ChemPubSoc Europe

DOI: 10.1002/cbic.201300486

Directed Evolution by Using Iterative Saturation Mutagenesis Based on Multiresidue Sites

Loreto P. Parra,[a, b, c] Rubén Agudo,[a, b] and Manfred T. Reetz*[a, b]

Iterative saturation mutagenesis (ISM) in combination with reduced amino acid alphabets has been shown to be an efficient method for directed evolution. In order to minimize the screening effort, the number of residues in a given randomization site has thus far been restricted to two or three; this prevents oversampling from reaching astronomical numbers when 95% library coverage is aimed for. In this study, ISM is applied for the first time by using randomization sites composed of five amino acid positions. The use of just two such sites (A and B) results in two different ISM pathways, A→B and B→A. A severely reduced amino acid alphabet (only five members) was employed for the building blocks—a minimal set of structurally representative amino acids. The Baeyer–Villiger monooxygenase PAMO was chosen as the enzyme for this proof-of-principle study. The test system employed tuning of activity and diastereoselectivity in the oxidation of 4-(bromomethylidene)cyclohexanone, which is not accepted by wild-type PAMO. Although only 8–9% library coverage was ensured (as calculated by traditional statistics), notable activity and 99% diastereoselectivity were obtained, thus indicating that such an ISM strategy is viable in protein engineering.

Introduction

Directed evolution has emerged as a powerful way to manipulate essentially any catalytic property of an enzyme, including stereo- and regioselectivity, substrate scope, rate, and thermal robustness.[1] It involves repeated cycles of gene mutagenesis, expression, and screening (selection). The most common mutagenesis methods are error-prone polymerase chain reaction (epPCR), saturation mutagenesis, and DNA shuffling. The degree of biocatalyst improvement depends on the amount of molecular biological work and screening effort the experimenter is willing to invest. As the screening step is the bottleneck in directed evolution, a great deal of research has been invested in the quest to generate higher-quality mutant libraries, which require less assay effort for a given property.[1,2] Our contribution to this task is iterative saturation mutagenesis (ISM),[3] according to which properly chosen single- or multiresidue sites (designated A, B, C, D, etc.) are first randomized by saturation mutagenesis. The gene of an improved mutant ("hit") in one library is then used as a template for saturation mutagenesis at other sites, and the process is continued until the desired degree of catalyst improvement has been achieved. The criteria for choosing appropriate sites depend upon the catalytic property of interest. In the case of stereo- or regioselectivity, rate, and/or substrate scope, the combinatorial active-site saturation test (CAST) is employed,[4] according to which residues surrounding or near the binding pocket are selected for saturation mutagenesis. When aiming for enhanced stability, the B-FIT method is called for: residues showing the highest average B-factors are chosen for saturation mutagenesis.[3,5] Both embodiments of ISM rely on structural biology and are thus knowledge-driven.

Saturation mutagenesis next to the binding pocket of an enzyme for the purpose of enhancing stereoselectivity was first reported in 2001: *Pseudomonas aeruginosa* lipase (PAL) was the catalyst in the hydrolytic kinetic resolution of a chiral ester,[6] although the acronym CAST was not used at the time. Interestingly, in that study a four-residue site was subjected to saturation mutagenesis by using NNK codon degeneracy encoding all 20 canonical amino acids (N: adenine/cytosine/guanine/thymine; K: guanine/thymine), yet only 5000 transformants were screened. Following the introduction of ISM into a CAST study aimed at enhancing the enantioselectivity of the epoxide hydrolase from *Aspergillus niger*,[7] several developments ensued aimed at improving the efficiency of saturation mutagenesis in general: 1) use of reduced amino acid alphabets,[8] 2) application of an improved PCR method for creating mutant libraries in cases of difficult-to-amplify templates,[9] 3) use of designed mixtures of primers for reducing codon redundancy (amino acid bias),[10] and 4) use of pooling techniques.[11] Coupled with statistical models for assessing the extent of oversampling necessary to ensure a defined degree of library coverage[8,12] (which had not been considered in the

[a] Dr. L. P. Parra, Dr. R. Agudo, 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] Dr. L. P. Parra, Dr. R. Agudo, Prof. Dr. M. T. Reetz
Fachbereich Chemie, Philipps-Universität Marburg
Hans-Meerwein-Strasse, 35032 Marburg (Germany)

[c] Dr. L. P. Parra
Department of Chemical and Bioprocesses Engineering
Pontificia Universidad Católica de Chile
Avenida Vicuña Mackenna 4860, Santiago (Chile)

Supporting information for this article is available on the WWW under http://dx.doi.org/10.1002/cbic.201300486.

original study),[6] all of these advances provide higher-quality libraries that require less screening. They therefore contribute to overcoming the numbers problem in directed evolution.[3,8]

When applying CAST, either single-residue or multiresidue sites are considered.[3] Although utilization of single-residue sites might well lead to improved enzyme mutants,[11b] we have shown that the use of two- and three-residue sites is the better option, because this increases sequence diversity while concomitantly reducing the number of steps (cycles) and therefore the total amount of screening in a multistep ISM process.[3,13] Such a strategy also allows epistatic effects in the tuning of an enzymatic property, especially cooperativity (more than mathematical additivity), not only between sets of mutations but also between the point mutations within a given set.[3,14] In most cases, docking the substrate into the X-ray structure of the enzyme (or homology model) will lead to the identification of many different residues as possible candidates for saturation mutagenesis. Utilization of all of these might well demand considerable experimental effort, and experience has shown that this is not necessary.[3] This means that some kind of a decision needs to be made regarding exclusion of certain residues from further consideration.

Guidelines on how to group single residues identified for a CAST process on the basis of structural data have not been proposed to date. The optimal choice and cost of the megaprimers can be expected to influence such decisions. As the degree of oversampling for 95% library coverage (as estimated on the basis of traditional statistics described by Patrick and Firth[12a] and incorporated in the CASTER computer manual)[15] rises drastically as the number of residues in a given site increases, we introduced reduced amino acid alphabets as building blocks when choosing two- or three-residue randomization sites.[3,8,13] This proved to be efficient: for example, NDT codon degeneracy (D: adenine/guanine/thymine; T: thymine) encodes 12 amino acids (Phe, Leu, Ile, Val, Tyr, His, Asn, Asp, Cys, Arg, Ser, Gly), a balanced mixture of polar/nonpolar, charged/uncharged, and aromatic/aliphatic building blocks.[3,8,13] When two- and three-residue sites are chosen, application of this particular codon degeneracy results in small but high-quality libraries, typically requiring a screening effort of only 2000–5000 transformants for high library coverage. Moreover, significantly smaller amino acid alphabets can be employed, as in the bioinformatics approach used in the ISM-based directed evolution of stereoselective Baeyer–Villiger monooxygenases; this likewise reduced the screening effort drastically.[16]

The number of upward pathways in an ISM system depends upon the number of sites (e.g., four-site system: $4! = 24$ pathways; three-site system: $3! = 6$ pathways; two-site system: 2 pathways). There then arises the question as to which pathway should be chosen. In more than two dozen ISM studies from various groups (including ours), more-or-less arbitrary choices were made, thus suggesting that this question is not crucial.[3] However, uncertainty regarding this issue remains (one of the reasons for performing the present study).

