Catalytic Versatility, Stability, and Evolution of the $(\beta\alpha)_8$-Barrel Enzyme Fold

Reinhard Sterner*,† and Birte Höcker‡

Institut für Biophysik und physikalische Biochemie, Universität Regensburg, Universitätsstrasse 31, D-93053 Regensburg, Germany, and Department of Biochemistry, Duke University Medical Center, Box 3711, Durham, North Carolina 27710

Received June 16, 2005

Contents

1. Introduction

Natural evolution has yielded enzymes that catalyze virtually all metabolic reactions under mild conditions with high specificities and enormous rate enhancements.[1] It has been a major goal of biochemistry for almost a century to understand the chemical and molecular principles that underlie this remarkable catalytic power and versatility.[2,3] During the past decades, the use of recombinant DNA technology has provided new insights into the structure−function relationship of enzymes. It allows us to produce large amounts of protein for X-ray crystallography or multidimensional NMR spectroscopy and to test the catalytic role of individual amino acids by site-directed mutagenesis. The recent advent of complete genome sequences and of bioinformatics tools to analyze them enhances the chance to discover novel enzymes and to gain insights into their natural evolution.[4,5] Moreover, new methods of directed laboratory evolution have the potential to tailor enzymes for their use in industrial processes.[6,7]

The $(\beta\alpha)_8$-barrel, which was found first in triosephosphate isomerase[8] and therefore is also known as TIM barrel, is the most common enzyme fold. The latest release (February 2005) of the SCOP database (http://scop.mrc-lmb.cam.ac.uk/scop/) lists 378 different $(\beta\alpha)_8$-barrels, and about 10% of all proteins with known three-dimensional structure contain at least one TIM barrel domain. Hence, $(\beta\alpha)_8$-barrels provide an excellent model to address the function, stability, and evolution of enzymes. This review begins with a general description of the topology of the $(\beta\alpha)_8$-barrel fold, its distribution in nature, and some of its most remarkable catalytic performances. We then discuss the mechanisms by which $(\beta\alpha)_8$-barrels can acquire high (thermo-)stability and what is known about the folding pathway from the unstructured polypeptide chain to the native structure. Attempts to design new $(\beta\alpha)_8$-barrels are presented, including experiments in which the evolution of $(\beta\alpha)_8$-barrels from $(\beta\alpha)_4$-half barrels was mimicked. Moreover, structural and mechanistic evidence are provided for the evolution of a large fraction of the known $(\beta\alpha)_8$-barrels from a common ancestor. Finally, directed laboratory evolution is described that generated $(\beta\alpha)_8$-barrel enzymes tailored with respect to turnover numbers, substrate spectra, and enantioselectivities. We focus mainly on work of the last 5 years, but earlier literature will be cited where relevant. Although we try to cover the field in a comprehensive manner, we refer the reader to some recent reviews that discuss specific aspects of the structure and evolution of $(\beta\alpha)_8$-barrel enzymes in more detail.[9−14]

2. Structure

The canonical $(\beta\alpha)_8$-barrel contains about 200 amino acids and is composed of eight units, each of which consists of a β-strand and an α-helix that are connected by a $\beta\alpha$-loop. The individual units are linked by $\alpha\beta$-loops (Figure 1a). The eight β-strands form a curved central parallel β-sheet, the barrel, which is surrounded by the α-helices (Figure 1b). The strands are tilted by approximately 36° with respect to the principal barrel axis leading to a staggered β-sheet with a shear number of 8, which denotes a shift of eight residues when moving around the barrel

* To whom correspondence should be addressed. Tel: +49-941-943 3015. Fax: +49-941-943 2813. E-mail: Reinhard.Sterner@biologie.uni-regensburg.de.
† Universität Regensburg.
‡ Duke University Medical Center.

10.1021/cr030191z CCC: $53.50
Published on Web 11/09/2005

Reinhard Sterner was born in 1961 in Eichstätt (Germany). He studied biology at the University of Munich, where he received his Ph.D. with Prof. Heinz Decker in 1991. After 6 years of postdoctoral work at the Biocenter of the University of Basel (Switzerland) with Prof. Kasper Kirschner, he completed his habilitation in biochemistry in 1997. He then moved to the University of Göttingen, where he was sponsored by a Heisenberg fellowship of the German Research Foundation. In 1999, he became associate professor at the University of Cologne, and since 2004, he has been full professor of biochemistry at the University of Regensburg (Germany). His main research interests are the evolution and the design of enzymes and the structural and functional peculiarities of enzymes from hyperthermophiles. He is the chairman of the priority program on "Directed Evolution To Optimize and Understand Molecular Biocatalysts", which is financed by the German Research Foundation.

Birte Höcker was born in 1974 in Bielefeld (Germany). She studied biology at the University of Göttingen and at Carleton University in Ottawa (Canada). She earned her Ph.D. in biochemistry from the University of Cologne under the direction of Prof. Reinhard Sterner in 2003. Currently, she is a postdoctoral research associate in the laboratory of Prof. Homme W. Hellinga at Duke University in Durham (NC) working on the computational and experimental design of enzymes. Her main research interests include the evolution and design of protein folds and functions.

by one turn.[15,16] Within the core of the barrel, side chains form three to four layers perpendicular to the principal barrel axis. Each layer consists of four alternating residues from the odd- (1, 3, 5, 7) or the even-numbered (2, 4, 6, 8) β-strands, which creates a 4-fold symmetry (Figure 1c). In all known ($\beta\alpha$)$_8$-barrel enzymes, the catalytically active residues are located at the C-terminal ends of the β-strands and in the $\beta\alpha$-loops ("catalytic face"), while residues maintaining the stability of the fold are found in the core and on the opposite end of the barrel, which includes the $\alpha\beta$-loops ("stability face"; Figure 1b).[17] This arrangement allows one to change catalytic activities without compromising stability, which is important for both natural and directed evolution (see sections 7 and 8).

The ($\beta\alpha$)$_8$-barrel fold does not appear to be dictated by details of sequence but rather by general characteristics such as the overall distribution of polar and nonpolar or charged residues. This finding is reflected in the overall low sequence conservation among ($\beta\alpha$)$_8$-barrels, which can make it difficult to detect evolutionary relatedness. Nevertheless, structure-based sequence alignments reveal clusters of similar residues at topologically equivalent positions that could direct and stabilize the common ($\beta\alpha$)$_8$-barrel folding pattern.[18] Besides their common characteristics, ($\beta\alpha$)$_8$-barrel structures are geometrically diverse.[11] The loss of hydrogen bonds between adjacent β-strands can lead to distortions of the barrel structure, as, for example, observed in a bacterial phosphatidylinositol-specific phospholipase C,[19] but deviations from the canonical topology are also known. For example, quinolinic acid phosphoribosyltransferase,[20] certain cellulases,[21,22] and flavoprotein 390[23] contain only seven β-strands. Enolases consist of a $\beta\beta\alpha\alpha(\beta\alpha)_6$-barrel domain, in which the direction of the first α-helix is reversed with respect to the other helices and the second β-strand runs antiparallel to the other strands, plus an $\alpha + \beta$ capping domain that is formed by the N- and C-termini of the protein.[24] The PLP-binding barrel starts with an α-helix and ends with a β-strand, which is probably the consequence of a circular permutation in the course of its evolution.[13] Furthermore, $\beta\alpha$-loops often contain insertions with additional secondary structure elements.

About two-thirds of the known TIM barrels are single-domain proteins.[11] However, many ($\beta\alpha$)$_8$-barrels assemble to homooligomers, and others are part of multienzyme complexes with complicated reaction mechanisms involving channeling of substrates. Examples are tryptophan synthase,[25,26] glutamate synthase,[27] and imidazoleglycerol phosphate synthase.[28,29]

3. Function

3.1. Range of Catalyzed Reactions

With few notable exceptions such as the storage proteins narbonin[30] and concavalin B,[31] all known ($\beta\alpha$)$_8$-barrels are enzymes. They cover five of the six classes as defined by the Enzyme Commission (EC), acting as oxidoreductases, transferases, lyases, hydrolases, and isomerases.[11,16] Overall, hydrolases (especially glycosidases) are the dominating class, comprising about half of the known TIM barrels (Figure 2a).[11] Of the more than 60 reactions performed by ($\beta\alpha$)$_8$-barrels, 85% are involved in energy metabolism, macromolecule metabolism, or small molecule metabolism (Figure 2b).[11]

A large fraction of the ($\beta\alpha$)$_8$-barrels require the presence of organic cofactors such as FMN, NADP, or PLP, and about half of them use divalent metal ions for catalysis.[12] Whereas the Mg^{2+}-binding ligands are conserved in the enolase superfamily (see section 7), the other ($\beta\alpha$)$_8$-barrels are specific but diverse in terms of position, type, and number of metal ions and their ligands. A very common active site motif is a phosphate binding moiety, as about two-thirds of the

