Protein science

DOI: 10.1002/anie.201207842

The Importance of Additive and Non-Additive Mutational Effects in Protein Engineering**

*Manfred T. Reetz**

directed evolution · enzymes · mutational additivity · mutational non-additivity · saturation mutagenesis

1. Introduction

Following the development of site-specific mutagenesis by Michael Smith more than three decades ago,[1] scientists have at their disposal a molecular biological method with which a given amino acid at a defined position of a protein can be substituted by any one of the other 19 canonical representatives.[2] This technique has proven to be invaluable in the study of enzyme mechanisms.[1,3] The enormous progress made in attempting to understand the intricacies of enzyme mechanisms by using this technique coupled with an arsenal of other experimental and computational methods developed during the last few decades has shown that the "secrets" behind these biocatalysts are more complex than originally thought. In addition to clearing up the gross features of a given enzyme mechanism, phenomena such as protein dynamics and flexibility, electrostatic preorganization, tunneling, and allosteric effects need to be considered and evaluated case by case.[4] This makes the detailed interpretation and prediction of mutational effects difficult and, as highlighted in this Essay, especially challenging if several point mutations are introduced.

Amino acid exchange events are also at the heart of protein engineering directed at improving the performance of enzymes as catalysts in synthetic organic chemistry and biotechnology, with increasing thermostability and stereoselectivity being of particular practical interest. Binding properties and expression rate can also be targeted. One option is "rational design" based on site-specific mutagenesis, which has proven to be successful in some cases,[2b,5] but not in a general way with present-day theoretical and computational techniques. The second option is directed evolution,[2b,6] which utilizes random or structure-based focused mutagenesis. It involves repeating cycles of gene mutagenesis, expression, and screening (or selection). The most often used gene mutagenesis methods in directed evolution are the error-prone polymerase chain reaction (epPCR, a "shotgun" technique),[2b,6] saturation mutagenesis (a combinatorial structure-based rational approach),[2,6,7] and DNA shuffling.[2b,6]

Directed evolution can always be expected to provide mutants (hits) with improved catalysis profiles. The degree of improvement depends upon the amount of mutagenesis and screening effort that the experimenter is willing to invest. Various approaches for maximizing the quality of mutant libraries and hence to reduce the screening effort (bottleneck of directed evolution) have been developed.[2b,6,8] It is currently impossible to predict how many point mutations are required for a given task, which is likely to depend, amongst other things, upon the mutagenesis method. Rational design and evolution in the laboratory have also been employed in an effort to understand and manipulate protein–protein interactions, an area of increasing importance.[9,10] In these cases, multiple mutations on the surface of the proteins are generally involved.

When attempting to understand the effect of mutational changes on a molecular level, the interpretation or prediction of a single mutation in a wild-type (WT) enzyme constitutes the simplest case. If two or more point mutations are introduced, the situation becomes considerably more complicated, because the question of additivity or non-additivity arises. Do the effects of two separate point mutations add up mathematically when they are combined in a double mutant? Or do they interact with one another in a non-additive manner and cause either cooperative (positive) or antagonistic (negative) effects?

This Essay does not constitute a comprehensive treatment of additive and non-additive mutational effects in proteins. Instead, illustrative examples from the literature are used to remind biochemists, molecular biologists, and biotechnologists of the need to focus on non-additivity as a challenging problem in diverse areas of protein science. Emerging structural insights regarding the origin of non-additive effects as well as future challenges are also highlighted.

[*] Prof. Dr. M. T. Reetz
Max-Planck-Institut für Kohlenforschung
Kaiser-Wilhelm-Platz 1, 45470 Mülheim an der Ruhr (Germany)
and
Philipps-Universität Marburg
Hans-Meerweinstrasse, 35032 Marburg (Germany)
E-mail: reetz@mpi-muelheim.mpg.de
Homepage: http://www.kofo.mpg.de/de/forschung/organische-synthese

[**] I thank all past and present co-workers who have participated in discussions regarding non-additive effects, especially J. D. Carballeira, J. Sanchis, C. Acevedo-Rocha, and S. Kille, as well as Prof. Yuval Nov (Haifa/Israel) for helpful comments. Support by the Max-Planck-Society and the Arthur C. Cope foundation is gratefully acknowledged.

2. Early Examples of Additivity

Early mechanistic studies by the research groups of Fersht,[11] Wells,[12] Horovitz,[13a] Shortle,[13b,c] and others,[13d–f] in which two point mutations were introduced in WT enzymes and then combined with formation of the double mutant, pointed to the prevalence of additivity when addressing enzyme activity or protein–protein interactions. However, exceptions were noted. A typical example of additivity is the case of the functional independence of two point mutations influencing the activity of tyrosyl-tRNA synthetase, as reported by Wells et al.[12] Based on the work of Fersht and co-workers,[11] they proposed in 1990 a simple mathematical expression describing such additive behavior in double-mutant free-energy cycles [Eq. (1)].[12b] Here, $\Delta\Delta G_{(X,Y)}$ represents the experimental change in free energy upon going from the WT to the double mutant (X,Y), while $\Delta\Delta G_{(X)}$ and $\Delta\Delta G_{(Y)}$ are the free-energy changes upon going from the WT to the single mutants X and Y, respectively; ΔG_I reflects the interaction of the two point mutations in rare cases. It was suggested that such non-additivity may occur when the side chains of two residues are in close contact with one another or when one or both of the point mutations cause a change in the rate-limiting step or a switch in the reaction mechanism. The results from several studies regarding the effect of mutational changes on activity using five different enzymes clearly pointed to additivity (Figure 1).[12b] The data chosen for the construction of this correlation originated from point mutations in which the distance between the respective amino acid side chains exceeded the van der Waals contact (> 4 Å).

$$\Delta\Delta G_{(X,Y)} = \Delta\Delta G_{(X)} + \Delta\Delta G_{(y)} + \Delta G_I \quad (1)$$

Two years later Mildvan et al. "dissected" the Fersht–Wells Equation, and suggested a set of terms which include additive effects, partially additive effects, synergistic effects, antagonistic effects, or no effects.[14a] Examples of all the cases from the literature were considered, and led to the conclusion that "*additive effects are the easiest to interpret and are commonly found, usually when the mutated residues are not in molecular contact*". Later Mildvan introduced the concept of "inverse thinking" for double mutant cycles of enzymes, in which the parameters of the double mutant instead of those of the WT enzyme are used as the reference point.[14b]

Manfred T. Reetz was born in 1943 in Germany and obtained his PhD in 1969 under the guidance of Ulrich Schöllkopf at Göttingen University. After postdoctoral studies with Reinhard W. Hoffmann and Habilitation in 1974 at Marburg University, he moved to the University of Bonn before becoming Full Professor of organic chemistry in Marburg (1980–1991). After 20 years as Director at the MPI für Kohlenforschung in Mülheim, he returned to Marburg as the first Hans-Meerwein-Research-Professor, financed by the Max-Planck-Society. His research focuses on the directed evolution of stereoselective enzymes and combinatorial transition-metal catalysis.