Relevant to this is the above-mentioned early report of saturation mutagenesis for enhancing the enantioselectivity of the lipase PAL; for this, a four-residue site was randomized by employing NNK codon degeneracy.[6] Despite the fact that only 5000 transformants were screened, significant improvement in enantioselectivity was observed, although oversampling was not considered at that time. According to traditional statistical analysis,[12a–d,15] about 3×10^6 transformants should have been screened for 95% library coverage. These and subsequent empirical results show that in order to generate notably improved enzyme mutants, it is not necessary to ensure 95% library coverage (as calculated on the basis of traditional statistics), although there is the probability of missing many other hits.[1,3,12] Nevertheless, at that time it was not clear whether the positive result was serendipitous.[6] In subsequent studies we therefore avoided sites of more than three amino acid positions. Noteworthy in this context is the report by Bäckvall and co-workers. They described the successful engineering of enantioselectivity and substrate scope of *Candida antarctica* lipase A (CALA) by randomization at a nine-residue CAST site, but with a highly reduced amino acid alphabet.[17]

In this study we decided to test a new strategy for applying ISM: focused CAST-type saturation mutagenesis at two large multiresidue sites (A and B) around the binding pocket of an enzyme, each of five amino acid positions, followed by exploration of the two possible ISM pathways: A→B and B→A (Figure 1). Thus, only four mutant libraries are involved. This strategy eliminates the uncertainty in choosing one of many upward pathways in a multisite ISM system. From a methodologically viewpoint it was also of interest to see whether the two pathways lead to equal success in terms of catalytic performance of the final best hits. In view of the statistics regarding the degree of screening, we expected to employ appropriately chosen reduced amino acid alphabets. In spite of this, high library coverage was not aimed for at the outset of the project. Rather, the strategy was to keep the screening effort to a minimum, and to see how things develop. In the event of unsatisfactory results, we planned to increase successively the amount of screening by covering more of the restricted protein sequence space, as defined by the number of randomization residues and the reduced amino acid alphabets employed in saturation mutagenesis.

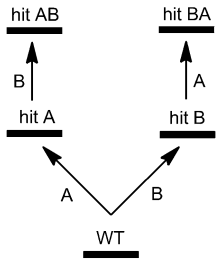
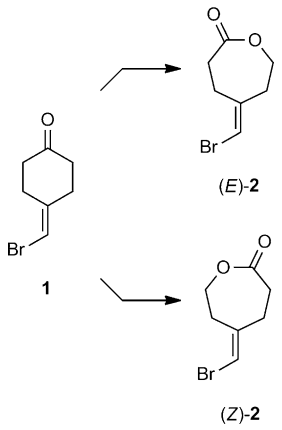


Figure 1. ISM with two multiresidue sites (A and B).

The thermostable Baeyer–Villiger monooxygenase (BVMO)[18] from *Thermobifida fusca*, phenylacetone monooxygenase[19] (PAMO; EC 1.14.13.92), was chosen as the catalyst; the model reaction was diastereoselective Baeyer–Villiger (BV) oxidative transformation of ketone **1** to lactones (*E*)-**2** or (*Z*)-**2** (Scheme 1). PAMO does not accept cyclohexanone or derivatives thereof, including substrates like **1**.[16,19,20] Thus, increasing substrate scope with reasonable enzyme activity was the primary parameter of interest, but *E*/*Z* diastereoselectivity was also considered. This unusual kind of stereoselectivity was recently achieved for the first time by using whole cells of



Scheme 1. Oxidative transformation of ketone **1** to lactones (*E*)-**2** and (*Z*)-**2**.

expressing the less-robust cyclohexanone monooxygenase (CHMO) as the biocatalyst.[21]

Results and Discussion

Choice of the two five-residue sites, rational library design, and screening platform

One of the PAMO 3D structures, PDB:2YLT,[22] crystallized in complex with 2-(*N*-morpholino)ethanesulfonic acid (MES) and $NADP^+$, was used as a guide for choosing the two five-residue sites. Removal of MES and docking of substrate **1** in its place produced the model shown in Figure 2, with 17 single residues within 5 Å of the binding pocket. In principle, any of these can be chosen. However, it was logical to utilize as much structural and mechanistic data as possible, as this might lead to elimination of some of these residues as candidates for the ten positions to be used in the CAST sites A and B. For example, D66 was not considered, because it participates in the binding of $NADP^+$ at its nicotinamide region; R337 was similarly excluded, because it is involved in stabilization of the catalytic intermediate.[19,22] Residues L340, I339 (both second sphere, >5 Å), W501 (close to $NADP^+$), N445, and M446 could have been chosen, but were arbitrarily eliminated. The remaining ten single residues were grouped into two multiresidue sites: A (L289/R258/I67/S444/A442) and B (S441L443/Y502/L338/P286). Decisions of this kind can be arbitrary, but we were guided by the sizes of the respective megaprimers, as these should be as similar as possible at a given site. It should be mentioned that positions 441–444 are part of a loop that had been targeted in a previous study; that investigation used a different substrate as the model compound, and did not consider other sites.[16]

For 95% library coverage, when five residues are randomized simultaneously with an amino acid alphabet of all 20 canonical members and NNK codon degeneracy, the screening effort amounts to more than 5×10^7 transformants, as calculated by traditional statistical methods.[3,8,12a–d] If the amino acid alphabet is reduced to 12 members (e.g., NDT codon degeneracy), the number collapses to 745433—still a formidable screening effort. In order to reduce both codon redundancy and the screening effort, a limited amino acid set (five members) was considered. It was chosen to yield a balanced mixture of amino acid properties: negatively charged (aspartate), positively charged (arginine), sterically bulky (phenylalanine), sterically small and hydrophobic (alanine), and special side chain (proline). The wild-type (WT) amino acid at each position was also considered. With six amino acids in each set, the calculated screening effort for 95% library coverage (using approximations based on traditional statistics) reduced considerably (Figure 3): 16176 clones (transformants) for library A and 19411 for library B.

The most convenient and cheapest medium-throughput screening system (used in this and many other studies) is automated gas chromatography (GC) or HPLC. Although the above