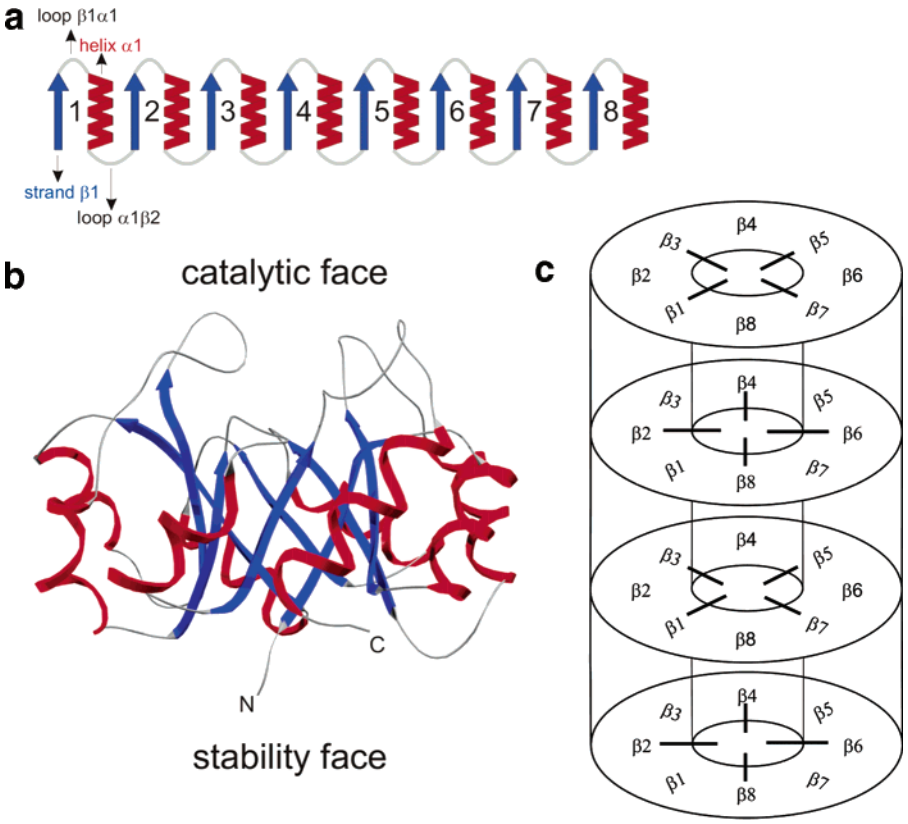

Figure 1. Schematic depiction of the $(\beta\alpha)_8$-barrel fold. (a) Representation of the eight $\beta\alpha$-units and the loops that connect the secondary structure elements within the units ($\alpha\beta$-loops) and between the units ($\beta\alpha$-loops). (b) Three-dimensional structure of a specific protein in ribbon representation, with the central eight-stranded parallel β-sheet (the "barrel") surrounded by the eight α-helices. The active site is formed by residues at the C-terminal ends of the β-strands and the $\beta\alpha$-loops ("catalytic face"). The remainder of the fold, including the opposite face of the barrel with the $\alpha\beta$-loops, is important for stability. The N- and C-terminal ends of the polypeptide chain are labeled. Reprinted with permission from ref 17. Copyright 2001 Elsevier. (c) Presentation of the four layers of the barrel from the catalytic face (top) to the stability face (bottom). Each layer is formed by the side chains of residues from the four odd- or even-numbered β-strands, which point to the interior of the barrel, leading to a 4-fold symmetric layer arrangement. The residues of a β-strand that are not involved in layer formation point toward the neighboring α-helices (not shown).

established $(\beta\alpha)_8$-barrels have substrates or cofactors with at least one phosphate group.[11]

It has been postulated that local microdipoles of individual α-helices[32] and global dipolar electrostatic field patterns along the barrel axis created by the combined contribution of α-helices[33] create a positive potential at the catalytic face of $(\beta\alpha)_8$-barrels (Figure 1b). The side chains appear to enhance the electrostatic field and focus it into a specific area near the active site, which could explain the preference of $(\beta\alpha)_8$-barrel enzymes for negatively charged metabolites.[33] It has also been shown that phylogenetic motifs (short sequence motifs that parallel the overall phylogeny of a protein family) correspond to active site electrostatic networks within different $(\beta\alpha)_8$-barrel families, indicating that conserved electrostatic properties are crucial for fine-tuning of catalysis.[34,35]

3.2. Selected Examples of Reaction Mechanisms

Of the many $(\beta\alpha)_8$-barrels for which catalytic mechanisms have been studied, three prominent examples will be discussed here in more detail. These are triosephosphate isomerase (TIM), orotidine-5′-monophosphate (OMP) decarboxylase, and ribulose-1,5 bisphosphate carboxylase (rubsico), which are among the most remarkable known enzymes with respect to importance, efficiency, and abundance.

Thirty years ago, TIM was the first enzyme for which the $(\beta\alpha)_8$-barrel structure was detected.[8] It catalyzes a central reaction within glycolysis, namely, the reversible interconversion of the ketose dihydroxyacetone phosphate (DHAP) and the aldose glyceraldehyde-3-phosphate (GAP). Formally, a proton is transferred from carbon 1 to carbon 2. Several sugar isomerases such as the xylose isomerase, which is also a $(\beta\alpha)_8$-barrel, catalyze a metal-dependent hydride transfer.[36] In contrast, the TIM mechanism is based on general acid−base catalysis, and it has been shown that four amino acid residues are important for the reaction. In the "classical" mechanism, an asparagine and a lysine residue in loop $\beta1\alpha1$ are engaged in substrate orientation, whereas a histidine residue in loop $\beta4\alpha4$ and a glutamate residue in loop $\beta6\alpha6$ are directly involved in proton transfer steps. The glutamate initiates the reaction by abstracting a proton from carbon 1 of the substrate DHAP. The negative charge that develops at the C-2 carbonyl group is stabilized by a proton that is donated by the histidine, which results in the formation of the enediol intermediate. The histidine then abstracts a pro-

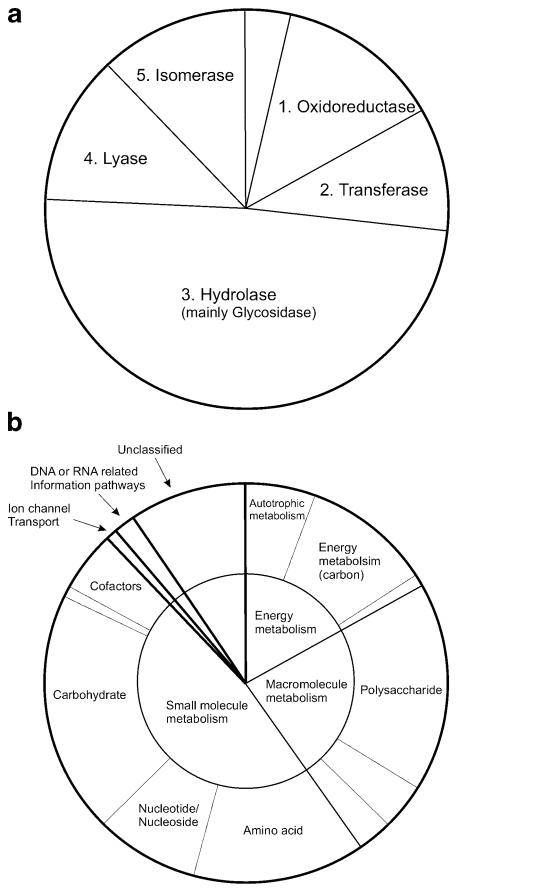

Figure 2. Distribution of chemical and biological functions of ($\beta\alpha$)$_8$-barrel enzymes represented by concentric pie charts. (a) Chemical classes. The numbers represent the first level of the EC hierarchy. The empty sector represents nonenzyme proteins. (b) Biological functions. Reprinted with permission from ref 11. Copyright 2002 Elsevier.

ton from the C-1 hydroxyl group, and the glutamate adds a proton to carbon 2, yielding GAP and regenerating the enzyme (Figure 3).[37]

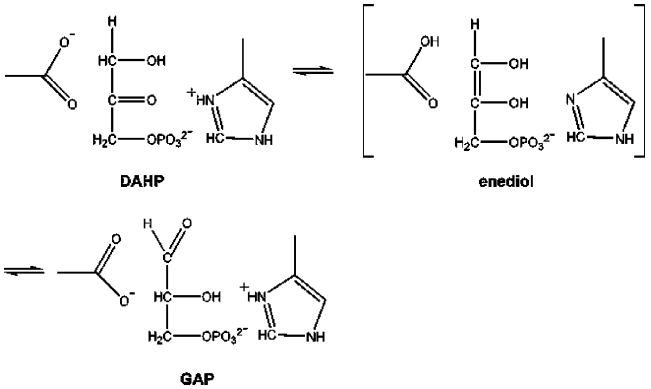

Figure 3. "Classical" reaction mechanism of TIM, shown in the physiological direction within glycolysis from left to right. The substrate DHAP, the enediol intermediate, and the product GAP are shown, together with the carboxylate and imidazole groups of the essential glutamate and histidine side chains. As the reaction proceeds, a proton is transferred from carbon 1 to carbon 2.

In the alternative "criss−cross" mechanism, all proton transfer steps are catalyzed by the glutamate while the developing negative charge on the substrate is stabilized by two different oxyanion holes, which are formed by the histidine and the lysine or the asparagine, respectively.[38,39] TIM is a remarkable catalyst for various reasons. In the (nonphysiological) direction from GAP to DHAP, it is a kinetically perfect enzyme, whose turnover is limited only by the diffusion-determined rate at which enzyme and substrate encounter each other.[37] Brownian dynamics simulations suggest that kinetic perfection is achieved through the steering of GAP by the local active site electrostatic potential, which has a similar value in various TIMs.[40] Moreover, TIM promotes catalysis and suppresses an undesired side reaction by a controlled movement of the protein chain during catalysis. In the absence of substrate, loop β6α6 interacts with the neighboring loop β5α5. Upon substrate binding, it positions itself as a lid on top of the active site, bringing the essential glutamate into position for catalysis. As a consequence, the productive protonation of the enediol intermediate is 10^5-fold favored over the nonproductive loss of its phosphate group by β-elimination, which in solution is >100-fold faster than the isomerization reaction.[41] Remarkably, the equivalent loops β6α6 of the ($\beta\alpha$)$_8$-barrels rubisco, inosine 5′-monophosphate dehydrogenase, the α-subunit of tryptophan synthase (TrpA), and phosphoribosyl anthranilate isomerase (TrpF) also undergo conformational transitions upon ligand binding.[42−45] The vast information on the TIM reaction mechanism was used in a computational design approach, by which the catalytically inert and structurally unrelated ribose-binding protein was turned into an enzyme that is highly active as triose phosphate isomerase.[46,47]