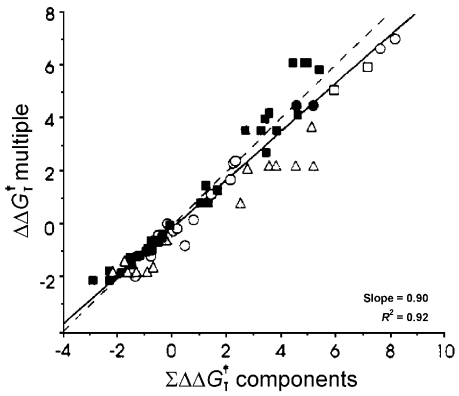


Figure 1. Changes in transition-state stabilization energies for the multiple mutant versus the sum of the component mutants.[12b] The data represent mutants from subtilisin (■), tyrosyl-tRNA synthetase (○), trypsin (□), DHFR (●), and glutathione reductase (Δ). The dashed line has a slope of 1, and the solid line corresponds to the best fit.

Following these early reports, further examples of mutational additivity accumulated. For example, an unambiguous case was reported by Skinner and Terwilliger in a 1996 study entitled "*Potential use of additivity of mutational effects in simplifying protein engineering*".[15] The influence of mutations on the stability and DNA binding affinity of the gene V protein (GVP) was systematically investigated, with all mutants as well as WT GVP being characterized by X-ray crystallography. Additivity as a consequence of separate localized effects was found for the two unrelated parameters (stability and DNA binding affinity). In 2000 Uchiyama and co-workers published an analysis of certain local fitness landscapes based on a rough Mt. Fuji-type model, with prolyl endopeptidase and thermolysin serving as the enzymes.[16] This study culminated in the conclusion that "*current biopolymers have statistical additivities*". The authors refer to a number of supporting studies regarding additivity and emphasize that "*If experiments are performed focusing on residues located at the surface of a protein and oriented outwardly, the mutual interference between residues changed as a result of multiple mutations must be small*".[16] The question whether this dogma remains valid today needs to be addressed.

Scientists at Novozyme reported additive effects when applying the consensus mutations concept to engineering proteins with an enhanced thermostability,[17a] as have Magliery and co-workers in a related, more recent study.[17b] Additive effects have also been noted or assumed in other studies.[17c–e] Combining positive mutations in a successful attempt to increase the thermostability of *Bacillus licheniformis* is an example.[17c] We have discovered in a directed evolution study that combining two positive mutations may indeed provide improved variants, but not in all cases (less than additivity)—which is an important finding.[18] Perhaps this kind of non-additivity is not routinely reported in directed evolution studies because such results do not reflect practical progress.

3. Examples of Mutational Non-Additivity in Mechanistic and Theoretical Studies

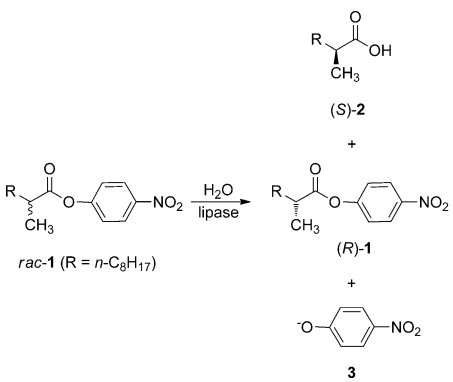
In the 1990s, studies regarding additive versus non-additive effects in protein engineering for the purpose of understanding enzyme mechanisms in greater detail began to appear.[13c] An early example is reported to Fersht and co-workers, who exploited double mutant cycles in their quest to estimate engineered surface electrostatic interactions in the ribonuclease from *Bacillus amyloliquefacies* (barnase).[19] Two residues in an α helix on the enzyme surface, Asp12 (charged) and Thr16 (polar), were mutated to Asp12Ala and Thr16Arg. The change in the free energy of reversible urea-induced unfolding relative to WT barnase for the three mutants was measured at different salt concentrations. Non-additivity was observed: for example, single mutant Asp12Ala (interaction energy = 0.43 kcal mol^{-1}), single mutant Thr16Arg (−0.48 kcal mol^{-1}), and double mutant Asp12Ala/Thr16Arg (0.28 kcal mol^{-1}).[19] Direct electrostatic interactions were invoked to explain the results. Subsequently, the same research group focused on protein–protein interactions by studying double mutant cycles involving barnase and its intracellular inhibitor barstar (a small protein).[10] In this case the coupling energy of two residues decreased as the distance between them increased, with maximum cooperativity being observed at distances less than 7 Å.

Another prime example of non-additivity was provided by the study of Benkovic and co-workers in 1995, who investigated the mutational changes in *Escherichia coli* dihydrofolate reductase (DHFR) and related these to the mechanism of the enzyme.[20] DHFR catalyzes the reduction of 7,8-dihydrofolate to 5,6,7-tetrahydrofolate using NADPH as a cofactor. From earlier studies it was known that the folate binding site is characterized by a 15 Å deep cavity. It is lined with the hydrophobic side chains of Leu28, Phe31, Ile50, and Leu54, which were known to interact with the *p*-aminobenzoyl-L-glutamate of folate. To shed light on the relationship between the binding pocket and catalytic function, several mutants were prepared by site-directed mutagenesis, including single mutants L28Y and L54F, as well as all the theoretically possible double mutants, including the combination L28Y/L54F. It was noted that residues L28 and L54 reside on different substructures at the active site and do not appear to interact with one another.[20] Three interrelated mutational cycles were analyzed, which allowed the identification of specific side-chain interactions in each catalytic step. As a consequence of the relatively large spatial separation of the respective side chains of residues 28 and 54 (ca. 8 Å), additive effects were expected for all parameters [$\Delta G_I = 0$ in Eq. (1)]. At some points in the scheme, additivity or partial additivity was in fact observed. However, significant non-additivity reflecting cooperativity as well as antagonism were detected in the thermodynamics of ligand binding and in the rate of hydride transfer and product release.[20] Conformational changes during the catalytic cycle were proposed. Although synergistically favorable double mutations were not found, the authors concluded that "*with an enhanced knowledge of the molecular origin of nonadditive effects, it may be possible to optimize an approach to improve the enzyme's efficiency by coupling mutations*".[20] While this goal has not been reached to date, the statement points the direction in which to go, namely to generate knowledge concerning the origin of non-additivity, which will hopefully lead to some degree of predictive power.

4. Examples of Mutational Non-Additivity in Directed Evolution

The question of strict additivity, as opposed to a cooperative or an antagonistic type of non-additivity, in directed evolution has seldom been addressed. Most researchers focus on obtaining improved enzymes for practical applications and are less interested in time-consuming deconvolution experiments.[2b,6] My research group has found such extra efforts to be rewarding because they reflect, for example, the efficiency of a given mutagenesis strategy.[21] The bottleneck in directed evolution is the screening step, which is the reason for the increased interest in developing methods for the generation of high-quality libraries, which require less screening.[2b,6,8,21] Our contribution in this regard is iterative saturation mutagenesis (ISM),[21,22] which led to the discovery of numerous cases of non-additive effects in the directed evolution of stereoselective lipases, epoxide hydrolases, enoate reductases, Baeyer–Villiger monooxygenases, and P450 enzymes.[6q,21] Only a few representative examples of non-additivity from our studies and from other research groups are presented here, all of which uncover what initially appeared to be puzzling results.