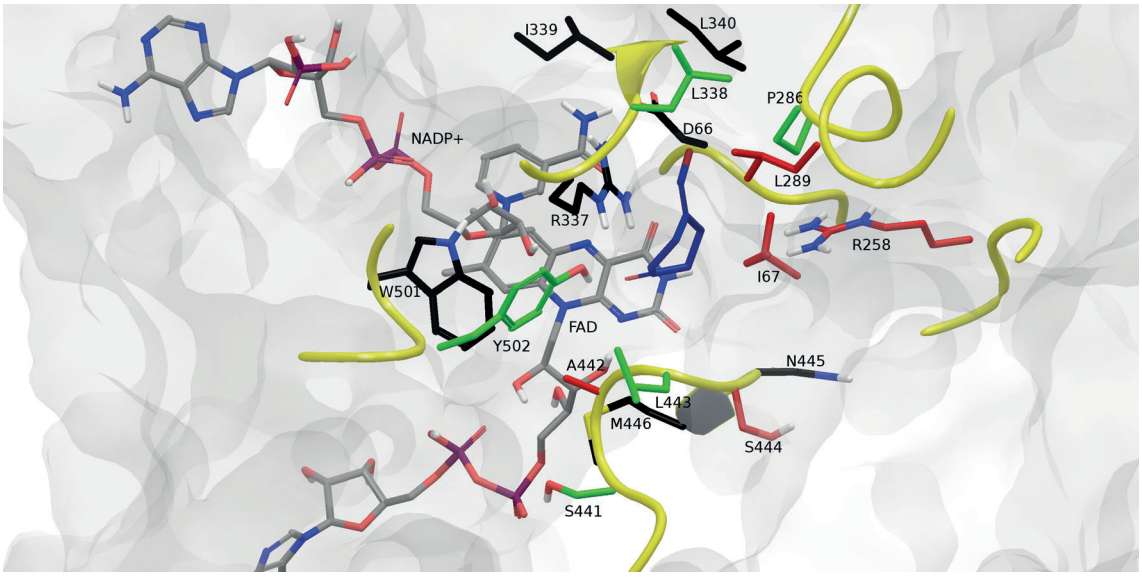


Figure 2. Choice of the two five-residue sites A and B for saturation mutagenesis showing substrate **1** (dark blue) docked in the published crystal structure of PAMO (PDB ID: 2YLT; co-crystalized MES removed):[22] Site A: L289/R258/I67/S444/A442 (red) and site B: S441L443/Y502/L338/P286 (green).

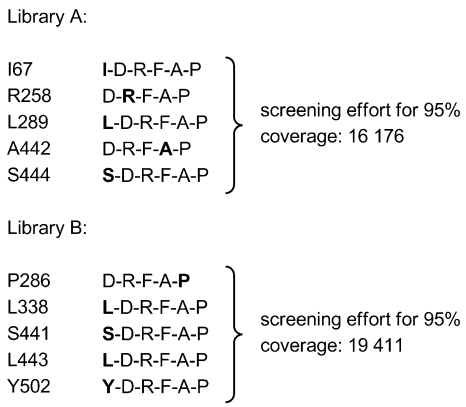


Figure 3. Screening effort for 95% coverage as calculated by traditional statistics (boldface letters are wild-type residues in PAMO).

numbers can be handled by GC, the screening process still requires considerable time if 95% library coverage is to be ensured. We therefore decided to adopt the following strategy: assay each library initially with just 16 microtiter plates of the 96-well format (1536 transformants per library, including four controls per plate); this corresponds to only 9% (library A) and 8% (library B) library coverage. In cases where improved mutants showing notably higher activity are not found, we expand the screening by increasing the degree of library coverage.

Generation of saturation mutagenesis libraries

In order to construct the library design as defined above, a two-step process for saturation mutagenesis was devised. First, two independent PCR steps were performed to amplify two regions at site A carrying the mutations: one containing two amino acids to be randomized (Megaprimer I) and the other containing three amino acids (Megaprimer II). Following purification of the megaprimer amplicons, these were used as primers in a second round of PCR (Figure 4). The PCR products were digested with DpnI and transformed in *Escherichia coli* TOP10. The same procedure was performed for the library at site B (Megaprimers I′ and II′). Prior to picking the colonies, Quick Quality Control (QQC)[11b] was applied (see the Supporting Information) to ensure that most of the diversity at the DNA level was in fact achieved, thereby avoiding screening of something that does not exist.

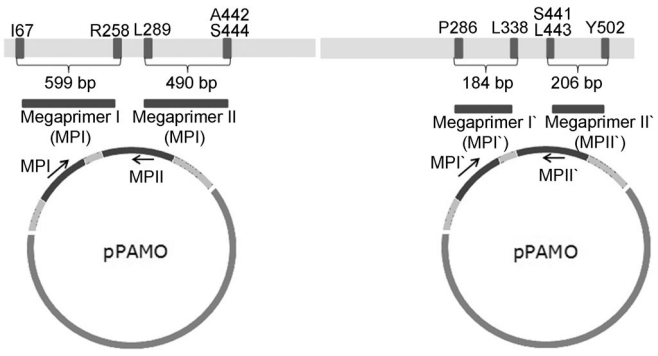


Figure 4. Strategy for the generation of mutant libraries based on the megaprimer method and traditional PCR for saturation mutagenesis at multiresidue sites A (left) and B (right).

Results

Upon successful library creation (Supporting Information), the library was transformed (Experimental Section), followed by colony picking and transferring into 96-well microtiter plates. When using WT PAMO in experimental reaction conditions (lysate) in microtiter plates (20 h incubation), no conversion of ketone **1** to the respective lactone **2** was observed. Rather, GC/MS analysis showed about 50% conversion to a product corresponding to reduction of the carbonyl moiety of **1**. We assume that this reaction is catalyzed by an unidentified alcohol dehydrogenase (ADH) in the lysate. Gratifyingly, upon screening 16 96-deep-well microtiter plates containing transformants generated by saturation mutagenesis at multiresidue site A with the designed reduced amino acid alphabets and by applying pooling (Experimental Section), five hits showing PAMO activity were identified (Table 1). In addition, about 46–47% conversion to the alcohol was observed.

Table 1. Results of screening a small portion (<9%) of the initial library generated by saturation mutagenesis at multiresidue site A (positions 67, 258, 289, 442, and 444) with reduced amino acid alphabets. Reactions (20 h, in microtiter plates) used crude lysate.

Mutant	Conversion[a] [%]	Mutant	Conversion[a] [%]
WT PAMO	0.1	A442P/S444A	10
R258D/L289A/A442P/S444A	4	L289A/A442P/S444A	6
I67A/R258A/A442P/S444A	12	L289A/A442P	5

[a] Conversion of ketone **1** to lactone **2**. In addition, 46–47% conversion to the alcohol corresponding to reduction of **1** was observed.

All the active variants showed >99% diastereoselectivity in favor of (*E*)-**2**. It is interesting to note that positions 442 and 444 seem to be especially "hot spots", with proline and alanine, respectively, being highly conserved. Based on the ability to convert **1** to the desired product **2**, the best hit proved to be the quadruple-mutant I67A/R258A/A442P/S444A. This was substantiated by using semipurified enzyme (Experimental Section).

These results are significant because WT PAMO does not accept **1** as a substrate, even after prolonged reaction times. The gene for the best variant obtained in library A (I67A/R258A/A442P/S444A) was then used as a template in ISM step A→B by performing saturation mutagenesis at multiresidue site B, again by using the designed reduced amino acid alphabet. Upon screening (again, 16 microtiter plates, 96-well format), about ten active clones were found. These were semipurified by first storing the lysates at 4 °C for 24 h and then in-

cubating them at 50 °C for 60 min. After centrifugation, the amount and degree of purity of enzyme were estimated by SDS-PAGE and by measuring the absorbance at 441 nm (Experimental Section). Subsequently, glucose dehydrogenase (GDH) and glucose were added, thus completing the experimental platform for performing the model BV reaction. The five best mutants are reported in Table 2 (see the Experimental Section).

Table 2. Results of screening a small portion (<9%) of the library obtained by saturation mutagenesis at multiresidue site B (positions 286, 338, 441, 443, and 502) with the previous mutant I67A/R258A/A442P/S444A as template and designed amino acid alphabets. New mutations introduced in this ISM step are marked in bold. Conditions: semipurified enzyme (5 μM), substrate (2 mM); 20 h reaction time; in all cases >99% diastereoselectivity in favor of (*E*)-**2**.

