Natl. Acad. Sci. USA* **1990**, *87*, 6659 – 6663; g) J. C. Hu, E. K. O'Shea, P. S. Kim, R. T. Sauer, *Science* **1990**, *250*, 1400 – 1403; h) S. Kamtekar, J. M. Schiffer, H. Xiong, J. M. Babik, M. H. Hecht, *Science* **1993**, *262*, 1680 – 1685; i) D. M. Kegler-Ebo, C. M. Docktor, D. DiMaio, *Nucleic Acids Res.* **1994**, *22*, 1593 – 1599; j) G. Chen, I. Dubrawsky, P. Mendez, G. Georgiou, B. L. Iverson, *Protein Eng.* **1999**, *12*, 349 – 356; first example of saturation mutagenesis in the quest to enhance enantioselectivity: k) M. T. Reetz, S. Wilensek, D. Zha, K.-E. Jaeger, *Angew. Chem.* **2001**, *113*, 3701 – 3703; *Angew. Chem. Int. Ed.* **2001**, *40*, 3589 – 3591.

[17] a) S. G. Peisajovich, D. S. Tawfik, *Nat. Methods* **2007**, *4*, 991 – 994; b) M. A. DePristo, *HFSP J.* **2007**, *1*, 94 – 98; c) J. D. Bloom, F. H. Arnold, *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 9995 – 10000.

[18] M. Eigen, J. McCaskill, P. Schuster, *J. Phys. Chem.* **1988**, *92*, 6881 – 6891.

[19] a) L. O. Emren, S. Kurtovic, A. Runarsdottir, A.-K. Larsson, B. Mannervik, *Proc. Natl. Acad. Sci. USA* **2006**, *103*, 10866 – 10870; b) S. Kurtovic, B. Mannervik, *Biochemistry* **2009**, *48*, 9330 – 9339; c) B. Mannervik, A. Runarsdottir, *FEBS Lett.* **2010**, *584*, 2565 – 2571; d) A. Runarsdottir, B. Mannervik, *J. Mol. Biol.* **2010**, *401*, 451 – 464.

[20] D. J. Bougioukou, S. Kille, A. Taglieber, M. T. Reetz, *Adv. Synth. Catal.* **2009**, *351*, 3287 – 3305.

[21] a) S. Kumar, C. Tsai, B. Ma, R. Nussinov, *J. Biol. Struct. Dynam.* **2000**, *11*, 79 – 85; b) A. H. Elcock, *J. Mol. Biol.* **1998**, *284*, 489 – 502; c) C. Vetriani, D. L. Maeder, N. Tolliday, K. S.-P. Yip, T. J. Stillman, K. L. Britton, D. W. Rice, H. H. Klump, F. T. Robb, *Proc. Natl. Acad. Sci. USA* **1998**, *95*, 12300 – 12305; d) J. H. G. Lebbink, S. Knapp, J. van der Oost, D. Rice, R. Ladenstein, W. M. de Vos, *J. Mol. Biol.* **1999**, *289*, 357 – 369; e) L. Xiao, B. Honig, *J. Mol. Biol.* **1999**, *289*, 1435 – 1444; f) G. I. Makhatadze, V.-V. Loladze, D. N. Ermolenko, X. Chen, S. T. Thomas, *J. Mol. Biol.* **2003**, *327*, 1135 – 1148; g) B. N. Dominy, H. Minoux, C. L. Brooks, *Proteins Struct. Funct.*

Bioinf. **2004**, *57*, 128 – 141; h) M. Robinson-Rechavi, A. Alibés, A. Godzik, *J. Mol. Biol.* **2006**, *356*, 547 – 557.
[22] Y. Gumulya, Doctoral Thesis, Ruhr-Universität Bochum (Germany), **2010**.
[23] a) A. Grinberg, R. Bernhardt, *Biochem. Biophys. Res. Commun.* **1998**, *249*, 933 – 937; b) O. Almog, D. T. Gallagher, J. E. Ladner, S. Strausberg, P. Alexander, P. Bryan, G. L. Gilliland, *J. Biol. Chem.* **2002**, *277*, 27553 – 27558.
[24] Recent examples of ISM (refs. [9,13]): a) S. Kille, F. E. Zilly, J. P. Acevedo, M. T. Reetz, *Nat. Chem.* **2011**, *3*, 738 – 743; b) M. A. Norrgard, B. Mannervik, *J. Mol. Biol.* **2011**, *412*, 111 – 120; c) W. L. Tang, Z. Li, H. Zhao, *Chem. Commun.* **2010**, *46*, 5461 – 5463; d) K. Engström, J. Nyhlen, A. G. Sandström, J.-E. Bäckvall, *J. Am. Chem. Soc.* **2010**, *132*, 7038 – 7042; e) M.-W. Bhuiya, C.-J. Liu, *J. Biol. Chem.* **2010**, *285*, 277 – 285; f) M. T. Reetz, S. Prasad, J. D. Carballeira, Y. Gumulya, M. Bocola, *J. Am. Chem. Soc.* **2010**, *132*, 9144 – 9152; g) K. Okrasa, C. Levy, M. Wilding, M. Goodall, N. Baudendistel, B. Hauer, D. Leys, J. Micklefield, *Angew. Chem.* **2009**, *121*, 7827 – 7830; *Angew. Chem. Int. Ed.* **2009**, *48*, 7691 – 7694; h) M. R. M. De Groeve, M. De Baere, L. Hoflack, T. Desmet, E. J. Vandamme, W. Soetaert, *Protein Eng. Des. Sel.* **2009**, *22*, 393 – 399.
[25] a) W. M. Patrick, A. E. Firth, *Biomol. Eng.* **2005**, *22*, 105 – 112; b) A. D. Bosley, M. Ostermeier, *Biomol. Eng.* **2005**, *22*, 57 – 61; c) M. Denault, J. N. Pelletier in *Protein Engineering Protocols, Vol. 352* (Eds.: K. M. Arndt, K. M. Müller), Humana, Totowa, **2007**, pp. 127 – 154; see also earlier approaches regarding this type of statistical analysis, for example: d) M. S. Warren, S. J. Benkovic, *Protein Eng.* **1997**, *10*, 63 – 68.

Received: June 27, 2011
Published online on September 13, 2011