In a proof-of-principle investigation concerning the directed evolution of enzyme stereoselectivity, the hydrolytic kinetic resolution of the ester *rac*-**1** using the lipase from *Pseudomonas aeruginosa* (PAL) was studied [Eq. (2)].[23a] It is characterized by the catalytic triad typical of lipases, Asp229, His251, and Ser82, the latter adding nucleophilically to the carbonyl function with rate-determining formation of the short-lived tetrahedral oxyanion intermediate. WT PAL is only slightly enantioselective in favor of (*S*)-**2**, with the selectivity factor amounting to a mere $E = 1.2$. Four successive cycles of epPCR provided a quadruple mutant with $E = 11$. After testing various alternative mutagenesis strategies, the combination of epPCR, saturation mutagenesis, and DNA shuffling generated a superior mutant characterized by six point mutations D20N/S53P/S155M/L162G/T180I/T234S,



(2)

with high *S* enantioselectivity ($E = 51$) being achieved.[23b] Surprisingly, only one mutation, Leu162Gly, proved to be near the binding pocket, while the other five were found on the surface of PAL.

This concept was subsequently generalized by us and by other research groups for the directed evolution of other stereoselective enzymes, and provided a new approach to generate catalysts for asymmetric transformations in organic chemistry.[2b,6,21] However, a quantum and molecular mechanics (QM/MM) study predicted that only two of the six point mutations should be important for enhancing the stereoselectivity, namely S53P on the enzyme's surface and L162G aligning the binding pocket.[24] Indeed, the double mutant S53P/L162G proved to be even more selective ($E = 64$). This was a triumph of theory, but it also demonstrated that the strategies based mainly on epPCR and DNA shuffling were far from optimal, since four superfluous mutations accumulated during the evolutionary process.

As a consequence of this study, we focused on method development in directed evolution.[6q,21,22] Although not anticipated at the time, ISM proved to be an ideal approach to protein engineering which induces pronounced cooperative effects operating between point mutations and sets of mutations (non-additivity). Appropriate single or multiple residue sites in an enzyme, comprising one or more amino acid positions, are first randomized with formation of focused libraries. The gene of a given hit is then used as a template for performing saturation mutagenesis at the other sites, and the process is continued until the desired degree of catalyst improvement has been achieved. When addressing stereo- or regioselectivity, activity, and/or substrate scope, sites lining or near the binding pocket are chosen in a process originally used in an unsystematic manner[23b] and later used more systematically as a combinatorial active-site saturation test (CAST).[22a] This is a convenient acronym for a strategy previously applied for tuning or reversing the enantioselectivity,[23b] increasing the activity, as well as controling the regioselectivity.[7] When aiming for thermostabilization, the sites for ISM are chosen by focusing on those residues which have the highest B factors (reflecting the degree of thermal motion) in a process dubbed B-FIT (B-factor iterative test).[22b] In both forms of ISM, the efficacy can be improved even more by using reduced amino acid alphabets, which minimizes the screening effort.[6q,21] Methods for reducing or eliminating amino acid bias caused by the degeneracy of the genetic code[25] also help to reduce the amount of screening, as do pooling techniques.[6e,25d] These techniques as well as appropriate statistical analyses of the degree of necessary oversampling,[8b–f] especially the Nov approach,[8e] are likewise of significant value in ISM studies.

It was essential to apply ISM to the lipase PAL,[26] because this allowed rigorous comparison with the earlier extensively applied mutagenesis methods.[23c] By using three CAST sites, A (Met16/Leu17), B (Leu159/Leu162), and C (Leu231/Val232), and appropriately chosen reduced amino acid alphabets while screening only small libraries along several of the six theoretically possible ISM pathways, a highly active and stereoselective triple mutant Met16Ala/Leu17Phe/Leu162Asn emerged along the trajectory B→A, which

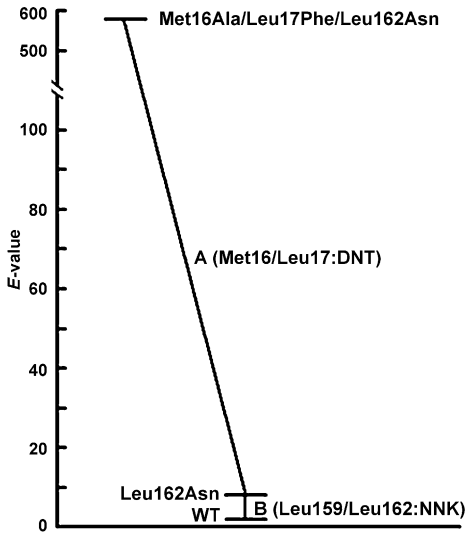


Figure 2. ISM scheme of the optimal pathway WT PAL→B→A.[26]

showed a selectivity factor of $E = 594$ (*S*; Figure 2).[26] In view of this excellent result, continuing to site C was not necessary. The mutant is clearly superior to the previous best hit produced by directed evolution obtained mainly by epPCR and DNA shuffling,[23b] while requiring significantly less screening (10000 versus 50000 transformants).

The superior performance of ISM demonstrated in the final PAL study was traced to strong cooperative epistatic effects, as revealed by deconvolution.[26] The first-round mutant Leu162Asn showed very low stereoselectivity ($E = 8$), and even lower selectivity was observed by deconvoluting the triple mutant and testing the double mutant Met16Ala/Leu17Phe ($E = 2.6$). Thus, the combined two sets of mutants finally leading to $E = 594$ interact in a strongly cooperative non-additive manner to the extent of about 8 kJ mol^{-1}. Cooperativity is also reflected in rate data, as shown by the kinetics of the favored reaction using (*S*)-**1**: Leu162Asn ($K_M = 5.6 \times 10^{-4}$ M; $k_{cat} = 42.1 \times 10^{-3}$ s^{-1}; $k_{cat}/K_M = 75.2$ s^{-1}M^{-1}); Met16Ala/Leu17/Phe ($K_M = 22.0 \cdot 10^{-4}$ M; $k_{cat} = 16.1 \times 10^{-3}$ s^{-1}; $k_{cat}/K_M = 7.3$ s^{-1}M^{-1}); and Met16Ala/Leu17Phe/Leu162Asn ($K_M = 3.4 \times 10^{-4}$ M; $k_{cat} = 1374 \times 10^{-3}$ s^{-1}; $k_{cat}/K_M = 4041$ s^{-1}M^{-1}). The results also prove that superfluous mutations are absent,[26] in contrast to the previous mutant characterized by six point mutations.[23b] Bäckvall and coworkers also observed non-additive effects on applying an extremely condensed saturation mutagenesis library to the directed evolution of the lipase CALA.[27]

Insight into the source of non-additivity in the PAL study was gained by performing molecular dynamics (MD) simulations and docking experiments.[26] Mutation Leu162Asn provides more space for the substrate to bind, primarily because the side chain of asparagine is found as a rotamer in which it "shifts" away from the binding pocket by forming a hydrogen bond with Ser158 and His83. This means that, in effect, Asn162 is a much smaller residue that Leu162, which is characterized by a somewhat space-filling hydrophobic isobutyl side chain pointing into the binding pocket of WT PAL. However, the creation of more space does not necessarily

lead to higher activity and enhanced stereoselectivity. This is where the mutational set Met16Ala/Leu17Phe comes into play. In the case of (*S*)-**1**, additional stabilization of the oxyanion is induced by new hydrogen bonds originating from His83 and Ala16, while Phe17 undergoes π,π interactions (Figure 3). These effects are not possible in the case of the disfavored (*R*)-**1** for steric reasons.[26] Analyses of this kind are necessary when striving to understand the origin of non-additivity on a molecular level.