Mutant	Conversion [%] to lactone **2**	to alcohol
I67A/R258A/A442P/S444A/**S441P**/**L443D**/**Y502F**	62	35
I67A/R258A/A442P/S444A/**P286A**/**S441P**/**L443A**/**Y502F**	79	17
I67A/R258A/A442P/S444A/**P286R**/**L443F**/**Y502F**	81	18
I67A/R258A/A442P/S444A/**P286A**/**S441P**/**L443A**	81	18
I67A/R258A/A442P/S444A/**P286R**/**S441P**/**L443D**/**Y502F**	79	20

For example, a variant with seven point mutations, I67A/R258A/A442P/S444A/**P286R**/**L443F**/**Y502F** (the three mutations introduced in the second mutagenesis round are shown in bold face), was an acceptable catalyst: 81% conversion to the desired lactone **2** with >99% *E* selectivity; only 18% competing reduction to the alcohol occurs when this more active PAMO mutant is used. Interestingly, no mutations occurred at L338 in any of the variants, thus suggesting that leucine at this position is essential for activity. In the absence of a detailed quantum mechanics/molecular mechanics (QM/MM) study, the high degree of observed diastereoselectivity is difficult to explain at the molecular level. However, point mutations such as I167A and Y502F indicate that more space is provided in the binding pocket, thus contributing to acceptance of the model substrate **1**.

The results are encouraging, because high conversion and essentially complete diastereoselectivity were obtained. Repeating the ISM steps with higher library coverage would probably provide even better results. Rather than testing such a strategy, we decided to explore the alternative ISM pathway B→A, with the same degree of oversampling in the screening step. Analogous procedures regarding the reaction conditions and the number of transformants were applied, and the results in the case of the initial library generated by saturation mutagenesis at multiresidue site B are summarized in Table 3. The result of the QQC test[11b] proved acceptable.

In all cases, *E* selectivity proved to be complete (>99%). The best variant was the double-mutant S441P/L443D (51% conversion of **1** to the desired lactone (*E*)-**2**). When comparing the point mutations at site B in these variants with those generated at site B by the first ISM-Scheme (A→B), some noteworthy trends are evident. Leucine at position 338 was not exchanged, as previously observed. In both libraries only phenylalanine (no other amino acid) was introduced at position 502. However, in no case were the mutational sets identical. The best mutant S441P/L443D was used as the template for subsequent randomization at multiresidue site A (Table 4).

Table 3. Results of screening a small portion (<9%) of the initial library generated by saturation mutagenesis at multiresidue site B (positions 286, 338, 441, 443, and 502) with reduced amino alphabets. Conditions as in Table 2; in all cases complete *E* selectivity was observed.

Mutant	Conversion [%][a]	Mutant	Conversion [%][a]
S441D/L443F	6	P286F/S441A/L443A/Y502F	6
S441P/L443D	51	S441A/L443D	14
S441D/L443F/Y502F	9	P286R/L443D	7

[a] To lactone **2**, with 20–45% conversion to the reduction product.

Table 4. Results of screening a small portion (<9%) of the library obtained by saturation mutagenesis at multiresidue site A (positions 67, 258, 289, 442, and 444) by using the previous mutant S441P/L443D as template and designed amino acid alphabets. New mutations introduced in this ISM step are shown in boldface. Conditions as in Table 2; in all cases complete *E* selectivity was observed.

Mutant	Conversion [%] to lactone **2**	to alcohol
S441P/L443D/**R258P**/**A442P**/**Y444F**	40	21
S441P/L443D/**A442P**/**Y444F**	27	59
S441P/L443D/**L289F**/**A442P**/**Y444F**	54	40
S441P/L443D/**R258D**/**A442P**/**Y444F**	23	61

Pathway A→B was more successful than B→A, at least when exploring a highly limited portion of the total defined protein sequence space (<9% library coverage, on the basis of conventional statistics).

Scale-up using resting cells

Nongrowing cells were employed in experiments aimed at scaling up the model reaction (Experimental Section). After a reaction time of only 8 h, complete conversion of substrate **1** was achieved (Figure 5). As expected, the best mutant proved to be I67A/R258A/A442P/S444A/**P286R**/**L443F**/**Y502F**, which showed good performance under the scale-up conditions (80% conversion to the desired lactone (*E*)-**2**, >99% stereoselectivity, only 20% alcohol as a side product). This contrasts significantly with the performance of the other mutants, which were less active. Formation of undesired reduction product can have various reasons, and this led us to perform additional experiments. Rather than relying on partial purification

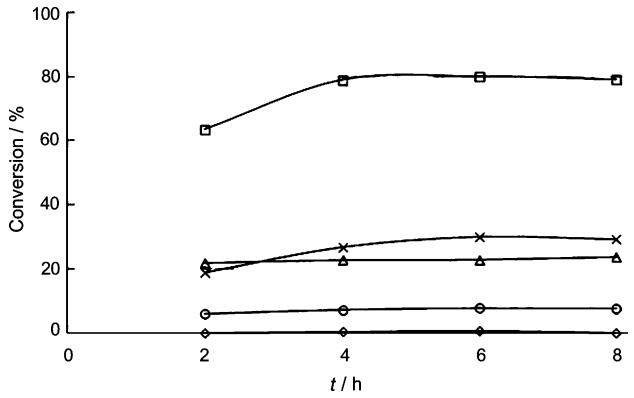


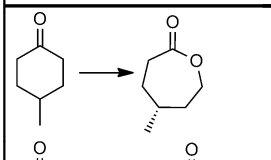
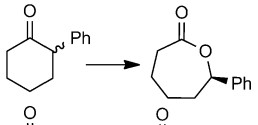
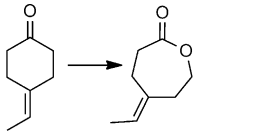
Figure 5. Scaling up of the **1**→**2** reaction by using non-growing cells harboring PAMO mutants lib A (×), lib AB (□), lib B (○), lib BA (△), and WT (◇), .

(as above), the best mutant was fully purified (Figure S1) and tested in the model reaction with GDH and glucose as the NADPH regeneration system (see the Experimental Section). The purified mutant I67A/R258A/A442P/S444A/**P286R**/**L443F**/**Y502F** achieved more than 99% conversion to lactone (*E*)-**2**, and minimal alcohol (<1%) after a reaction time of 16 h.

Testing the best PAMO mutants as catalysts in reactions with other substrates not accepted by the WT enzyme

The substrate scopes of the two best mutants obtained by pathway A→B were studied by testing several substrates that are also not accepted by WT PAMO: one variant in the initial library at site A (I67A/R258A/A442P/S444A), and the other in the library from the subsequent ISM step (variant I67A/R258A/A442P/S444A/**P286R**/**L443F**/**Y502F**). The three ketones were accepted by the mutants, unlike WT PAMO (Table 5). However, different PAMO mutants[20] and WT CHMO[18] are generally better suited, as these lead to higher conversion and excellent enantioselectivity. Nevertheless, the best mutant is potentially a promising starting point for fine-tuning the enzymes by performing additional mutagenesis experiments. Also noteworthy is the observation regarding stereoselectivity in the oxidative kinetic resolution of 2-phenylcyclohexanone (Table 5). From the initial mutant in the first library at site A to the second-generation mutant (ISM pathway A→B), the selectivity factor *E* (relative rate of reaction of one enantiomer with respect to the other) increased from 33 to ~400. This remarkable result invites a thorough theoretical analysis in future studies.