The OMP decarboxylase is engaged in nucleotide biosynthesis. It accelerates the interconversion of OMP to UMP by a factor of 10^{17}, which is higher than for any other known enzyme.[48] The two identical subunits of the OMP decarboxylase homodimer contact each other via their catalytic faces and are related by a 2-fold rotation axis.[49] At the C-terminal ends of strands β2 and β3, each of the two identical active sites harbors a charged network of two lysine and two aspartate residues that are conserved and crucial for catalysis.[50,51] Extensive investigations have focused on the enormous catalytic proficiency of OMP decarboxylase,[52] but several of the hitherto proposed mechanisms seem unlikely in the light of recently determined X-ray structures of the four different OMP decarboxylases in complex with various ligands.[53] In a plausible mechanism, the formation of a high-energy vinyl carbanion is prevented by a bimolecular electrophilic substitution in which carbon 6 is simultaneously decarboxylated and protonated, with the conserved lysine residue at the C-terminal end of strand β3 acting as the general acid. In this mechanism, the negatively charged aspartate in strand β3 interacts more favorably with the transition state than with the substrate (Figure 4).[49]

Molecular dynamics simulations have related this "ground state destabilization" to a distortion of the protein conformation in the presence of OMP, which would be relieved upon reaching the transition state.[53,54] This mechanism would explain how a

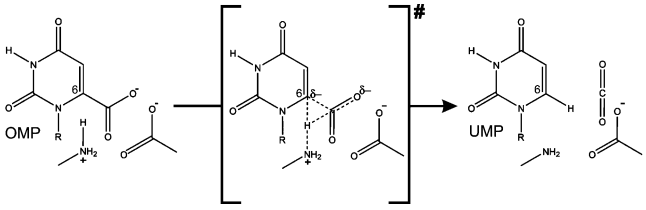

Figure 4. Plausible reaction mechanism of OMP decarboxylase. The substrate OMP ("ground state"), the proposed transition state, and the products UMP and CO_2 are shown, together with the amino and carboxylate groups of the essential lysine and aspartate side chains. As the decarboxylation reaction proceeds, a proton of the side chain of the lysine residue is transferred to carbon 6 of the pyrimidine. The ground state is destabilized by electrostatic repulsion between the substrate carboxylate and the aspartate carboxylate. This repulsion is reduced in the transition state by shifting the negative charge from the carboxylate to carbon 6. Reprinted with permission from ref 49. Copyright 2000 The National Academy of Sciences.

tyrosine and an arginine residue, which form hydrogen bonds with the phosphate moiety of OMP but are far apart from the disappearing carboxylate group, can contribute significantly to the rate enhancement of the enzyme.[55]

Rubisco catalyzes the fixation of carbon dioxide within the Calvin cycle of photosynthesis. This reaction occurs in the chloroplasts of plants and is rate limiting for hexose synthesis. Rubisco from plants is a complex enzyme that is composed of eight large and eight small subunits. Rubisco from photosynthetic bacteria forms a simple homodimer. The structure of its subunits is similar to the large subunit of the plant enzyme. The large subunits are composed of a β-sheet domain and a $(\beta\alpha)_8$-barrel domain that harbors the active site. The function of the small subunits, whose central part folds into a four-stranded antiparallel β-sheet covered on one side by two α-helices,[56] is unknown but seems to improve catalysis in some still undefined way.[57] In the first step of catalysis, the carbon 3 atom of ribulose-1,5-bisphosphate is deprotonated by the carboxylated lysine. The double bond between carbon 2 and carbon 3 of the resulting enediolate starts a nucleophilic attack onto the bound carbon dioxide, leading to an unstable six-carbon β-ketoacid intermediate that is hydrolyzed to yield two molecules of 3-phosphoglycerate. A Mg^{2+} ion is bound to the catalytic center through the side chains of a glutamate, an aspartate, and a carboxylated lysine residue. It is crucial for catalysis, because it stabilizes the enediol intermediate and orients it productively for its nucleophilic attack onto carbon dioxide (Figure 5).

X-ray crystallography has shown that binding of a transition state analogue induces the transition from an "open" to a "closed" conformation, due to the movement of loop $\beta6\alpha6$ toward the active site. This finding suggests that catalysis goes along with the closure of the active site, which is accompanied by a distinct shift in the position of the two phosphate binding sites, leading to a shortening of the interphosphate distance.[42] Given this complicated reaction mechanism, rubisco is a surprisingly inefficient catalyst that fixes only about three molecules of carbon dioxide per second. Moreover, it competes with the

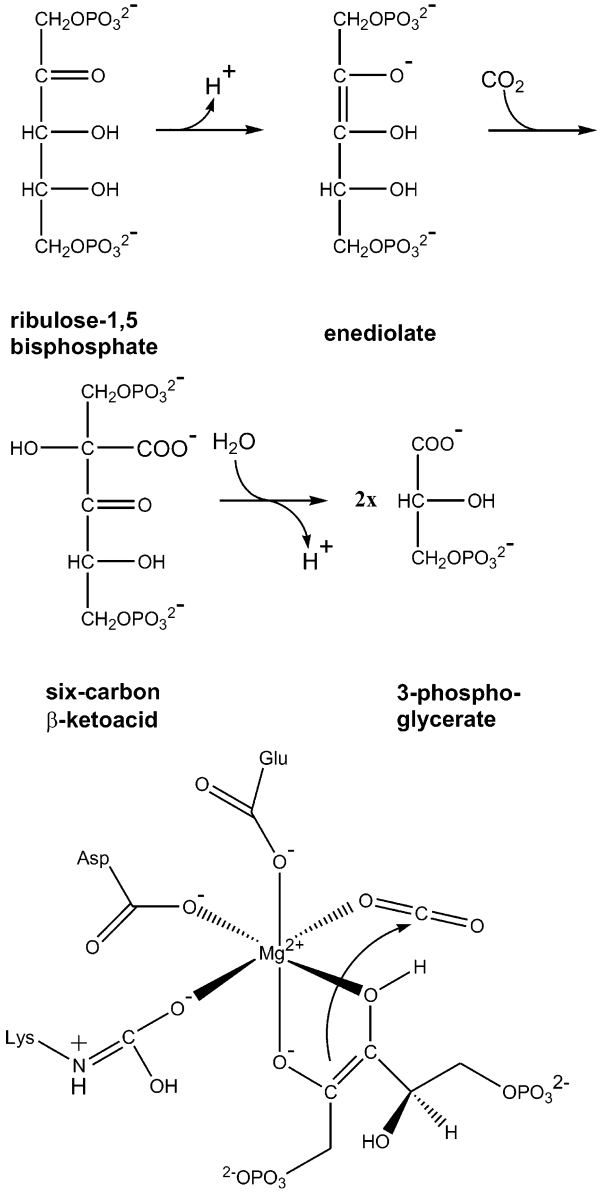

Figure 5. Reaction mechanism of rubisco. Upper two panels: The overall pathway from the substrate (ribulose-1,5-bisphosphate) via the intermediates (enediolate and six-carbon β-ketoacid) to the product (two molecules of 3-phosphoglycerate). Lower panel: The reactions take place on the magnesium ion, which is complexed to rubisco through a glutamate residue, an aspartate residue, and the carboxylated lysine. The nucleophilic attack of the enediolate onto carbon dioxide is shown.

nonproductive oxygenase side reaction that produces phosphoglycolate and 3-phosphoglycerate. The low turnover is compensated by the high concentration of rubisco in the chloroplast stroma (250 mg/mL, corresponding to an active site concentration of about 4 mM). Hence, rubisco represents about 50% of the soluble protein in chloroplasts, making it the most abundant enzyme on earth.

4. Stability

4.1. Structural Determinants of Stability

Numerous studies were performed to identify structural elements and individual amino acids that are

crucial for the stability of ($\beta\alpha$)$_8$-barrels. In a recently performed comprehensive search for potentially stabilizing amino acid residues, the structures of about 70 ($\beta\alpha$)$_8$-barrel proteins were analyzed.[58] Residues were assumed to be important for stability, depending on their involvement in long-range interactions, the hydrophobicity of their environment, and their conservation. Almost 1000 stabilizing residues were identified, which corresponds to about 5% of all inspected positions. Most of the bona fide stabilizing residues contribute to the eight-stranded β-sheet (Figure 1), with about one-half oriented toward the interior of the barrel and the other half oriented toward the surrounding α-helices. In contrast, only very few of the stabilizing residues were localized in α-helices.

In an experimental search for amino acid variations that allow one to maintain the stability of the ($\beta\alpha$)$_8$-barrel fold, about 180 out of the 250 sequence positions within yeast TIM were randomized using combinatorial mutagenesis.[59] Catalytically active TIM variants were isolated by functional selection. The analysis showed that a subset of seven amino acids (Phe, Val, Leu, Ala, Lys, Glu, and Gln) is a sufficiently large reservoir to replace about 80% of the randomized positions without losing the ability to form a stable and catalytically active TIM. This result suggests that a large fraction of the residues in α-helices or $\beta\alpha$-loops are highly mutable and therefore are not crucial for stability. In contrast, the disruption of a buried salt bridge and the exchange of amino acids in the core of the β-barrel resulted in inactive TIM variants, demonstrating that the optimal packing of the central β-sheet is essential for the formation of a stable ($\beta\alpha$)$_8$-barrel fold.