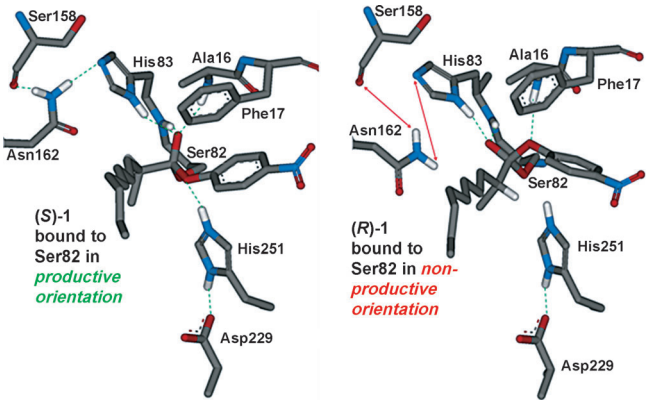
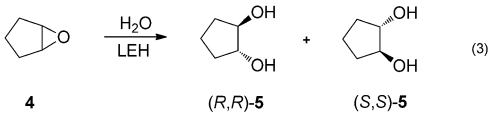


Figure 3. Comparison of oxyanion-stabilizing hydrogen bonds (green lines) in PAL mutant Met16Ala/Leu17Phe/Leu162Asn. Left: in a productive mode for (*S*)-**1**; right: in a nonproductive mode for (*R*)-**1**.[26]

We have consistently also found non-additivity when applying ISM to other types of enzymes.[6q,21] This was also observed by Bornscheuer and co-workers on applying CAST to an esterase as a catalyst in the hydrolytic kinetic resolution of racemic tertiary acetates.[28] The double mutant Glu188Trp/Met193Cys led to *S* selectivity ($E = 64$) in a particular case, but the deconvolution led to a surprising result: Whereas mutant Glu188Trp showed the expected moderate *S* selectivity ($E = 26$), the other single mutant Met193Cys proved to be *R* selective ($E = 16$).[28] If the reverse process had been relevant, namely, if the single mutants had been generated separately, one would never have combined an *S*- and an *R*-selective point mutation to achieve higher *S* selectivity!

Even more extreme is a puzzling observation made upon deconvoluting the first-generation *S*,*S*-selective double mutant Met32Leu/Leu35Phe of limonene epoxide hydrolase (LEH) obtained by saturation mutagenesis; this enzyme acts as a catalyst in the desymmetrization of cyclopentane oxide **4**→**5** [24% *ee*; Eq. 3].[29] The two respective single mutants Met32Leu and Leu35Phe were found to be *R*,*R* selective (14% and 2% *ee*, respectively)! The cooperative interaction between the two point mutations amounts to 2.0 kJ mol^{-1}. In the case of a much better first-generation *S*,*S*-selective double mutant, the respective deconvoluted single mutants are both *S*,*S*-selective, but they communicate with each other in a synergistic manner to the extent of 2 kJ mol^{-1}. In subsequent ISM experiments, enantioselectivities of up to 93% *ee* were achieved, and selectivities with other prochiral epoxides reached 99% *ee*.[29] A model was proposed to explain cooperative non-additivity on a molecular level.[29] However, it would have been difficult to predict such a behavior.



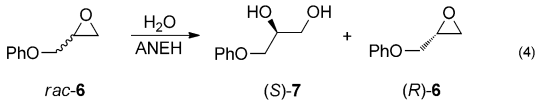
Livesay and co-workers have obtained new insight into non-additivity within protein double-mutant free-energy cycles with an emphasis on long-range effects.[30] They conclude that a statistically significant bias toward non-additivity occurs whenever the two residues, although not in direct contact, are located within the same rigid cluster, whereas additivity can be expected when they are in different clusters. It will be interesting to see whether this prediction holds up as more data become available.

5. Non-Additive Mutational Effects as Revealed by Fitness Landscapes

Directed evolution generally requires *n* cycles of mutagenesis/expression/screening that lead to the accumulation of *n* sets of mutations, with each set comprising one or more point mutations.[2b,6,21] With the exception of the initial set, the separate catalytic performance of each of the subsequent sets remains unknown unless deconvolution is performed. Moreover, systematic deconvolution by preparing and testing all the combinations of mutational sets allows information regarding additivity or non-additivity to be gained at all stages of the evolutionary process. When, for example, five successive rounds of mutagenesis are performed, then $n = 5$, and $5! = 120$ possible pathways from the WT to the final originally evolved (best) mutant exist. The complete deconvolution data enables the construction of a type of fitness landscape which enables additive and non-additive effects operating between all combinations of mutational sets at all stages of a given pathway to be identified. We have described this type of fitness landscape as being "constrained", because at each evolutionary stage new mutants are involved, but not new point mutations. An early example was reported by Weinreich et al., who considered a β-lactamase mutant characterized by five point mutations previously obtained in a stepwise manner by DNA shuffling, and constructed the 120 relevant pathways from the WT to the final mutant.[31] Additive and non-additive effects on β-lactamase activity in a selection system were identified. Of the 120 pathways, 102 were reported to be inaccessible because of the occurrence of local minima, thereby leading the authors to conclude that "*few pathways in Darwinian evolution lead to fitter proteins*".[31]

We have taken a different approach to constructing "constrained" fitness landscapes, which likewise reveal mutational additive as well as non-additive behavior. In the initial ISM study concerning the evolution of stereoselective mutants of the epoxide hydrolase from *Aspergillus niger*

(ANEH) as catalysts in the hydrolytic kinetic resolution of *rac*-**6** [Eq. (4)], five CAST sites were considered in a five-step upward climb.[22a] Whereas WT ANEH is only slightly *S*-selective ($E = 4.6$), the final mutant LW202 characterized by five sets of point mutations led to a selectivity factor of $E = 115$ (*S*).



(4)

In a follow-up study, systematic deconvolution of the five sets of mutations obtained at the five CAST sites B, C, D, E, and F (site A was not considered) allowed the construction of 120 pathways to link WT ANEH with the best variant LW202 (Figure 4).[32] None of the five sets of mutations proved to be superfluous. Moreover, the majority (55) of pathways turned out to be energetically favorable, as they did not contain local minima.