Table 5. Results of testing two of the best-evolved PAMO mutants as semipurified catalysts in stereoselective Baeyer–Villiger reactions with three other substrates (also not accepted by WT PAMO).[a]

Substrate	PAMO-mutant Lib A.[b]	PAMO-mutant Lib A→B.[c]
	20% conv. to lactone >99% *ee* (*R*)	32% conv. to lactone >99% *ee* (*R*)
	50% conv. to lactone *E*=33 (*R*)	48% conv. to lactone *E*=415 (*R*)
	11% conv. to lactone >99% *E* selective	29% conv. to lactone >99% *E* selective

[a] Ph: Phenyl; conditions: 20 h at 30 °C. [b] Mutant I67A/R258A/A442P/S444A. [c] Mutant I67A/R258A/A442P/S444A/P286R/L443F/Y502F.

Conclusions

The goal of this study was methodology development in directed evolution. The thermostable Baeyer–Villiger monooxygenase PAMO served as the catalyst in the Baeyer–Villiger reaction of ketone **1** with stereoselective formation of lactone (*E*)-**2**, a diastereoselective transformation that we have previously investigated with a different BVMO (CHMO).[21] Although ISM has emerged as a particularly effective method in directed evolution,[3] the question as to the ideal number of amino acid positions to be included in a randomization site remains. Thus far ISM has been restricted to sites comprising a maximum of three residues, and the screening effort has been held to a minimum by the use of reduced amino acid alphabets. The present study was designed to address this question.

Consider a system in which ten CAST residues have been identified. Grouping them into five two-residue sites would lead to an ISM system of 5!=120 different five-step upward pathways, and thus uncertainty concerning the choice of the best upward trajectory. In sharp contrast, grouping them into two five-residue sites involves only two libraries in each of the two possible pathways. However, such large sites require highly reduced amino acid alphabets in order to keep the screening effort to a reasonable level; this, of course, limits structural diversity. The objective of this proof-of-concept study was to test such a system. Both pathways A→B and B→A were explored by using a Baeyer–Villiger monooxygenase (PAMO) and a substrate that is not accepted by the enzyme. This means that only four mutant libraries had to be generated and screened. The results show that such a strategy is indeed viable. As oversampling in the screening steps with conventional NNK codon degeneracy (encoding all 20 canonical acids) would be an astronomical task for achieving only moderate library coverage (as calculated by traditional statistics),[3,8,12a–d] a highly reduced amino acid alphabet of only five members was employed, and this proved successful.

Having little mechanistic or mutational data, we chose for the building blocks a structurally balanced set of five amino acids: asparagine (negatively charged), arginine (positively charged), phenylalanine (bulky), alanine (small and hydrophobic), and proline (special rigid structure), in addition to the WT residues at each position. This was a rational choice, because one can consider this a minimal amino acid alphabet with full representative structural diversity. We recommend this minimal alphabet for future directed evolution studies. Nevertheless,

we do not maintain that it is the only possibility, even under the given circumstances. For example, instead of choosing phenylalanine as a fairly bulky and hydrophobic amino acid, isoleucine could have been used (or certain other amino acids). Moreover, if some of the chosen single residues are part of a loop (as in our case), proline is a logical choice. In future studies the experimenter following the kind of approach introduced here should embrace as much information from structural biology as possible, including any available mutational data from previous protein engineering studies.

Of the two ISM pathways here, A→B proved to be more prolific under the conditions used. This does not mean that A→B is fundamentally better; if larger portions of the libraries had been screened, a different conclusion might have been reached. The quality of the libraries is crucial in such assessments, so Quick Quality Control (QQC) should always be applied.[11b] There is no reason to expect identical mutants from the two pathways. The screening was designed to focus primarily on activity; in all cases *E* selectivity resulted in the BV reaction with the model compound, **1**. In a previous study with WT CHMO and mutants under the same reaction conditions, *E* selectivity was the preferred diastereoselectivity, and reversing it (formation of (*Z*)-**2**) by directed evolution required a fairly intensive search.[21] Detailed QM/MM studies of the enantioselectivity of CHMO and mutants as catalysts in the oxidative desymmetrization of 4-methyl- and 4-hydroxycyclohexanone have been performed;[23] however, PAMO has yet to be studied theoretically. This should be followed by QM/MM and kinetics analysis of the best mutants described herein. In the present study we discovered that simple docking experiments are not sufficient to understand the unusual diastereoselectivity of the reaction **1**→(*E*)-**2**.

The assaying step was purposely restricted to 16 microtiter plates (96-well format 1536 samples per library, including four controls per plate): 1472 transformants. This translates into just 8–9% library coverage, as estimated by conventional statistics. The actual screening effort with automated GC was actually less because of the pooling procedure: 384 samples per library, 1536 GC runs for all four libraries in both ISM pathways. Thus, a very limited portion of the protein sequence space (determined by codon degeneracy) was in fact considered, yet notably improved PAMO variants were obtained. As noted in several previous studies that used saturation mutagenesis, it is not at all necessary to ensure 95% library coverage (as calculated by traditional statistics),[3,12] but it is helpful to know the statistical numbers as an orientational guide. Indeed, screening only a fraction of the relevant protein sequence space usually provides improved (albeit not necessarily optimal) mutants, as first demonstrated more than 12 years ago.[6] A logical explanation for such empirical observations emerges upon applying the recently published statistical model of Nov,[12e] in which the second best (or *n*th best) mutants are considered. Settling for the second (or *n*th) best variant reduces the computed amount of screening drastically, in contrast to the more extensive screening required for 95% library coverage as calculated by traditional statistics.[3,12]

Concerning practical questions in the present study, we note that the use of an in vitro system based on isolated PAMO mutants and an NADPH regeneration system leads to the elimination of undesired side products (reduction of the carbonyl function).

We expect that the underlying concept presented here will be successful in other cases. We plan to investigate the effect of increasing the number of members in a multiresidue site significantly beyond five in two-site ISM systems, while reducing the amino acid alphabet to a bare minimum. Comparison of simultaneous randomization at all identified residues with formation of a single mutant library as an alternative strategy will reveal the relative efficacy of these two approaches.