The application of knowledge-based potentials, which contain information on the forces and energies in proteins from high-resolution X-ray structures, suggests that αβ-loops are more important for the stability of the fold than $\beta\alpha$-loops.[60] This finding that suggests a division of labor between the two faces of the barrel (Figure 1) is supported by protein engineering studies. For example, an artificial TrpF protein with an internal duplication of the fifth $\beta\alpha$-module was enzymatically active, which proves that it has a wild-type ($\beta\alpha$)$_8$-core.[61] The duplicated segment was found to be inserted between strand β5 and helix α5, suggesting that the neighboring loops α4β5 and α5β6 are important for the stability of the structure. Along the same lines, a circularly permuted TrpF variant in which loop β6α6 was distorted by the insertion of the newly generated N- and C-termini was somewhat more stable and unfolded more cooperatively than a variant in which the new termini were inserted in the loop α6β7.[62] However, the analysis of active complexes between N- and C-terminal fragments of TrpA indicates that the role of $\beta\alpha$-loops for stability has to be judged individually.[63] Gene libraries comprising mixtures of randomly digested N- and C-terminal fragments were constructed, and clones expressing enzymatic activity were isolated by functional selection. Some of the catalytically active noncovalent complexes exhibited either cleavage, duplication, or fragment overlaps in loop α5β6. Moreover, a noncovalent complex between the four N- and the four C-terminal ($\beta\alpha$)-elements of imidazoleglycerol phosphate synthase (HisF) was stable and catalytically active, although it lacked an intact loop α4β5,[64] confirming that not all αβ-loops are equally important for a stable ($\beta\alpha$)$_8$-scaffold.

4.2. Lessons from Thermostable Representatives

Important insights into the structural basis of the stability of ($\beta\alpha$)$_8$-barrels were obtained by comparing the X-ray structures of orthologous enzymes from mesophiles and hyperthermophiles. Well-studied examples are TrpF and indoleglycerol phosphate synthase (TrpC) from tryptophan biosynthesis, as well as TIM. In most mesophiles, TrpF either forms a monomer or constitutes the C-terminal domain in a bifunctional TrpC−TrpF fusion protein.[62,65] In contrast, TrpF from *Thermotoga maritima* is an extremely thermostable homodimer, in which the contact interface between the two subunits is formed by two long symmetry-related αβ-loops that protrude reciprocally into cavities of the other β-barrel.[66,67] The intentional weakening of these interactions by shortening of the αβ-loops and by introducing electrostatic repulsion at the 2-fold symmetry axis resulted in a monomeric TrpF variant with wild-type catalytic activity but drastically reduced thermostability.[68] This result supports the postulated crucial role of αβ-loops for the integrity of ($\beta\alpha$)$_8$-barrels[61] and shows that an increased oligomerization is one possible strategy that allows proteins from hyperthermophiles to fold and function at temperatures around the boiling point of water.[69,70] Analogously, TIM forms a homodimer in eucarya, bacteria, and mesophilic archaea but a homotetramer in hyperthermophilic archaea.[71−73] Interestingly, TIM from *Thermoproteus tenax*, which forms an equilibrium between inactive dimers and active tetramers, is less stable than TIM from *Pyrococcus woesei*, which exclusively forms tetramers.[73−75] A different stabilization strategy is used by the TrpC proteins from *Sulfolobus solfataricus* and *T. maritima*. Both thermostable TrpC homologues are monomers but contain many more salt bridges (17) than the thermolabile TrpC from *Escherichia coli* (10), which forms the N-terminal domain of the bifunctional, monomeric TrpC−TrpF fusion protein.[65,76,77] However, only three out of the 17 salt bridges are topologically conserved between the two thermostable homologues, the major difference being the preference for intrahelical salt bridges in TrpC from *T. maritima* and for interhelical salt bridges in TrpC from *S. solfataricus* (Figure 6).

Obviously, the stabilizing effect of salt bridges depends on the sum of their individual contributions rather than on their location. The disruption by site-directed mutagenesis of a salt bridge that fixes the N-terminus to the core of the ($\beta\alpha$)$_8$-barrel of TrpC from *T. maritima* was less destabilizing than the disruption of a second one, which serves as a clamp between helices α1 and α8.[78] This finding demonstrates that individual salt bridges contribute differently to the thermostability of ($\beta\alpha$)$_8$-barrels.

The systematic analysis of X-ray structures and comprehensive site-directed mutagenesis studies have

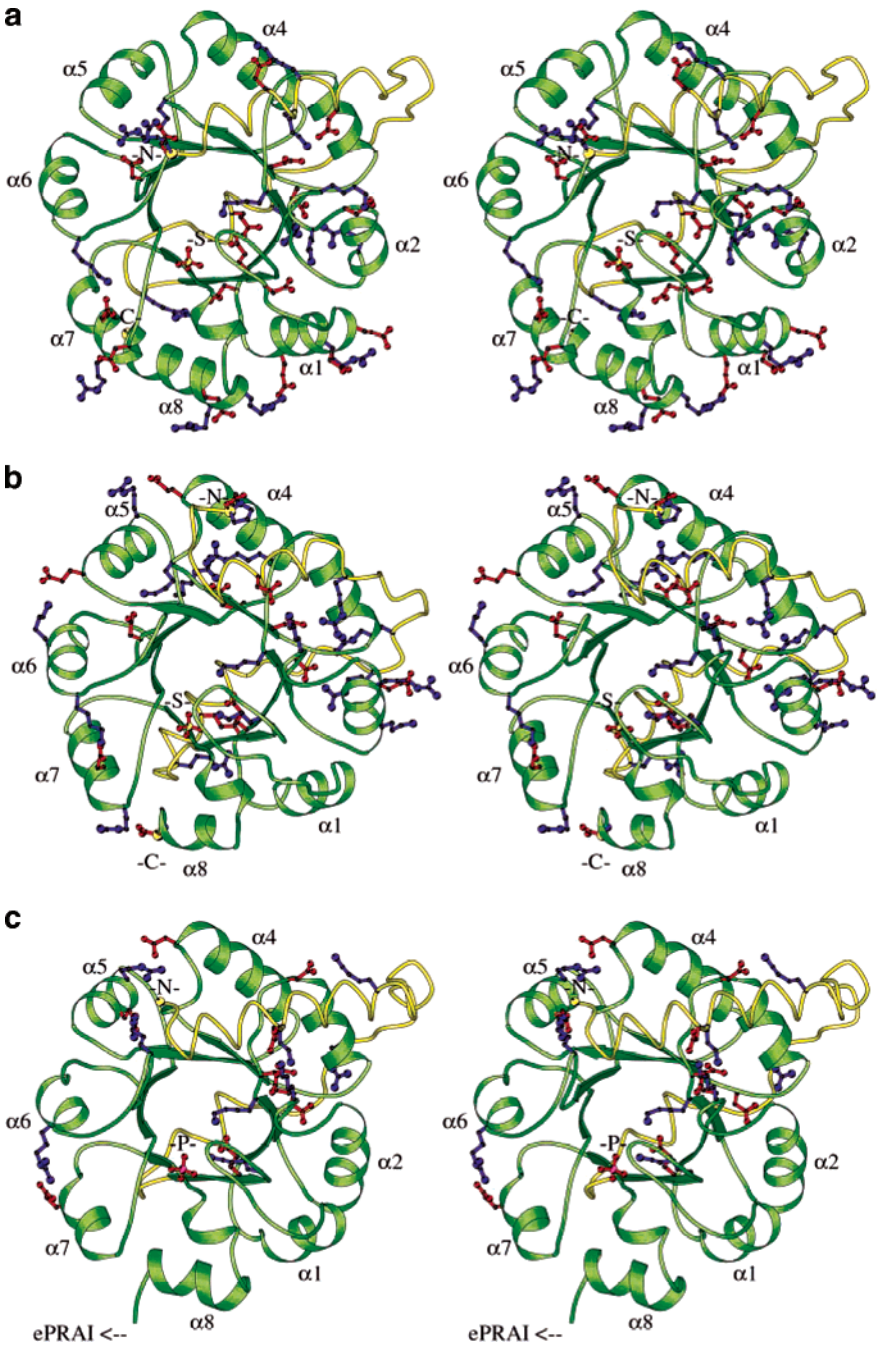

Figure 6. Distribution of salt bridges in the TrpC proteins from (a) *T. maritima*, (b) *S. solfataricus*, and (c) *E. coli*. The $(\beta\alpha)_8$-barrel fold is drawn in green, and the N-terminal extension preceding strand $\beta1$ is drawn in yellow. Negatively charged side chains are drawn as ball-and-stick models and colored red, and positively charged ones are colored blue. Both TrpC from *T. maritima* and *S. solfataricus* have more salt bridges than TrpC from *E. coli*. Reprinted with permission from ref 77. Copyright 2002 The American Society for Biochemistry and Molecular Biology.

shown that besides higher oligomerization states and additional salt bridges, proteins from hyperthermophiles can be intrinsically stabilized by an increased number of side chain−side chain hydrogen bonds, a decreased fraction of thermolabile residues such as glutamine and asparagine, an improved packing of the hydrophobic core, disulfide bonds, and the optimization of the global electrostatic potential.[69,70,79−82] This knowledge was used to engineer stabilized $(\beta\alpha)_8$-barrel proteins. In most known TIMs, the side chain of a glutamate residue from loop $\alpha3\beta4$ is completely buried within the dimer interface and involved in a conserved intersubunit hydrogen-bonding network. In TIM from *Leishmania mexicana*, this glutamine residue is replaced by a glutamate; therefore, the hydrogen-bonding network is distorted to some extent. Establishing this network in *L. mexicana* TIM by exchanging the glutamate for a glutamine residue resulted in an increase of the melting temperature from 57 to 83 °C, with practically unchanged catalytic activity.[83] In an attempt to stabilize the labile TrpC domain of the TrpC−TrpF fusion protein from *E. coli*, a variant linking residues 3 and 189 via a disulfide bridge was generated, which fastens the N-terminal extension of the protein to the core of the $(\beta\alpha)_8$-barrel. The thermal inactivation of the variant with a closed

disulfide was 65-fold slower than that of the reference dithiol form but only 13-fold slower than that of the parental protein. The decreased reactivity of parental buried cysteines with Ellman's reagent indicates that the stabilized variant with the closed disulfide is rigidified,[84] which might be responsible for its turnover number being decreased 4-fold as compared to the reduced variant and 9-fold as compared to native TrpC.