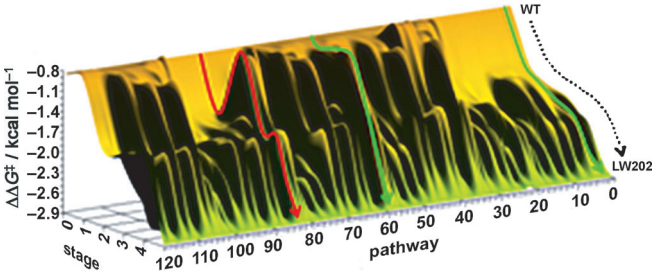


Figure 4. Energy profile of the two types of pathways leading from WT ANEH to mutant LW202 in the model reaction involving epoxide *rac*-**6**: Energetically favored (green) as in the original B→C→D→F→E (pathway 2) or D→C→F→E→B (pathway 60) and disfavored (red) as in E→C→F→D→B (pathway 84).[32]

The experimental results allowed the free energy of interaction $\Delta G^{\#}(i,j)$ between any two sets of mutations *i* and *j* in any one of the 120 pathways to be calculated[32] [Eq. (5)],

$$\Delta G^{\#}(i,j) = \Delta\Delta G^{\#}(\mathrm{exp}) - [\Delta\Delta G^{\#}(i) + \Delta\Delta G^{\#}(j)] \quad (5)$$

where $\Delta\Delta G^{\#}(\mathrm{exp})$ is the difference in the activation energy between the two enantiomers obtained experimentally for the binary combination, and $\Delta\Delta G^{\#}(i)$ and $\Delta\Delta G^{\#}(j)$ are the experimental energies obtained for each set of mutants separately. $\Delta G^{\#}(i,j)$ is a measure of cooperative effects [$\Delta G^{\#}(i,j) < 0$], of additive effects [$\Delta G^{\#}(i,j) = 0$ (no interaction)], of partially additive effects [$\Delta G^{\#}(i,j) > 0$ and $|\Delta\Delta G^{\#}(i)|$ and $|\Delta\Delta G^{\#}(j)| < |\Delta\Delta G^{\#}(\mathrm{exp})|$], or it is a measure of antagonistic effects [$\Delta G^{\#}(i,j) > 0$ and $|\Delta\Delta G^{\#}(i)|$ or $|\Delta\Delta G^{\#}(j)| > |\Delta\Delta G^{\#}(\mathrm{exp})|$].

The result of such a procedure is shown in Figure 5 for the original trajectory B→C→D→F→E. It can be seen that non-additive behavior pertains at all the evolutionary stages. For example, mutational set D alone actually reduces the enantioselectivity in an antagonistic manner, but in concert with those at B, C, and F at stage BCDF results in more than additivity. The other 119 pathways reveal similar non-additive effects, which means that the sets of mutations interact with one another, although the respective residues are not always direct neighbors in the enzyme. Exact additivity is just one point on a scale of many possibilities.

Application of Mildvan's technique of "inverse thinking"[14] to the ANEH data likewise uncovered non-additive effects.[32] Non-additivity appears to play a crucial role in ISM, which helps to explain its efficacy.[21] Unveiling how it functions on a molecular level is a fascinating task, which needs to be performed on a case by case basis by using

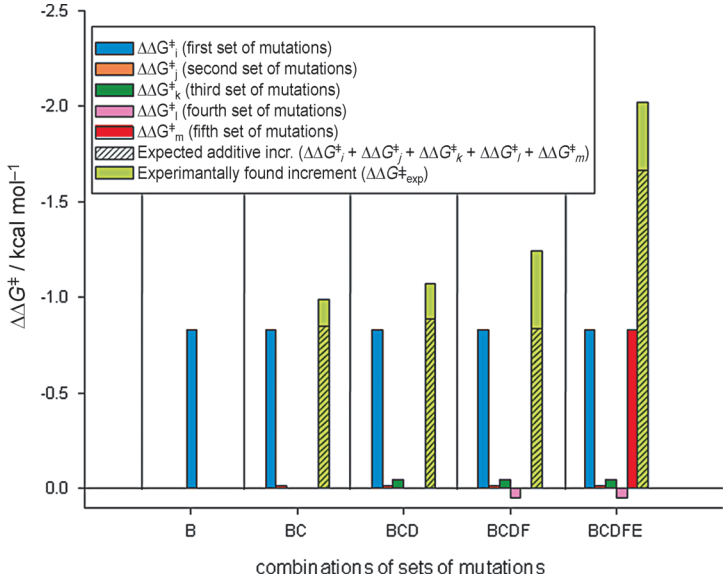


Figure 5. Thermodynamic cycle [Eq. (5)] for the interaction of the sets of mutations involved at every stage along the energetically favored pathway B→C→D→F→E.[32]

biophysical methods in combination with QM/MM calculations.

The number of favored pathways identified in the ANEH study[32] that lack local minima is different from that in the Weinreich case.[31] The percentage of favorable trajectories is likely to depend upon the mutagenesis method employed in a given experimental setup and upon the particular enzyme. We suggest that the experimenter can learn from natural evolution, but that the reverse, namely the use of data from laboratory evolution to make generalizations regarding true Darwinian evolution in nature, is problematic, although not impossible if performed carefully.

Even greater cooperative effects were encountered in a B-FIT/ISM-based study of the heat stabilization of the lipase from *Bacillus subtilis*.[22b,33] Biophysical characterization of the best mutants, including by protein NMR spectroscopy, uncovered the factors contributing to the increased retention of activity at high temperatures, which point to inhibition of protein–protein interactions leading to undesired aggregation and precipitation upon heating.[33a] Cooperative long-range interactions between point mutations on the surface of the lipase mutants are in line with the dramatic non-additive effects.[33b]

Recently, we reported an "unconstrained" system by using ANEH and the same model reaction in which all $4! = 24$ pathways of a four-site ISM system were explored.[34] This allowed amino acid exchange events at every stage, as defined by the applied codon degeneracy (reduced amino acid alphabet). We found the majority of trajectories to be energetically favorable (16 of the 24), with no local minima. In the case of the eight disfavored pathways, it was possible to escape from local minima by choosing an inferior mutant in the respective library as a template for continuing the upward climb.[34] Failing to find an improved mutant in a given library traditionally forces the experimenter to increase the size of the library, to end the search, or to try a different mutagenesis strategy.[2b,6] Our study shows that it might well be useful in such cases to utilize deleterious mutations as stepping stones,[34] the success of which is probably due to cooperative effects (non-additivity). Neutral drift[35] and/or the Eigen–Schuster concept of quasi species,[36] as invoked by Mannervik and co-workers[37] in other directed evolution studies, may well involve non-additive effects.[34]

5. Conclusions and Perspectives

This Essay is not meant to be a comprehensive treatment of additivity or non-additivity in protein engineering studies. Rather, its purpose is to point out the crucial role of non-additivity in protein engineering, with the analysis being based on a select number of examples from the literature. It is hoped that a deeper awareness of the importance of non-additive effects in the entire field of protein research will emerge. If additivity were to be universal, it would be easier (although not trivial) to develop a general theory of functional proteins as part of a linear system. However, the increasing number of cases describing non-additivity shows that proteins are more complicated than one might wish to assume. Making reliable predictions in such nonlinear regimes is difficult. It is also not fully clear whether true Darwinian evolution in nature utilizes non-additive effects, although it is likely. Evolutionary biologists have recently demonstrated that epistasis, not just adaptation, is a major factor in molecular evolution.[38]