Experimental Section

General procedures: DpnI was purchased from New England Biolabs. KOD Hot Start Polymerase, 10× KOD buffer, dNTPs, and $MgSO_4$ were obtained from Novagen. HPLC-purified primers (Tables S1–S4) were synthetized by Eurofins (Ebersberg, Germany). GDH, NADPH, and $NADP^+$ were purchased from Codexis (Redwood City, CA); L-arabinose was from Sigma–Aldrich; D-glucose monohydrate, DNase I, lysozyme, lysogeny broth (LB, premixed powder), and carbenicillin (CB) were from Applichem (Darmstadt, Germany). TB medium contained yeast extract (24 g L^{-1}), peptone (12 g L^{-1}), glycerol (4 mL L^{-1}), KH_2PO_4 (0.017 M), and K_2HPO_4 (0.072 M). SOC medium contained yeast extract (20 g L^{-1}), peptone (5 g L^{-1}), glucose (20 mM), NaCl (10 mM), KCl (2.5 mM), $MgCl_2$ (10 mM), and $MgSO_4$ (10 mM). Plasmids were purified on Qiagen spin-columns. Electro-competent cells were prepared according to standard protocols (Bio-Rad manual; MicroPulser Electroporation Apparatus Operating Instructions and Applications Guide). Starting substrates and products have been described previously.[20,21]

Library creation: Plasmid pPAMO[19] was used as the template to create saturation mutagenesis libraries A (L289/R258/I67/S444/A442) and B (S441/L443/Y502/L338/P286). Saturation mutagenesis libraries were created in two steps. First, two parallel PCR reactions were performed to amplify two regions carrying the mutations; for site A, one contained two amino acids to randomize (Megaprimer I), and the other containing three amino acids (Megaprimer II); the analogous procedure was followed for site B with Megaprimers I′ and II′. After PCR purification, a second PCR was performed with the two purified megaprimers (Figure 4). The amplification reaction for megaprimer amplification contained dNTPs (2 mM each), $MgCl_2$ (25 mM), mutagenic primers (3.5 μM each), template plasmid (50 ng), and KOD Hot Start polymerase (0.5 U) in KOD buffer (50 μL). Optimized PCR conditions: 95 °C for 3 min, then 27 cycles of 95 °C for 60 s, 62 °C for 60 s, and 68 °C for 45 s, and a final extension step of 68 °C for 10 min. Amplified products (megaprimers) were purified by using a PCR purification kit (Qiagen). For the second PCR reaction, two megaprimers (5 pmol each) were used with dNTPs (2 mM each), $MgCl_2$ (25 mM), template plasmid (50 ng), and KOD Hot Start polymerase (0.5 U) in a KOD buffer (20 μL). Optimized PCR conditions: 1 cycle at 95 °C for 3 min; 5 cycles of 95 °C for 30 s, 68 °C for 60 s, and 72 °C for 60 s; 24 cycles of 95 °C for 30 s and 68 °C for 6 min; and final extension at 68 °C for 10 min. In order to digest the template plasmid, the PCR products (20 μL) were incubated with DpnI (1 μL) at 37 °C for 1.5 h, then further DpnI (1 μL) was added, and incubation was continued for 90 min. An aliquot (2 μL) was used directly to transform *E. coli* TOP10 electrocompetent cells (50 μL). The transformation mixture was incu-

bated with SOC medium (1 mL) at 37 °C with shaking. After 1 h, an aliquot (30 μL) was spread onto LB-agar plates supplemented with CB (100 μg mL^{-1}). More than 5000 colonies were obtained per transformation.

Integrity and library quality control: The integrity and quality of the coding region was verified by DNA sequencing as previously described (Table S5).[11b] Briefly, colonies obtained after transformation were scratched from the plate with a Drigalsky spatula and placed in LB medium (4 mL). The pool of DNA was extracted from the collected cells by using the QIAprep Miniprep Kit (Qiagen) and analyzed by sequencing. In addition, ten colonies from each library were randomly picked. Plasmids were sequenced to confirm that all five sites were mutated simultaneously and that all five (or six) amino acids were present. For this reason, the majority of the primers inserting WT amino acids were designed with a different codon to recognize the mutations (Table S6).

Library screening: For each library, 1472 individual colonies were randomly picked and used to inoculate 16 2.2 mL 96-deep-well plates containing LB medium (800 μL per well) supplemented CB (100 μg mL^{-1}). Four wells in the last column of each plate were inoculated with the parental strain (1536 colonies per library). The plates were incubated overnight at 37 °C with shaking (800 rpm). After cell growth, aliquots (70 μL) of overnight culture were transferred to a 96-well plate filled with glycerol solution (120 μL; 70%) and kept at −80 °C. Other aliquots (100 μL) were used to inoculate fresh TB medium (700 μL) supplemented with CB (100 μg mL^{-1}), with L-arabinose (0.1%) as the inducer. After 20 h of expression at 37 °C, four expression cultures were pooled in one 96-deep-well plate by combining 200 μL of expression cultures belonging to the same well (four pool plates (800 μL of expression culture) for each library). The cells of the 16 plates plus the four pool plates were harvested by centrifugation (3220 *g*, 5 min), and pellets from the original 16 plates were stored at −80 °C. Each cell pellet from the four pooling plates was resuspended in lysis buffer (600 μL; Tris·HCl buffer (50 mM, pH 8.0)) containing lysozyme (1.4 mg mL^{-1}) and DNase I (6 U)). Lysis was performed at 50 °C (800 rpm) for 60 min. Cell debris was precipitated by centrifugation (3220 *g*, 30 min). Aliquots of cleared supernatant (75 μL) were transferred to 2.2 mL 96-deep-well plates containing reaction buffer (425 μL; Tris·HCl (50 mM, pH 8.0) containing D-glucose monohydrate (11.77 mM), GDH (4 U), $NADP^+$ (0.235 mM), and substrate (2.35 mM)). The reaction plates were incubated (30 °C, 800 rpm) for 20–24 h, and ethyl acetate (400 μL) was added to each well to extract the substrate and product. After centrifugation (3220 *g*, 15 min), the organic phases (200 μL) from the wells were transferred to a new glass 96-deep-well plate, and subjected to automated GC analysis for screening. For each library, clone pools capable of producing the lactone were chosen for pooling: lysis buffer (600 μL) was added directly to wells containing the desired clone pellets (previously stored at −80 °C), and then the lysate was transferred to new 2.2 mL 96-deep-well plates. Lysis, reaction, and screening were performed as described for the pooling step. Clones capable of converting the substrate into lactone (hits) were selected. LB medium (4 mL) was inoculated with the appropriate hit clone from the −80 °C glycerol stock plates. Plasmids were extracted from overnight cultures and sequenced.

Reactions with semipurified enzymes: After sequencing, reactions were performed with semipurified enzymes of the best mutants to select the best hit to continue ISM. Semipurification was performed by inoculating LB medium (5 mL; supplemented with CB (100 μg mL^{-1})) with the corresponding PAMO mutant from the glycerol stock. After overnight incubation at 37 °C with shaking, an aliquot of this saturated culture (2 mL) was used to inoculate TB medium (100 mL; supplemented with CB (100 μg mL^{-1})) in a 500 mL Erlenmeyer flask. After incubation (3 h, 37 °C, 200 rpm) induction was performed by adding L-arabinose (0.1%). After a further 20 h of incubation, cells were harvested by centrifugation (3220 *g*, 37 °C, 15 min). Cell lysis was performed in 50 mL Falcon tubes by adding lysis buffer (4 mL) to the pellet and incubating (50 °C, 60 min). Cell debris was precipitated by centrifugation (8290 *g*, 20 min). The lysate was stored at 4 °C for 24 h, and then incubated at 50 °C for 60 min. The purity of the enzymes was assessed by SDS-PAGE. Concentration of semipurified enzyme was determined by measuring absorbance ($\varepsilon_{441} = 12.4\ mM^{-1}cm^{-1}$). Reactions were performed in Tris·HCl (50 mM, pH 8.0; 500 μL) containing semipure enzyme (5 μM), substrate (2 mM), $NADP^+$ (200 μM), GDH (2 U), and D-glucose monohydrate (10 mM) at 30 °C with agitation. After 20 h, the reaction was stopped by adding ethyl acetate (400 μL). Analysis of lactone production was performed by GC.