5. Folding

($\beta\alpha$)$_8$-barrel enzymes have similar structures but diverse amino acid sequences. They are therefore well-suited for testing whether folding mechanisms are dictated by the native topology or whether conserved amino acids are essential for the formation of crucial folding intermediates.[85,86] Fragmentation studies have suggested the existence of distinct folding intermediates of ($\beta\alpha$)$_8$-barrel enzymes. For example, TrpF from yeast forms a stable fragment consisting of the N-terminal ($\beta\alpha$)$_{1-6}$ units, which associates with the unstructured C-terminal ($\beta\alpha$)$_{7-8}$ units to form a functional complex.[87] These findings suggest a 6 + 2-folding mechanism, where the prefolded ($\beta\alpha$)$_{1-6}$ module serves as a template for the association and folding of the unstructured ($\beta\alpha$)$_{7-8}$ module. This model was supported by equilibrium and kinetic studies on the unfolding of the isolated TrpF domain from *E. coli*.[88] However, in another study, out of three different combinations of TrpF fragments [($\beta\alpha$)$_{1-2}$ + ($\beta\alpha$)$_{3-8}$; ($\beta\alpha$)$_{1-4}$ + ($\beta\alpha$)$_{5-8}$; and ($\beta\alpha$)$_{1-6}$ + ($\beta\alpha$)$_{7-8}$)] coexpressed in vivo, only the noncovalent complex of ($\beta\alpha$)$_{1-4}$ with ($\beta\alpha$)$_{5-8}$ yielded a functional enzyme that was able to complement a *trp*F deficiency strain.[89] Moreover, upon coexpression in vivo or joint refolding in vitro, the separately produced ($\beta\alpha$)$_{1-4}$ and ($\beta\alpha$)$_{5-8}$ units of HisF from *T. maritima* assembled to a catalytically fully active noncovalent complex, which is in favor of a 4 + 4-folding mechanism.[64] Along the same lines, the separately produced N- and C-terminal halves of chicken TIM assembled to a functional noncovalent complex when refolded together from guanidinium chloride.[90] The 4 + 4-folding mechanism is also supported by folding and unfolding studies with rabbit muscle TIM, which were performed by a combination of amide hydrogen exchange and mass spectrometry.[91] Whereas no unfolding intermediates were detected in either kinetic or equilibrium experiments, a highly populated refolding intermediate could be identified. This intermediate is composed of the folded C-terminal half of the sequence and the unfolded N-terminal half. Taken together, these findings indicate that TIM is composed of two half-barrel domains, which are structured sequentially on a 4 + 4-folding pathway to yield folded monomers, which then dimerize to native TIM. The unfolding of TIM was also studied by a native state exchange analysis using misincorporation proton-alkyl exchange (MPAX).[92] MPAX utilizes translational misincorporation of cysteine residues to generate an ensemble of proteins with single cysteine residue substitutions, which are distributed throughout the structure.[93] When exposed to solvent, the cysteines are alkylated by iodoacetamide, which prevents cleavage of the protein backbone by 2-nitro-5-thiocyanobenzoic acid. Thus, the loss of cleavage products over time reflects the rate of alkylation, which is a measure for fractional unfolding of the region around the incorporated cysteine. This technique identified three structurally distinct unfolding intermediates of TIM and supports a 3 + 3 + 2-folding pathway. The various models for the folding of ($\beta\alpha$)$_8$-barrels are compared in Figure 7.

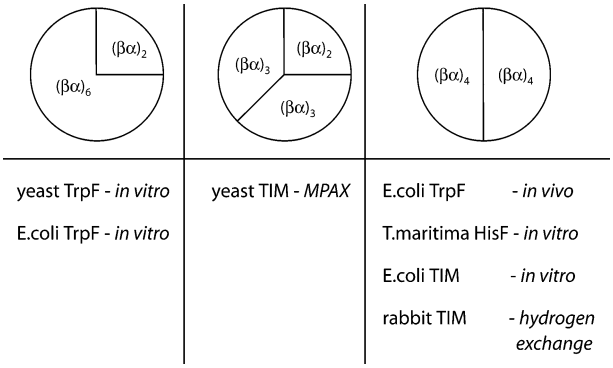

Figure 7. Models describing the folding of ($\beta\alpha$)$_8$-barrels. From left to right: 6 + 2, 3 + 3 + 2, and 4 + 4 model. Top row: pie charts depicting ($\beta\alpha$)-units that fold in a concerted way; bottom row: examples for the various models. See the text for details.

It is not clear whether and how the proposed 6 + 2-, 4 + 4-, and 3 + 3 + 2-folding models can be fully reconciled to generate a general model for the folding of TIM barrel proteins. Indeed, comprehensive investigations of the equilibrium and kinetic folding of TrpA from *E. coli* suggest that the folding process is more complex than described by these models. According to those studies, folding of TrpA involves four parallel pathways with on- and off-pathway sets of intermediates. An obligate set of late on-pathway equilibrium and kinetic intermediates comprises the folded N-terminal ($\beta\alpha$)$_{1-3}\beta_4$ fragment, which seems to direct the folding of the remaining C-terminal sequences.[94−96] Much of the complexity of the folding of TrpA is due to slow cis/trans prolyl isomerization reactions, which control the interconversions between the parallel pathways. Along these lines, the specific replacement of several proline residues leads to a simplified pathway in which refolding of more than 90% of the unfolded protein was no longer limited by proline isomerization.[97] Equilibrium and kinetic simulations of TrpA folding confirmed the existence of experimentally determined on- and off-pathway intermediates, although they were unable to detect parallel folding pathways resulting from prolyl isomerisation.[98] A parallel pathway mechanism controlled by proline isomerization has also been found for fructose-1,6-bisphosphate aldolase from *Staphylococcus aureus*.[99] In a comprehensive study, the folding of TrpC from *S. solfataricus* was analyzed and compared to TrpA.[86] These two evolutionarily related ($\beta\alpha$)$_8$-barrels catalyze successive steps in tryptophan biosynthesis, but their overall sequence identities are insignificant.[100] In TrpA, the fastest detectable folding event leads to the off-pathway intermediate,

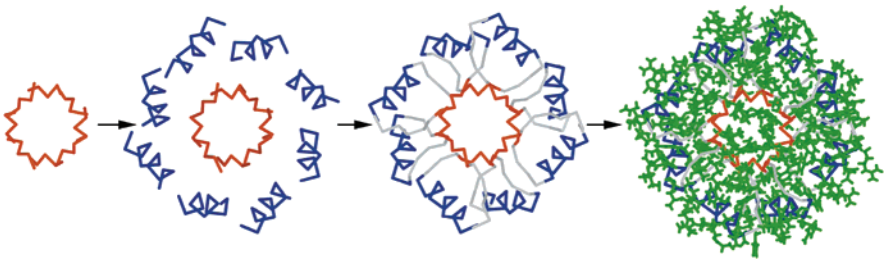

Figure 8. De novo design of a $(\beta\alpha)_8$-barrel. Representation (from left to right) of the β-barrel (red) and the surrounding α-helices (blue) with idealized geometric parameters, which were connected by short structural motifs (grey). An automated selection algorithm was used to find from different side chain rotamer libraries the amino acid sequence that was best compatible with the defined main chain scaffold (green). Reprinted with permission from ref 108. Copyright 2003 Elsevier.

which has to unfold before the obligate on-pathway intermediate can form. In contrast, the burst phase folding reaction of TrpC leads directly to its stable on-pathway intermediate. Moreover, the folding of TrpC is not complicated by cis/trans prolyl isomerization reactions but follows a simple sequential mechanism. Nevertheless, the stable folding intermediates of both TrpC and TrpA represent similar well-defined thermodynamic states that contain about 50% of the secondary structure and 25% of the buried surface area of the respective native proteins.[86] These findings suggest that topology dictates the population of similar folding intermediates in $(\beta\alpha)_8$-barrel proteins, whose exact structures, however, are determined by the specific amino acid sequences.[96]

6. Design

The regular order of the $(\beta\alpha)$-units and the overall symmetry of the fold have prompted attempts to design new $(\beta\alpha)_8$-barrels. Here, the reshuffling of individual $(\beta\alpha)$-units and the de novo design of a $(\beta\alpha)_8$-barrel from first principles will be discussed. Moreover, experimental evidence is presented suggesting that the $(\beta\alpha)_8$-barrel fold evolved by the fusion, mixing, and matching of $(\beta\alpha)_4$-half barrels.

6.1. Redesign and De Novo Design

Variants of TIM from yeast were constructed, where the native order of $(\beta\alpha)$-units (1−2−3−4−5−6−7−8) was changed as follows: full-length proteins with the orders (2−4−6−8−1−3−5−7) and (1−3−5−7−2−4−6−8) as well as a half-barrel (2−4−6−8) and a quarter-barrel (1−3) variant were generated.[101] It was found that the (2−4−6−8−1−3−5−7) variant is almost as stable as native TIM and is also primarily dimeric as is the wild type, whereas the other variants form higher order oligomers. Moreover, (2−4−6−8−1−3−5−7) possesses a comparable amount of secondary structure as native TIM and shields its tryptophan residues equally well from solvent. Although it remains to be shown whether (2−4−6−8−1−3−5−7) actually forms a regular $(\beta\alpha)_8$-barrel, the results indicate that the fold can tolerate profound alterations of strand−strand interactions responsible for the creation of the central β-barrel and the geometry of nonpolar side chains inside the hydrophobic core.