Directed evolution provides an ideal tool for studying non-additive effects. Hopefully, scientists employing this form of protein engineering in future research will invest more time and effort in uncovering potential non-additivity, as revealed by deconvolution studies of the type featured in this Essay.[21,26–29,32–34] With the emergence of more data coupled with detailed biophysical characterization of all mutants appearing in an upward evolutionary climb as well as QM/MM analyses, it will be possible to deepen our understanding of non-additivity on a molecular level. This will also increase our knowledge of functional proteins, and ultimately lead to more efficient rational design. Unveiling cooperative or antagonistic mutational effects as a function of the mutagenesis method and strategy also helps to assess the viability and efficacy of the different approaches currently known for protein tuning.[2b,6,21] Iterative saturation mutagenesis (ISM) has emerged as a particularly efficient approach, in which cooperative non-additivity plays a crucial role, with the probability of such effects being high when this method is applied.[21] It will be interesting to see whether de novo computational design of enzymes can be extended to include such cooperative mutational effects.[2b,39] One possibility would be to subject such enzymes to further tuning by applying ISM. Harnessing the full power of non-additivity in protein engineering[40] continues to be a rewarding goal from a mechanistic, theoretical, and practical viewpoint.

Received: September 28, 2012
Revised: December 19, 2012
Published online: February 4, 2013

[1] a) C. A. Hutchison III, S. Phillips, M. H. Edgell, S. Gillam, P. Jahnke, M. Smith, *J. Biol. Chem.* **1978**, *253*, 6551–6560; b) Nobel Lecture: M. Smith, *Angew. Chem.* **1994**, *106*, 1277–1284; *Angew. Chem. Int. Ed. Engl.* **1994**, *33*, 1214–1221.

[2] a) The QuikChange protocol of Stratagene is widely used for site-specific mutagenesis: H. H. Hogrefe, J. Cline, G. L. Youngblood, R. M. Allen, *BioTechniques* **2002**, *33*, 1158–1165; b) other methods for saturation mutagenesis, such as megaprimer and overlap extension PCR, as well as various assembly strategies can also be used:[7] S. Lutz, U. T. Bornscheuer, *Protein Engineering Handbook*, *Vol. 1–2*, Wiley-VCH, Weinheim, **2009**.

[3] A. Fersht, *Structure and Mechanism in Protein Science*, W. H. Freeman, New York, **1999**.

[4] See for example: a) G. G. Hammes, S. J. Benkovic, S. Hammes-Schiffer, *Biochemistry* **2011**, *50*, 10422–10430; b) A. J. Adamczyk, J. Cao, S. C. L. Kamerlin, A. Warshel, *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 14115–14120; c) J. Liu, R. Nussinov, *Proc. Natl. Acad. Sci. USA* **2008**, *105*, 901–906; d) M. Garcia-Viloca, J. Gao, M. Karplus, D. G. Truhlar, *Science* **2004**, *303*, 186–195; e) S. D. Schwartz, V. L. Schramm, *Nat. Chem. Biol.* **2009**, *5*, 551–558; f) Z. D. Nagel, J. P. Klinman, *Nat. Chem. Biol.* **2009**, *5*, 543–550; g) N. Boekelheide, R. Salomón-Ferrer, T. F. Miller III, *Proc. Natl. Acad. Sci. USA* **2011**, *108*, 16159–16163; h) S. Hay, N. S. Scrutton, *Nat. Chem.* **2012**, 4, 161–168.

[5] Reviews of rational design of enzymes using site-specific mutagenesis:[2b] a) J. Pleiss in *Enzyme Catalysis in Organic Synthesis*, *3rd ed.* (Eds.: K. Drauz, H. Gröger, O. May), Wiley-VCH, Weinheim, **2012**, pp. 89 – 117; b) T. Ema, *Curr. Org. Chem.* **2004**, *8*, 1009 – 1025; c) K. Hult, P. Berglund, *Curr. Opin. Biotechnol.* **2003**, *14*, 395 – 400; d) L. G. Otten, F. Hollmann, I. W. C. E. Arends, *Trends Biotechnol.* **2010**, *28*, 46 – 54.

[6] Recent reviews of directed evolution:[2b,8] 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. Biomol. Struct.* **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* (Eds.: K. Drauz, H. Gröger, O. May), Wiley-VCH, Weinheim, **2012**, pp. 119 – 190.

[7] Various older and more recent versions and applications of saturation mutagenesis have been described:[2b,6] a) J. A. Wells, M. Vasser, D. P. Powers, *Gene* **1985**, *34*, 315 – 323; b) K. M. Derbyshire, J. J. Salvo, N. D. F. Grindley, *Gene* **1986**, *46*, 145 – 152; c) A. Wissmann, I. Meier, W. Hillen, *J. Mol. Biol.* **1988**, *202*, 397 – 406; d) R. Higuchi, B. Krummel, R. K. Saiki, *Nucleic Acids Res.* **1988**, *16*, 7351 – 7367; e) M. A. Vandeyar, M. P. Weiner, C. J. Hutton, C. A. Batt, *Gene* **1988**, *65*, 129 – 133; f) D. K. Dube, L. A. Loeb, *Biochemistry* **1989**, *28*, 5703 – 5707; g) G. Sarkar, S. S. Sommer, *BioTechniques* **1990**, *8*, 404 – 407; h) S. Climie, L. Ruiz-Perez, D. Gonzalez-Pacanowska, P. Prapunwattana, S.-W. Cho, R. Stroud, D. V. Santi, *J. Biol. Chem.* **1990**, *265*, 18776 – 18779; i) A. V. Teplyakov, J. M. van der Laan, A. A. Lammers, H. Kelders, K. H. Kalk, O. Misset, L. J. S. M. Mulleners, B. W. Dijkstra, *Protein Eng.* **1992**, *5*, 413 – 420; j) L. D. Graham, K. D. Haggett, P. A. Jennings, D. S. Le Brocque, R. G. Whittaker, P. A. Schober, *Biochemistry* **1993**, *32*, 6250 – 6258; k) M. S. Warren, S. J. Benkovic, *Protein Eng.* **1997**, *10*, 63 – 68; l) R. D. Kirsch, E. Joly, *Nucleic Acids Res.* **1998**, *26*, 1848 – 1850; m) K. Miyazaki, F. H. Arnold, *J. Mol. Evol.* **1999**, *49*, 716 – 720; n) S. V. Taylor, P. Kast, D. Hilvert, *Angew. Chem.* **2001**, *113*, 3408 – 3436; *Angew. Chem. Int. Ed.* **2001**, *40*, 3310 – 3335; o) N. M. Antikainen, P. J. Hergenrother, M. M. Harris, W. Corbett, S. F. Martin, *Biochemistry* **2003**, *42*, 1603 – 1610; p) C. N. Dominy, D. W. Andrews, *Methods in Molecular Biology*, *Vol. 235* (Eds.: N. Casali, A. Preston), Humana, Totowa, **2003**, pp. 209 – 223; q) R. Georgescu, G. Bandara, L. Sun, *Directed Evolution Library Creation*, *Vol. 231* (Eds.: F. H. Arnold, G. Georgiou), Humana, Totowa, **2003**, pp. 75 – 83; r) E. M. Gabor, D. B. Janssen, *Protein Eng. Des. Sel.* **2004**, *17*, 571 – 579; s) A. R. Schmitzer, F. Lépine, J. N. Pelletier, *Protein Eng. Des. Sel.* **2004**, *17*, 809 – 819; t) T. S. Wong, K. L. Tee, B. Hauer, U. Schwaneberg, *Nucleic Acids Res.* **2004**, *32*, e26; u) L. Zheng, U. Baumann, J.-L. Reymond, *Nucleic Acids Res.* **2004**, *32*, e115; v) A. Herman, D. S. Tawfik, *Protein Eng. Des. Sel.* **2007**, *20*, 219 – 226; w) W.-C. Tseng, J.-W. Lin, T.-Y. Wei, T.-Y. Fang, *Anal. Biochem.* **2008**, *375*, 376 – 378; x) A. Hidalgo, A. Schließmann, R. Molina, J. Hermoso, U. T. Bornscheuer, *Protein Eng. Des. Sel.* **2008**, *21*, 567 – 576; y) 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, M. T. Reetz, *Appl. Microbiol. Biotechnol.* **2008**, *81*, 387 – 397.