Experiments using resting cells and up-scaling: Induced cultures (50 mL) were prepared for WT PAMO and four PAMO mutants as described in the text. Before centrifugation, the OD_{600} of each culture was measured. Then, for each OD_{600} unit, 10 volumes of reaction buffer (Tris·HCl (50 mM, pH 8.0) with CB (100 μg mL^{-1}), glucose (20 mM), and substrate (2 mM)) were added. Reactions were performed at 30 °C in 300 mL flasks. Aliquots (600 μL) were collected after 2, 4, 6, and 8 h and extracted with ethyl acetate (400 μL). After centrifugation (8610 *g*, 15 min), the organic phase (100 μL) was transferred to a glass vial, and formation of products was monitored by GC.

Enzyme purification: Induced cultures (250 mL) were prepared for WT PAMO and PAMO mutants I67A/R258A/A442P/S444A/P286R/L443F/Y502F and I67A/R258A/A442P/S444A as described in the text. The cell pellet was resuspended in lysis buffer (25 mL; Tris·HCl (50 mM, pH 7.5), lysozyme (1.4 mg mL^{-1}), and DNase I (6 U)). Lysis was performed at 50 °C for 1 h with agitation. Cell debris was removed by centrifugation (12900 *g*, 30 min). In order to change the lysis buffer, supernatants were filtered through a 10 K MWCO membrane protein concentrator in binding buffer (50 mL; sodium phosphate (20 mM, pH 7.4), NaCl (0.5 M), imidazole (5 mM)). Proteins were loaded on HisTrap FF crude columns (5 mL; GE Healthcare) pre-equilibrated with binding buffer and with an ÄKTA purifier. PAMO WT and mutants were eluted with a gradient (12–18%) of imidazol (2 M, pH 7.4). Enzymes were desalted and concentrated by ultracentrifugation (18500 *g*) in a 10 K MWCO membrane protein concentrator with Tris·HCl (50 mM, pH 7.5). Enzyme purity was assessed by SDS-PAGE. The concentration of the purified enzyme was determined by measuring absorbance ($\varepsilon_{441} = 12.4\ mM^{-1}cm^{-1}$).

In vitro reaction: In order to see whether the undesired side product (formed by some kind of ADH-catalyzed reduction of the carbonyl function of **1**) can be suppressed in vitro, the best PAMO mutant (I67A/R258A/A442P/S444A/P286R/L443F/Y502F) was isolated and purified as described above (purification of other PAMO mutants as previously described),[20b] and used as the catalyst in the BV model reaction with the GDH/glucose NADPH regeneration system. Reaction with this purified enzyme mutant was carried out as described for semipurified enzymes, but with less GDH (0.1 U). After 16 h, the reaction was stopped by adding ethyl acetate (400 μL). Analysis of alcohol and lactone production was performed by GC measurement. Only 1% undesired reduction product was observed (99% conversion to (*E*)-**2**); diastereoselectivity was complete.

GC analyses: The organic phase was analyzed by gas chromatography. Chiral and achiral GC analyses were performed on a G-TA (30 m × 0.25 mm, 0.125 μm; Alltech) or an HP1 (30 m × 0.25 mm, 0.25 μm; Agilent Technologies) column, respectively.

Acknowledgements

We thank the Max Planck Society and the Arthur C. Cope Foundation for generous support. We are also grateful to Carlos G. Acevedo-Rocha for many helpful discussions.

Keywords: Baeyer–Villiger reaction · directed evolution · enzyme models · iterative saturation mutagenesis · stereoselectivity

[1] Recent reviews of directed evolution: a) M. B. Quin, C. Schmidt-Dannert, *ACS Catal.* **2011**, *1*, 1017 – 1021; b) E. M. Brustad, F. H. Arnold, *Curr. Opin. Chem. Biol.* **2011**, *15*, 201 – 210; c) N. J. Turner, *Nat. Chem. Biol.* **2009**, *5*, 567 – 573; d) C. Jäckel, P. Kast, D. Hilvert, *Annu. Rev. Biophys.* **2008**, *37*, 153 – 173; e) A. S. Bommarius, J. K. Blum, M. J. Abrahamson, *Curr. Opin. Chem. Biol.* **2011**, *15*, 194 – 200; f) S. Bershtein, D. S. Tawfik, *Curr. Opin. Chem. Biol.* **2008**, *12*, 151 – 158; g) P. A. Dalby, *Curr. Opin. Struct. Biol.* **2011**, *21*, 473 – 480; h) L. G. Otten, F. Hollmann, I. W. C. E. Arends, *Trends Biotechnol.* **2009**, *28*, 46 – 54; i) A. V. Shivange, J. Marienhagen, H. Mundhada, A. Schenk, U. Schwaneberg, *Curr. Opin. Chem. Biol.* **2009**, *13*, 19 – 25; j) A. Kumar, S. Singh, *Crit. Rev. Biotechnol.* **2012**, DOI: 10.3109/07388551.2012.716810; k) A. Bolt, A. Berry, A. Nelson, *Arch. Biochem. Biophys.* **2008**, *474*, 318 – 330; l) N. E. Labrou, *Curr. Protein Pept. Sci.* **2010**, *11*, 91 – 100; m) D. P. Nannemann, W. R. Birmingham, R. A. Scism, B. O. Bachmann, *Future Med. Chem.* **2011**, *3*, 809 – 819; n) N. U. Nair, C. A. Denard, H. Zhao, *Curr. Org. Chem.* **2010**, *14*, 1870 – 1882; o) S.-C. Lee, J.-H. Kim, H.-S. Kim, *Curr. Org. Chem.* **2010**, *14*, 1894 – 1901; p) R. M. P. Siloto, R. J. Weselake, *Biocatal. Agric. Biotechnol.* **2012**, *1*, 181 – 189; q) M. T. Reetz in *Enzyme Catalysis in Organic Synthesis*, 3rd ed. (Eds.: K. Drauz, H. Gröger, O. May), Wiley-VCH, Weinheim, **2012**, pp. 119 – 190.

[2] Reviews of directed evolution emphasizing library quality rather than quantity, thus leading to less screening: a) S. Lutz, W. M. Patrick, *Curr. Opin. Biotechnol.* **2004**, *15*, 291 – 297; b) M. T. Reetz, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 5716 – 5722; c) R. J. Fox, G. W. Huisman, *Trends Biotechnol.* **2008**, *26*, 132 – 138; d) A. Herman, D. S. Tawfik, *Protein Eng. Des. Sel.* **2007**, *20*, 219 – 226; e) J. D. Bloom, M. M. Meyer, P. Meinhold, C. R. Otey, D. MacMillan, F. H. Arnold, *Curr. Opin. Struct. Biol.* **2005**, *15*, 447 – 452; f) T. S. Wong, D. Roccatano, M. Zacharias, U. Schwaneberg, *J. Mol. Biol.* **2006**, *355*, 858 – 871.