The ultimate goal of enzyme design is to generate de novo a polypeptide chain that folds into a predetermined three-dimensional structure with a given function. Geometrical characteristics of $(\beta\alpha)_8$-barrels have been parametrized in detail establishing the theoretical basis for such an approach.[102] However, early attempts for the de novo design of $(\beta\alpha)_8$-barrels generated sequences that at best are molten globules. These contained a considerable amount of secondary structure but did not fold into well-defined three-dimensional structures.[103−107] Important progress has been achieved in a recent two-step approach, which generated an idealized $(\beta\alpha)_8$-barrel protein of 216 residues.[108] First, 10 different geometric parameters were used to design an idealized artificial backbone, which contained an eight-stranded β-sheet with 4-fold symmetry surrounded by eight α-helices. The $(\beta\alpha)$-units were connected by short conserved $\alpha\beta$-turns.[109] In the second step, an optimal amino acid sequence fitting the idealized backbone was chosen with an automated selection algorithm that systematically searches side chain rotamer libraries (Figure 8).[110]

The designed protein was produced in *E. coli* by expression of a synthetic gene, purified, and characterized. Far-UV CD spectroscopy detected α-helical structure, and the presence of a pronounced near-UV signal indicated the immobilization of aromatic side chains. The chemical and thermal unfolding transitions were cooperative and yielded a high free energy of stabilization (ΔG_{H_2O} = 35 kJ/mol) and a high thermal stability (T_m = 65 °C). Moreover, ^{1}H NMR provided evidence for specific interactions between side chains. Together, these data suggest that the designed protein adopts a well-defined three-dimensional structure in solution. X-ray crystallography or multidimensional NMR are now required to show whether the artificial protein is a regular $(\beta\alpha)_8$-barrel.

6.2. Mimicking the Evolution of the Fold

The $(\beta\alpha)_8$-barrel enzymes N′-[(5′-phosphoribosyl)-formimino]-5-aminoimidazole-4-carboxamide-ribonucleotide isomerase (HisA) and HisF catalyze successive steps in histidine biosynthesis. The amino acid sequences and X-ray structures of HisA and HisF from *T. maritima* show an internal 2-fold symmetry.[111,112] The pairs of N-terminal halves (designated HisA-N and HisF-N) that consist of the first four $\beta\alpha$-units and the pairs of C-terminal halves (designated HisA-C and HisF-C) that consist of the last four $\beta\alpha$-units show sequence identities between 16 and 26%, and RMS deviations of their main chain nonhydrogen atoms between only 1.4 and 2.1 Å,

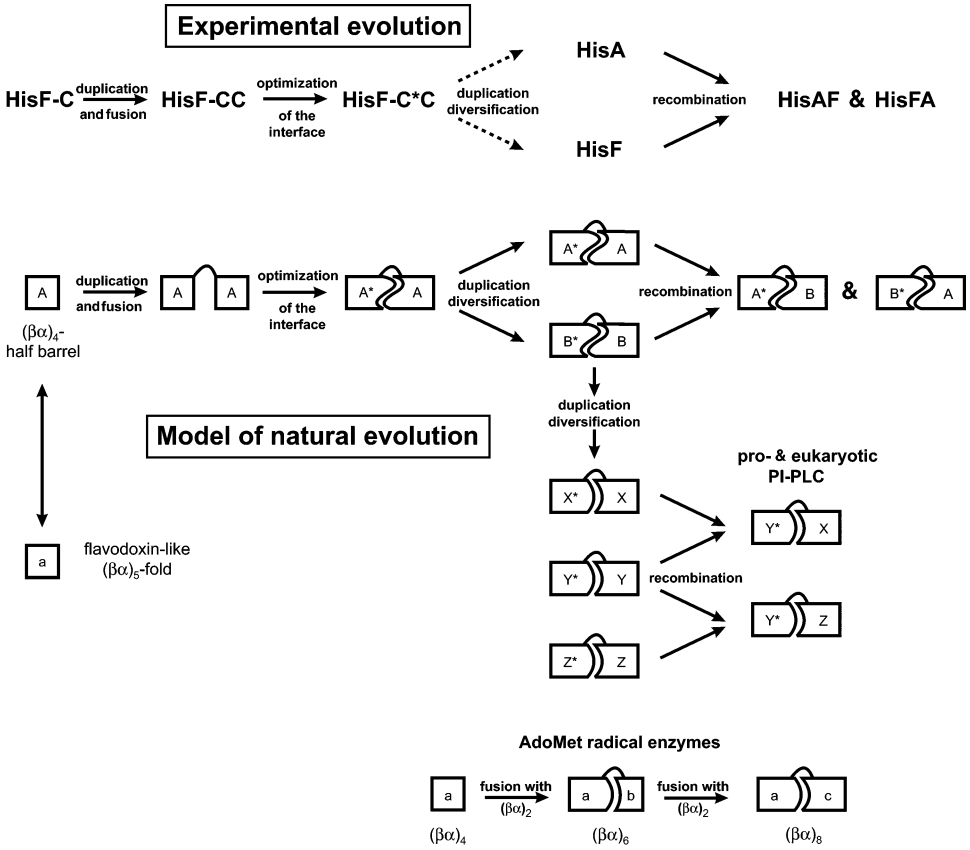

Figure 9. Evolving ($\beta\alpha$)$_8$-barrels by duplicating and fusing, and mixing and matching ($\beta\alpha$)$_4$-half-barrels. Model of the natural evolution of ($\beta\alpha$)$_8$-barrel enzymes from half-barrels (lower panel), and its experimental reconstruction (upper panel). The primordial ($\beta\alpha$)$_4$-half-barrel was mimicked by HisF-C. The duplication of its gene and the subsequent fusion of the daughter genes yielded the gene for HisF-CC, and the subsequent optimization of the interface between the two identical halves resulted in HisF-C*C. HisF-C*C mimicks an ancestral ($\beta\alpha$)$_8$-barrel from which HisA and HisF might have evolved by further gene duplication and diversification events. Recombination of ($\beta\alpha$)$_4$-half-barrels, mimicked by HisA-N, HisF-N, HisA-C, and HisF-C, leads to new ($\beta\alpha$)$_8$-barrels, mimicked by HisAF and HisFA. Through further steps of duplication and diversification, a repertoire of less symmetrical ($\beta\alpha$)$_8$-barrel enzymes evolved, which was extended by recombination. An example is provided by the pro- and eukaryotic phosphoinositide-specific phospholipases C (PI-PLC).[115] A comprehensive database search suggests an evolutionary linkage between ($\beta\alpha$)$_4$-half barrels and the ($\beta\alpha$)$_5$-flavodoxin-like fold family.[116] The analysis of sequences and structures of S-adenosyl-L-methionine (AdoMet) radical proteins suggests that the fold evolved from a ($\beta\alpha$)$_4$-half barrel to a ($\beta\alpha$)$_6$-three-quarter barrel to a ($\beta\alpha$)$_8$-full barrel structure.[117] Reprinted with permission from ref 113. Copyright 2004 The National Academy of Sciences.

respectively. When produced separately, HisF-N and HisF-C are predominantly homodimeric proteins with native secondary and tertiary structures but without measurable catalytic activity. When coexpressed in vivo or refolded together in vitro, however, the two half-barrels assemble to a catalytically fully active HisF-NC heterodimer. These results suggest an evolutionary scenario according to which a primordial gene coding for a ($\beta\alpha$)$_4$-half barrel (as a subunit of a homodimeric enzyme) was duplicated and fused in tandem to yield a monomeric, ancestral ($\beta\alpha$)$_8$-barrel, from which HisF, HisA, and other ($\beta\alpha$)$_8$-barrel enzymes may have evolved by a series of further gene duplication and diversification events.[64] This evolutionary strategy was mimicked by generating in vitro ($\beta\alpha$)$_8$-barrels from the ($\beta\alpha$)$_4$-half barrel HisF-C (Figure 9).[113] To this end, the gene for HisF-C was duplicated and fused in tandem to yield HisF-CC, which is more stable than HisF-C. In the next step, a salt bridge cluster that is present in the central barrel of native HisF was reconstructed within HisF-CC by introducing two amino acid exchanges, which yielded the monomeric and compact ($\beta\alpha$)$_8$-barrel HisF-C*C. Furthermore, in an attempt to construct ($\beta\alpha$)$_8$-barrels from different half-barrels, the N- and C-terminal halves of HisA and HisF were fused crosswise to yield the chimeric proteins HisAF and HisFA. Whereas HisFA contains native secondary structure elements but adopts ill-defined association states, HisAF is a stable and monomeric protein. These findings show that new ($\beta\alpha$)$_8$-barrels can form by mixing and matching existing ($\beta\alpha$)$_4$-half barrels,[114] which suggests a previously undescribed dimension for the evolution of enzymatic activities (Figure 9).[113]

When the Protein Data Bank was searched with DALI (http://www.ebi.ac.uk/dali/) for structures similar to HisF-N and HisF-C, several members of the flavodoxin-like fold family were identified.[116] Four of the five $\beta\alpha$-units of the flavodoxin-like fold are topologically equivalent to the four $\beta\alpha$-units of HisF-N and HisF-C while the fifth element corresponds to an additional small two-stranded β-sheet that is located in the first $\beta\alpha$-loop of each half-barrel (loops β1α1 and β5α5). These findings suggest that ($\beta\alpha$)$_4$-half barrels are independently evolving units, a

conclusion that is further supported by an inspection of the structures of the phosphoinositide-specific phospholipases C (PI-PLC) and S-adenosyl-L-methionine (AdoMet) radical proteins (Figure 9).[115,117] The structures of the C-terminal halves of eukaryotic and prokaryotic PI-PLCs are unrelated, whereas their N-terminal $(\beta\alpha)_4$-half barrels, which contain all catalytically essential residues, superimpose with an RMS deviation of only 1.85 Å for 104 equivalent Cα-atoms.[115] The members of the AdoMet radical protein superfamily have three putative architectures, which comprise four, six, and eight βα-units. The N-terminal half or three-quarter barrels contain the elements for radical generation, whereas the remainder of the protein provides the determinants for the binding of the different substrates that range in size from 10 atoms to 608 residue proteins. These findings led to the conclusion that AdoMet radical proteins may have evolved from a $(\beta\alpha)_4$-half barrel to a $(\beta\alpha)_6$-three-quarter barrel to a $(\beta\alpha)_8$-full barrel structure.[117] This scenario would imply that $(\beta\alpha)_2$-units might constitute the smallest evolving entity, which is in accordance with the observation that many $(\beta\alpha)_8$-barrels contain a conserved Gly-X-Asp motif in loops α1β2, α3β4, α5β6, and α7β8 but not in loops α2β3, α4β5, and α6β7.[11]