[8] a) Review of directed evolution emphasizing library quality rather than quantity, which leads to less screening:[21] S. Lutz, W. M. Patrick, *Curr. Opin. Biotechnol.* **2004**, *15*, 291 – 297; statistical approaches to assess oversampling in the screening process:[21,25a] b) A. E. Firth, W. M. Patrick, *Bioinformatics* **2005**, *21*, 3314 – 3315; c) A. E. Firth, W. M. Patrick, *Nucleic Acids Res.* **2008**, *36*, W281 – W285; d) M. Denault, J. N. Pelletier, *Protein Engineering Protocols*, *Vol. 352* (Eds.: K. M. Arndt, K. M. Müller), Humana, Totowa **2007**, pp. 127 – 154; e) Y. Nov, *Appl. Environ. Microbiol.* **2012**, *78*, 258 – 262; f) J. F. Chaparro-Riggers, K. M. Polizzi, A. S. Bommarius, *Biotechnol. J.* **2007**, *2*, 180 – 191.

[9] a) W. Cai, H. Hong, *Protein-Protein Interactions—Computational and Experimental Tools*, InTech, Rijeka, **2012**; b) D. A. Bonsor, E. J. Sundberg, *Biochemistry* **2011**, *50*, 2394 – 2402; c) "Protein-Protein Interactions": M. Werther, H. Seitz, *Advances in Biochemical Engineering/Biotechnology*, Springer, Berlin, **2010**; d) O. Keskin, B. Ma, K. Rogale, K. Gunasekaran, R. Nussinov, *Phys. Biol.* **2005**, *2*, S24 – S35; e) G. Pál, J.-L. K. Kouadio, D. R. Artis, A. A. Kossiakoff, S. S. Sidhu, *J. Biol. Chem.* **2006**, *281*, 22378 – 22385; f) S. Byrum, S. K. Smart, S. Larson, A. J. Tackett, *Methods Mol. Biol.* **2012**, *833*, 143 – 152; g) Z. Qiu, X. Wang, *J. Theor. Biol.* **2012**, *293*, 143 – 150; h) Q. C. Zhang, D. Petrey, L. Deng, L. Qiang, Y. Shi, C. A. Aye, B. Bisikirska, C. Lefebvre, D. Accili, T. Hunter, T. Maniatis, A. Califano, B. Honig, *Nature* **2012**, *490*, 556 – 560.

[10] G. Schreiber, A. R. Fersht, *J. Mol. Biol.* **1995**, *248*, 478 – 486.

[11] D. C. Carter, G. Winter, A. J. Wilkinson, A. R. Fersht, *Cell* **1984**, *38*, 835 – 840.

[12] a) J. A. Wells, D. B. Powers, R. R. Bott, T. P. Graycar, D. A. Estell, *Proc. Natl. Acad. Sci. USA* **1987**, *84*, 1219 – 1223; b) J. A. Wells, *Biochemistry* **1990**, *29*, 8509 – 8517.

[13] a) A. Horovitz, *Curr. Biol.* **1996**, *6*, R121 – R126; b) D. Shortle, A. K. Meeker, *Proteins Struct. Funct. Genet.* **1986**, *1*, 81 – 89; c) D. Shortle, *Quart. Rev. Biophys.* **1992**, *25*, 205 – 250; d) G. K. Ackers, F. R. Smith, *Annu. Rev. Biochem.* **1985**, *54*, 597 – 629; e) J. Laskowski, M. I. Kato, W. Ardelt, J. Cook, A. Denton, M. W. Empie, W. J. Kohr, S. J. Park, K. Parks, B. L. Schatzley, O. L. Schoenberger, M. Tashiro, G. Vichot, H. E. Whatley, A. Wieczorek, M. Wieczorek, *Biochemistry* **1987**, *26*, 202 – 221; f) H. C. M. Nelson, R. T. Sauer, *Cell* **1985**, *42*, 549 – 558.

[14] a) A. S. Mildvan, D. J. Weber, A. Kuliopulos, *Arch. Biochem. Biophys.* **1992**, *294*, 327 – 340; b) A. S. Mildvan, *Biochemistry* **2004**, *43*, 14517 – 14520.

[15] M. M. Skinner, T. C. Terwilliger, *Proc. Natl. Acad. Sci. USA* **1996**, *93*, 10753 – 10757.

[16] T. Aita, H. Uchiyama, T. Inaoka, M. Nakajima, T. Kokubo, Y. Husimi, *Biopolymers* **2000**, *54*, 64 – 79.

[17] a) M. Lehmann, C. Loch, A. Middendorf, D. Studer, S. F. Lassen, L. Pasamontes, A. P. G. M. van Loon, M. Wyss, *Protein Eng.* **2002**, *15*, 403 – 411; b) B. J. Sullivan, T. Nguyen, V. Durani, D. Mathur, S. Rojas, M. Thomas, T. Syu, T. J. Magliery, *J. Mol. Biol.* **2012**, *420*, 384 – 399; c) N. Declerck, M. Machius, P. Joyet, G. Wiegand, R. Huber, C. Gaillardin, *Biologia Bratislava* **2002**, *57*, 203 – 211; d) Y. Qvarnström, G. Swedberg, *Microbiology* **2000**, *146*, 1151 – 1156; e) J. D. Bloom, J. J. Silberg, C. O. Wilke, D. A. Drummond, C. Adami, F. H. Arnold, *Proc. Natl. Acad. Sci. USA* **2005**, *102*, 606 – 611.

[18] M. T. Reetz, J. D. Carballeira, J. J.-P. Peyralans, H. Höbenreich, A. Maichele, A. Vogel, *Chem. Eur. J.* **2006**, *12*, 6031 – 6038.