[3] a) Review of directed evolution of stereoselective enzymes with emphasis on iterative saturation mutagenesis (ISM): M. T. Reetz, *Angew. Chem.* **2011**, *123*, 144 – 182; *Angew. Chem. Int. Ed.* **2011**, *50*, 138 – 174; b) Guidelines on how to avoid possible pitfalls when applying ISM: C. G. Acevedo-Rocha, S. Kille, M. T. Reetz, *Methods in Molecular Biology,* Humana, Totowa, **2013**, in press.

[4] 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.

[5] M. T. Reetz, J. D. Carballeira, *Nat. Protoc.* **2007**, *2*, 891 – 903.

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

[8] M. T. Reetz, D. Kahakeaw, R. Lohmer, *ChemBioChem* **2008**, *9*, 1797 – 1804.

[9] J. Sanchis, L. Fernández, J. D. Carballeira, J. Drone, Y. Gumulya, H. Höbenreich, D. Kahakeaw, S. Kille, R. Lohmer, J. J.-P. Peyralans, J. Podtetenieff, S. Prasad, P. Soni, A. Taglieber, S. Wu, F. E. Zilly, M. T. Reetz, *Appl. Microbiol. Biotechnol.* **2008**, *81*, 387 – 397.

[10] a) S. Kille, C. G. Acevedo-Rocha, L. P. Parra, Z.-G. Zhang, D. J. Opperman, M. T. Reetz, J. P. Acevedo, *ACS Synth. Biol.* **2012**, *2*, 83 – 92; b) L. Tang, H. Gao, X. Zhu, X. Wang, M. Zhou, R. Jiang, *BioTechniques* **2012**, *52*, 149 – 158.

[11] a) K. M. Polizzi, M. Parikh, C. U. Spencer, I. Matsumura, J. H. Lee, M. J. Realff, A. S. Bommarius, *Biotechnol. Prog.* **2006**, *22*, 961 – 967; b) D. J. Bougioukou, S. Kille, A. Taglieber, M. T. Reetz, *Adv. Synth. Catal.* **2009**, *351*, 3287 – 3305.

[12] Statistical approaches to assess oversampling in the screening process: a) A. E. Firth, W. M. Patrick, *Bioinformatics* **2005**, *21*, 3314 – 3315; b) A. E. Firth, W. M. Patrick, *Nucleic Acids Res.* **2008**, *36*, W281 – W285; c) M. Denault, J. N. Pelletier in *Protein Engineering Protocols, Vol. 352* (Eds.: K. M. Arndt, K. M. Müller), Humana Press, Totowa, NJ, **2007**, pp. 127 – 154; d) J. F. Chaparro-Riggers, K. M. Polizzi, A. S. Bommarius, *Biotechnol. J.* **2007**, *2*, 180 – 191; e) Y. Nov, *Appl. Environ. Microbiol.* **2012**, *78*, 258 – 262.

[13] a) M. T. Reetz, S. Prasad, J. D. Carballeira, Y. Gumulya, M. Bocola, *J. Am. Chem. Soc.* **2010**, *132*, 9144 – 9152; b) Y. Gumulya, J. Sanchis, M. T. Reetz, *ChemBioChem* **2012**, *13*, 1060 – 1066.

[14] M. T. Reetz, *Angew. Chem.* **2013**, *125*, 2720 – 2729; *Angew. Chem. Int. Ed.* **2013**, *52*, 2658 – 2666.

[15] The computational tools CASTER and B-FITTER (useful in the design of saturation mutagenesis libraries) are available free of charge from the author: http://www.kofo.mpg.de/en/research/organic-synthesis.

[16] M. T. Reetz, S. Wu, *Chem. Commun.* **2008**, 5499 – 5501.

[17] A. G. Sandström, Y. Wikmark, K. Engström, J. Nyhlén, J.-E. Bäckvall, *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 78 – 83.

[18] Reviews of Baeyer–Villiger monooxygenases as catalysts in organic chemistry: a) M. M. Kayser, *Tetrahedron* **2009**, *65*, 947 – 974; b) J. D. Stewart, *Curr. Org. Chem.* **1998**, *2*, 195 – 216; c) M. D. Mihovilovic in *Enzyme Catalysis in Organic Synthesis*, 3rd ed. (Eds.: K. Drauz, H. Gröger, O. May), Wiley-VCH, Weinheim, **2012**, pp. 1439 – 1485; d) V. Alphand, G. Carrea, R. Wohlgemuth, R. Furstoss, J. M. Woodley, *Trends Biotechnol.* **2003**, *21*, 318 – 353; e) M. D. Mihovilovic, *Curr. Org. Chem.* **2006**, *10*, 1265 – 1287; f) R. Wohlgemuth, *Eng. Life Sci.* **2006**, *6*, 577 – 583; g) D. E. Torres Pazmiño, H. M. Dudek, M. W. Fraaije, *Curr. Opin. Chem. Biol.* **2010**, *14*, 138 – 144; h) H. Leisch, K. Morley, P. C. K. Lau, *Chem. Rev.* **2011**, *111*, 4165 – 4222; i) F. Hollmann, I. W. C. E. Arends, K. Buehler, A. Schallmey, B. Bühler, *Green Chem.* **2011**, *13*, 226 – 265.

[19] Isolation and substrate range of PAMO: a) M. W. Fraaije, J. Wu, D. P. H. M. Heuts, E. W. van Hellemond, J. H. L. Spelberg, D. B. Janssen, *Appl. Microbiol. Biotechnol.* **2005**, *66*, 393 – 400; b) F. Zambianchi, M. W. Fraaije, G. Carrea, G. de Gonzalo, C. Rodriguez, V. Gotor, G. Ottolina, *Adv. Synth. Catal.* **2007**, *349*, 1327 – 1331; c) C. Rodriguez, G. de Gonzalo, D. E. Torres Pazmiño, M. W. Fraaije, V. Gotor, *Tetrahedron: Asymmetry* **2008**, *19*, 197 – 203; d) Z.-G. Zhang, L. P. Parra, M. T. Reetz, *Chem. Eur. J.* **2012**, *18*, 10160 – 10172.

[20] a) M. Bocola, F. Schulz, F. Leca, A. Vogel, M. W. Fraaije, M. T. Reetz, *Adv. Synth. Catal.* **2005**, *347*, 979 – 986; b) M. T. Reetz, S. Wu, *J. Am. Chem. Soc.* **2009**, *131*, 15424 – 15432.

[21] Z.-G. Zhang, G.-D. Roiban, J. P. Acevedo, I. Polyak, M. T. Reetz, *Adv. Synth. Catal.* **2013**, *355*, 99 – 106.

[22] E. Malito, A. Alfieri, M. W. Fraaije, A. Mattevi, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 13157 – 13162.

[23] a) I. Polyak, M. T. Reetz, W. Thiel, *J. Am. Chem. Soc.* **2012**, *134*, 2732 – 2741; b) I. Polyak, M. T. Reetz, W. Thiel, *J. Phys. Chem. B* **2013**, *117*, 4993 – 5001.

Received: July 29, 2013
Published online on October 18, 2013