7. Natural Evolution

7.1. Global Comparisons of Sequences and Structures

The conserved location of the active site at the C-terminal end of the central β-barrel within all known $(\beta\alpha)_8$-barrel enzymes ("catalytic face", Figure 1) as well as similarities of three-dimensional structures and amino acid sequences suggests that a large fraction of the $(\beta\alpha)_8$-barrel enzymes have divergently evolved from a common ancestor by gene duplication and diversification.[118−120] Along the same lines, the recent releases of SCOP (http://scop.mrc-lmb.cam.ac.uk/scop/) and CATH (http://www.biochem.ucl.ac.uk/bsm/cath/) list 28 superfamilies and superfamily-like H-level families, respectively, of $(\beta\alpha)_8$-barrels, whose members presumably have a common evolutionary origin. To detect evolutionary links between these families, global comparisons of the known $(\beta\alpha)_8$-barrel sequences and X-ray structures were performed with an earlier version of CATH that contained 21 homologous families.[11] In the most conservative interpretation, which considers only significant sequence similarities, six of the 21 families were linked in one cluster and two families in another. In the most speculative interpretation, however, which considers all detected similarities of sequences, global structures, phosphate binding motifs, and location of catalytic and metal-binding residues, 17 families were linked in one large cluster. Most remarkable similarities were detected between $(\beta\alpha)_8$-barrels that use phosphate-containing substrates. Members from four different superfamilies contain a "standard phosphate binding motif", which is composed of residues that are located in strands β7 and β8 and in an additional helix α8′ that precedes the regular helix α8. These residues are conserved and superimpose with RMS deviations of only 1.2−2.2 Å. Members from seven other superfamilies seem to bind the phosphate moieties of their substrates at the same site or with similar geometry at another site, which suggests that all phosphate-binding $(\beta\alpha)_8$-barrels might be evolutionarily related. Comprehensive amino acid sequence comparisons of many $(\beta\alpha)_8$-barrels with phosphorylated substrates lead to a similar conclusion.[121]

7.2. Comparisons within Selected TIM-Barrel Families

Whereas global comparisons of structures and sequences can provide insights into higher order evolutionary networks, comparisons between closely related representatives are required to understand the mechanistic details of the divergent evolution of $(\beta\alpha)_8$-barrels. Well-studied examples are the enolase[10] and the amidohydrolase[122] superfamilies (members share a common reaction mechanism), the OMP decarboxylase suprafamily[13,123] (members share the same overall structure of the active site without similarities in reaction mechanism or substrate specificity), and the phosphate-binding $(\beta\alpha)_8$-barrel family of tryptophan and histidine biosynthesis (members share a similar substrate specificity).[120] These groups are discussed here in some detail, because they provide insightful examples of how existing catalytic machineries or ligand binding pockets can be used to establish new substrate specificities or reaction mechanisms.

The members of the enolase superfamily consist of a central $(\beta\alpha)_8$-barrel domain plus an $\alpha + \beta$ capping domain that is formed from the N- and C-termini of the protein.[13] Whereas the capping domain determines substrate specificity, the barrel domain contains the catalytic residues. All members of the enolase superfamily catalyze a common partial reaction, namely, the abstraction of the α-proton of the carboxylate substrate catalyzed by a general base. The resulting endiol intermediate is stabilized by a Mg^{2+} ion, which is liganded by three acidic residues that are located at the C-terminal ends of strands β3, β4, and β5. The intermediates are then further converted by unrelated mechanisms to different products in the various enzymes. The enolase superfamily is divided into three major subgroups [enolase, mandelate racemase (MR), and muconate lactonizing enzyme (MLE); Figure 10]. The division is based on the location and identity of the general base and acid catalysts (Table 1).

In the isofunctional enolase subgroup, which contains more than 600 identified members, a conserved lysine residue at the end of strand β6 acts as the general base, and a glutamate residue in loop β2α2 acts as the general acid. The hallmark of the MR subgroup is a His(β7)−Asp(β6) dyad, which acts as the general base. In addition, the active site contains a general acid at the C-terminal end of either strand β2, β3, or β5, depending on the identity and the stereochemical course of the reaction. The MR subgroup contains more than 400 identified members, about half of which have been functionally assigned as MR or one of the following acid sugar dehy-

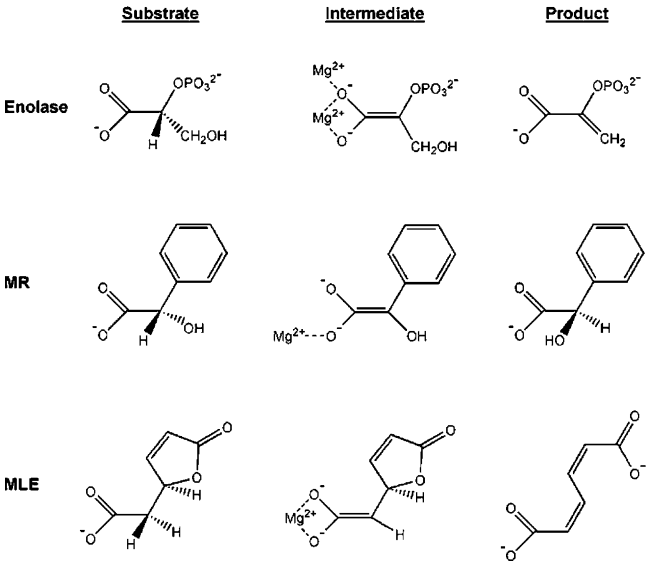

Figure 10. Members of the enolase superfamily share a common partial reaction: The α-proton of a carboxylic acid is abstracted, and the resulting enolic intermediate is stabilized by an active site Mg^{2+} ion. The reactions catalyzed by the naming representatives of the major subgroups enolase, mandelate racemase (MR), and muconate lactonizing enzyme (MLE) are shown.

dratases: D-glucarate/L-idarate dehydratase (GlucD), D-mannonate dehydratase (ManD), D-galactonate dehydratase (GalD), D-gluconate dehydratase (GlcD), and L-rhamnonate dehydratase (RhamD). The various acid sugar dehydratases catalyze syn- (RhamD) or anti- (GalD, GlcD) dehydration reactions, in which a proton is abstracted from carbon 2 and an OH leaving group departs from carbon 3. The syn dehydration requires a general base and a general acid at the same face of the active site, whereas the anti dehydration requires the general base and acid to be located on opposite faces. A specific equipment with catalytic residues makes GlucD functional promiscuous in that it can catalyze both syn and anti dehydration reactions, as the protons are abstracted from carbons of opposite configuration when D-glucarate or D-idarate are used as substrates.[10,124]

The members of the MLE subgroup contain at the ends of strands β2 and β6 lysine (or arginine) residues, one of which acts as the general base. This subgroup comprises more than 300 identified members, about half of which have been functionally assigned as MLE (cycloisomerization reaction), *o*-succinylbenzoate synthase (OSBS; dehydration reaction), or L-Ala-D/L-Glu epimerase (AE Epim; 1,1 proton transfer reaction) (Figure 11a). Despite this mechanistic diversity, the successful redesign of substrate specificities and the occurrence of natural catalytic promiscuities suggest that the members of the MLE subgroup are closely evolutionarily related. Along these lines, OSBS activity could be established on the AE Epim and MLE II scaffolds by rational design and directed evolution, respectively.[125] Space for the binding of the unnatural substrate was created by replacing a single acidic residue at the C-terminal end of strand β8 through a glycine (Figure 11b). The catalytic activities of the wild-type reactions were decreased but still measurable.

The replacement obviously enables the OSBS substrate to bind to the active sites of the MLE and AE Epim variants, which is prevented in the wild-type enzymes by steric and electrostatic repulsion. Moreover, OSBS is naturally promiscuous for catalyzing the racemization of N-acylamino acids,[126] because the corresponding substrates can be productively bound between the two catalytic lysine residues.[127] The efficiency of the promiscuous reaction was enhanced with an N-amino acid substrate that is more similar to the natural OSBS substrate.[128]

The members of the amidohydrolase (metallo-dependent hydrolase) superfamily catalyze the hydrolysis of a wide range of substrates of tetrahedral phosphorus and trigonal carbon centers.[9,122] The hallmark of this superfamily, which currently comprises more than 1000 identified members, is a mononuclear or binuclear metal center (Zn^{2+} or Ni^{2+} or Fe^{2+}) that activates the scissile bond of the substrate for cleavage and a hydrolytic water molecule for nucleophilic attack. There are 16 nonredundant X-ray structures of amidohydrolases available, which reveal numerous variations in the identity of the specific metal amino acid ligands (Figure 12).