[19] L. Serrano, A. Horovitz, B. Avron, M. Bycroft, A. R. Fersht, *Biochemistry* **1990**, *29*, 9343 – 9352.
[20] a) Z. Huang, C. R. Wagner, S. J. Benkovic, *Biochemistry* **1994**, *33*, 11576 – 11585; b) C. R. Wagner, Z. Huang, S. F. Singleton, S. J. Benkovic, *Biochemistry* **1995**, *34*, 15671 – 15680.
[21] Review of the directed evolution of stereoselective enzymes with an emphasis on iterative saturation mutagenesis (ISM):[6q] M. T. Reetz, *Angew. Chem.* **2011**, *123*, 144 – 182; *Angew. Chem. Int. Ed.* **2011**, *50*, 138 – 174.
[22] First ISM study for enhancing stereoselectivity: a) M. T. Reetz, L.-W. Wang, M. Bocola, *Angew. Chem.* **2006**, *118*, 1258 – 1263; *Angew. Chem. Int. Ed.* **2006**, *45*, 1236 – 1241; Erratum: M. T. Reetz, L.-W. Wang, M. Bocola, *Angew. Chem.* **2006**, *118*, 2556; *Angew. Chem. Int. Ed.* **2006**, *45*, 2494; b) ISM for enhancing stability: M. T. Reetz, J. D. Carballeira, *Nat. Protoc.* **2007**, *2*, 891 – 903; c) ISM study for controlling regioselective P450-catalyzed oxidative hydroxylation: S. Kille, F. E. Zilly, J. P. Acevedo, M. T. Reetz, *Nat. Chem.* **2011**, *3*, 738 – 743.
[23] a) M. T. Reetz, A. Zonta, K. Schimossek, K. Liebeton, K.-E. Jaeger, *Angew. Chem.* **1997**, *109*, 2961 – 2963; *Angew. Chem. Int. Ed. Engl.* **1997**, *36*, 2830 – 2832; b) M. T. Reetz, S. Wilensek, D. Zha, K.-E. Jaeger, *Angew. Chem.* **2001**, *113*, 3701 – 3703; *Angew. Chem. Int. Ed.* **2001**, *40*, 3589 – 3591; c) M. T. Reetz, *Proc. Natl. Acad. Sci. USA* **2004**, *101*, 5716 – 5722.
[24] a) M. Bocola, N. Otte, K.-E. Jaeger, M. T. Reetz, W. Thiel, *ChemBioChem* **2004**, *5*, 214 – 223; b) M. T. Reetz, M. Puls, J. D. Carballeira, A. Vogel, K.-E. Jaeger, T. Eggert, W. Thiel, M. Bocola, N. Otte, *ChemBioChem* **2007**, *8*, 106 – 112.
[25] 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**, DOI: 10.1021/sb300037w; b) L. Tang, H. Gao, X. Zhu, X. Wang, M. Zhou, R. Jiang, *BioTechniques* **2012**, *52*, 149 – 158; c) Y. Gumulya, *Dissertation*, Ruhr-Universität Bochum, Bochum, **2010**; d) D. J. Bougioukou, S. Kille, A. Taglieber, M. T. Reetz, *Adv. Synth. Catal.* **2009**, *351*, 3287 – 3305.
[26] M. T. Reetz, S. Prasad, J. D. Carballeira, Y. Gumulya, M. Bocola, *J. Am. Chem. Soc.* **2010**, *132*, 9144 – 9152.
[27] A. G. Sandström, Y. Wikmark, K. Engström, J. Nyhlén, J.-E. Bäckvall, *Proc. Natl. Acad. Sci. USA* **2012**, *109*, 78 – 83.
[28] S. Bartsch, R. Kourist, U. T. Bornscheuer, *Angew. Chem.* **2008**, *120*, 1531 – 1534; *Angew. Chem. Int. Ed.* **2008**, *47*, 1508 – 1511.
[29] H. Zheng, M. T. Reetz, *J. Am. Chem. Soc.* **2010**, *132*, 15744 – 15751.
[30] A. Y. Istomin, M. M. Gromiha, O. K. Vorov, D. J. Jacobs, D. R. Livesay, *Proteins Struct. Funct. Bioinf.* **2007**, *70*, 915 – 924.
[31] D. M. Weinreich, N. F. Delaney, M. A. DePristo, D. L. Hartl, *Science* **2006**, *312*, 111 – 114.
[32] a) M. T. Reetz, J. Sanchis, *ChemBioChem* **2008**, *9*, 2260 – 2267; b) M. T. Reetz, D. Kahakeaw, J. Sanchis, *Mol. BioSyst.* **2009**, *5*, 115 – 122.
[33] a) W. Augustyniak, A. A. Brzezinska, T. Pijning, H. Wienk, R. Boelens, B. W. Dijkstra, M. T. Reetz, *Protein Sci.* **2012**, *21*, 487 – 497; b) M. T. Reetz, P. Soni, J. P. Acevedo, J. Sanchis, *Angew. Chem.* **2009**, *121*, 8418 – 8422; *Angew. Chem. Int. Ed.* **2009**, *48*, 8268 – 8272.
[34] Y. Gumulya, J. Sanchis, M. T. Reetz, *ChemBioChem* **2012**, *13*, 1060 – 1066.
[35] a) S. G. Peisajovich, D. S. Tawfik, *Nat. Methods* **2007**, *4*, 991 – 994; b) M. A. DePristo, *HFSP J.* **2007**, *1*, 94 – 98.
[36] M. Eigen, J. McCaskill, P. Schuster, *J. Phys. Chem.* **1988**, *92*, 6881 – 6891.
[37] a) S. Kurtovic, B. Mannervik, *Biochemistry* **2009**, *48*, 9330 – 9339; b) S. Kurtovic, A. Shokeer, B. Mannervik, *J. Mol. Biol.* **2008**, *382*, 136 – 153.
[38] a) M. S. Breen, C. Kemena, P. K. Vlasov, C. Notredame, F. A. Kondrashov, *Nature* **2012**, *490*, 535 – 538; b) G. P. Wagner, *Nature* **2012**, *490*, 493 – 494.
[39] De novo computational design of enzymes based on the Rosetta algorithm: a) L. Jiang, E. A. Althoff, F. R. Clemente, L. Doyle, D. Röthlisberger, A. Zanghellini, J. L. Gallaher, J. L. Betker, F. Tanaka, C. F. Barbas III, D. Hilvert, K. N. Houk, B. L. Stoddard, D. Baker, *Science* **2008**, *319*, 1387 – 1391; b) N. Koga, R. Tatsumi-Koga, G. Liu, R. Xiao, T. B. Acton, G. T. Montelione, D. Baker, *Nature* **2012**, *491*, 222 – 227.
[40] I would like to add that the traditional term "protein engineering" is an unfortunate misnomer. According to Webster's Dictionary, engineering is defined as "*the application of science and mathematics by which the properties of matter and the sources of energy in nature are made useful to people*". However, the vast majority of science in the field of today's protein engineering involves basic research. Not just in this Essay, but also in thousands of other articles, it has become abundantly clear that we are far from understanding how enzymes function in detail, and we still do not know how to predict mutations for improving catalytic behavior in a general and reliable way.