Phosphotriesterase (PTE), dihydroorotase, urease, iso-aspartyl dipeptidase, and hydantoinase contain two Me^{2+} ions, which are complexed by four histidines and one aspartate ligand from strands β1, β5, β6, and β8.[122] The Me^{2+} ions of the PTE subgroup are bridged by a hydroxide ion and a carboxylated lysine residue from strand β4 (Figure 12a), which in the PTE homology protein has been replaced by glutamate (Figure 12b). PTE shows a marked stereoselectivity for pairs of chiral substrates, which seems to be dictated by the orientation of the side chains that form the substrate binding cavity and could be enhanced, relaxed, and inverted by the introduction of only a few amino acid exchanges.[129,130] The *N*-acetyl glucosamine-6-phosphate (Figure 12d) and the D-amino acid deacetylase (Figure 12e) contain all ligands found in the PTE subgroup except for the bridging carboxylate from strand β4, but the histidine residues from strand β1 are not engaged in binding of the single metal ion.[131,132] Instead, a cysteine residue in strand β2 acts as a ligand in the D-amino acid deacetylase. In the adenosine and

Table 1. Hallmarks of the Enolase Superfamily

subgroup	location of Mg^{2+}-binding Glu or Asp	reactions catalyzed	identity (location) of general base	identity (location) of general acid
enolase	$\beta3 + \beta4 + \beta5$	enolase	Lys (β6)	Glu (loop $\beta2\alpha2$)
MR	$\beta3 + \beta4 + \beta5$	MR, sugar dehydration	His(β7) − Asp(β6)	LysXN[a] (β2)[b]
MLE	$\beta3 + \beta4 + \beta5$	various (see Figure 11)	2 Lys[c] ($\beta2 + \beta6$)	

[a] N = Arg, Lys, Asp, Tyr, His. [b] In some cases, the general acid is located in strand β3 or β5. [c] The specific function of the individual lysine residues is unclear in most cases; occasionally, one lysine is replaced by an arginine.

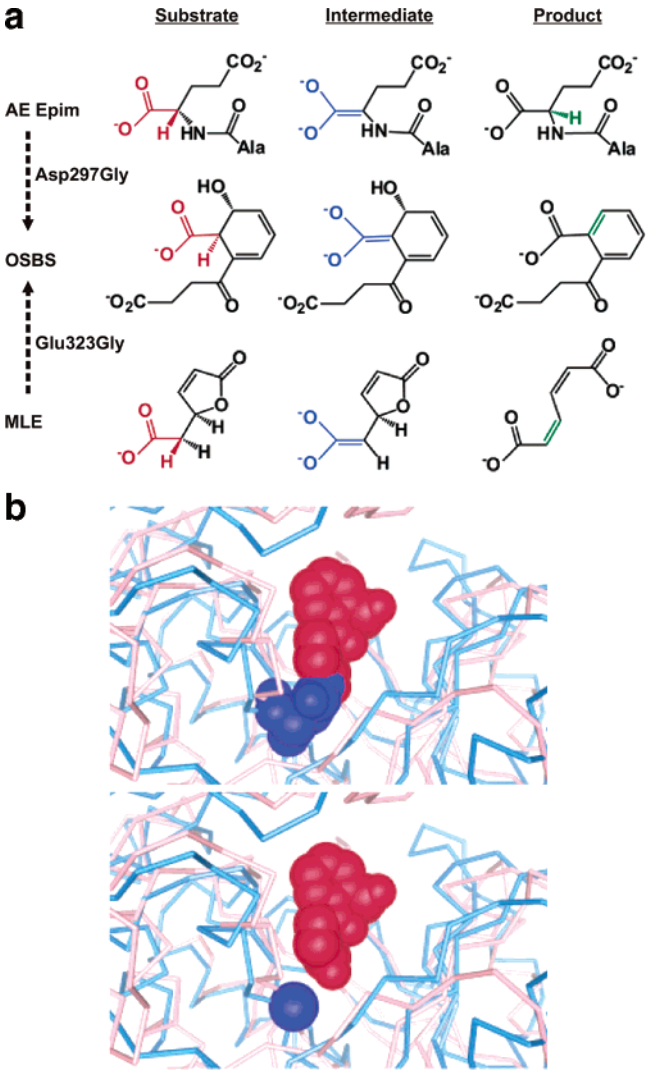

Figure 11. Extending catalytic activities within the MLE subgroup of the enolase superfamily. (a) Substrates, intermediates that are generated by the abstraction of the α-proton, and products of the AE Epim, OSBS, and MLE reactions. The indicated single amino acid exchanges are sufficient to establish OSBS activity on the AE Epim and the MLE scaffolds. (b) Upper panel: superposition showing the overlap of Asp297 (blue space-filling model of the side chain) in the active site of AE Epim (backbone represented by blue sticks) with the OSBS product (red space-filling model) in the active site of OSBS (backbone represented by red sticks); lower panel: predicted effect of the Asp297Gly exchange of AE Epim on removal of the structural overlap between amino acid side chain and bound ligand. Reprinted with permission from ref 125. Copyright 2003 American Chemical Society.

cytosine deaminases, no bridging ligand and only one divalent cation is present at the active site[133] (Figure 12c). As a consequence, a histidine from strand β6 is free to act as the general base for the activation of water. In the human renal dipeptidase (Figure 12f), the bridging lysine was replaced by glutamate on strand β3, and one of the histidines from strand β1 was replaced by an aspartate.[134]

The examples of the enolase and amidohydrolase superfamilies illustrate the plasticity of the active sites of related $(\beta\alpha)_8$-barrels, where the mechanisms of (partial) enzymatic reactions are retained while substrate specificities and stereoselectivities diverge in the course of evolution.

In other families of $(\beta\alpha)_8$-barrels, the substrate binding pockets are conserved, while catalytic mechanisms have diverged. A prominent example is the triad TrpF, TrpC, and TrpA, which belong to the ribulose-phosphate-binding barrel superfamily as defined by SCOP. These enzymes catalyze three successive steps in tryptophan biosynthesis, involving similar monophosphorylated substrates, and contain the above-mentioned conserved standard phosphate-binding motif at the ends of strands β7 and β8 and helix α8′. Therefore, it is plausible that they evolved from a common ancestor, although their reaction mechanisms are diverse.[100,120] The $(\beta\alpha)_8$-barrels HisA and HisF bind similar biphosphorylated substrates and, accordingly, contain an additional phosphate binding motif at the ends of strands β3 and β4.[112] Moreover, both enzymes use superimposable aspartate residues at the C-terminal ends of strands β1 and β5 for acid−base catalysis of their reactions.[45,135] The same aspartate residues are also essential for the promiscuous HisA activity of HisF, supporting divergent evolution from a common precursor.[111] Interestingly, TrpF and HisA catalyze similar Amadori rearrangements of aminoaldoses into the corresponding aminoketoses. A combination of random mutagenesis and selection in vivo revealed that the exchange of the essential aspartate in strand β5 by valine was sufficient to establish TrpF activity not only on the HisA but also on the HisF scaffold[136,137] (Figure 13). The newly established catalytic abilities resulted in an almost complete loss of the original activities.

Obviously, the removal of the negative charge enables the binding of the TrpF substrate phosphoribosyl anthranilate to the active sites of HisA and HisF, allowing its conversion into product by the aspartate residue of strand β1.[137] These findings suggest an evolutionary network, which links the histidine and tryptophan biosynthetic pathways. Along these lines, a novel isomerase (PriA) was recently discovered, which has a relaxed substrate specificity that allows it to catalyze both the HisA and the TrpF reaction (Figure 13).[138,139] PriA is a homologue of HisA, and further investigations are necessary to explain why it is able to bind and productively convert phosphoribosyl anthranilate without removal of the aspartate residue at the C-terminal end of strand β5.

The OMP decarboxylase (OMPDC) "suprafamily" denotes a group of $(\beta\alpha)_8$-barrel enzymes whose similarities in quaternary structure and active site architecture suggest a common evolutionary origin, although they catalyze mechanistically distinct reactions in different metabolic pathways.[140] The suprafamily comprises OMPDC, 3-keto-L-gulonate 6-phosphate decarboxylase (KGPDC), and D-arabino-hex-3-ulose 6-phosphate synthase (HPS). OMPDC uses a metal-independent reaction mechanism, which has been discussed in section 3 (Figure 4). KGPDC and HPS catalyze aldol condensation and decarboxylation reactions, respectively, using Mg^{2+} to stabilize an enediol intermediate that is derived from a ketose phosphate substrate (Figure 14).

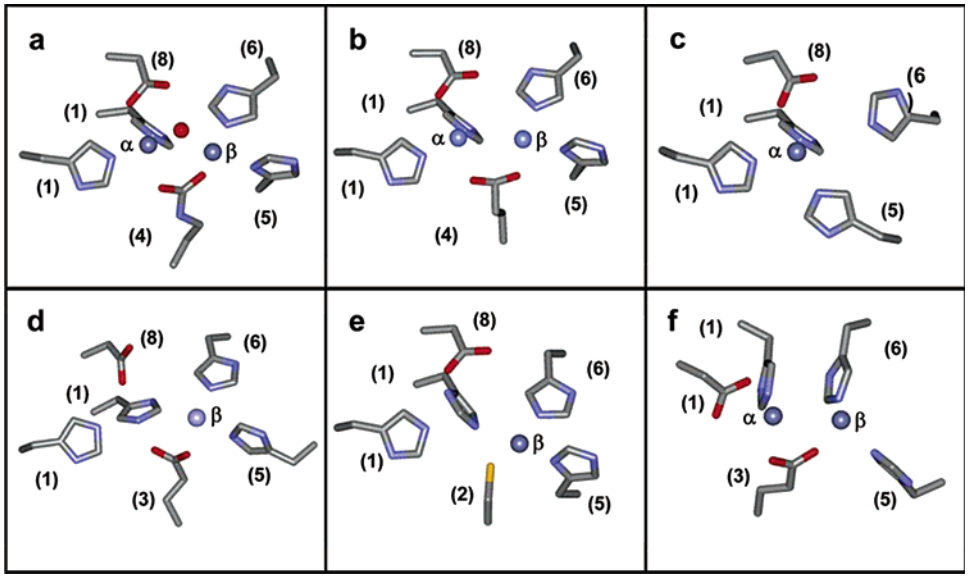

Figure 12. Comparison of the six subtypes for the binding of one or two divalent cations (α,β) to the active site of enzymes within the amidohydrolase superfamily. The subtypes are represented by (a) phosphotriesterase (red sphere is a hydroxide ion), (b) phosphotriesterase homology protein, (c) adenosine deaminase, (d) *N*-acetyl glucosamine-6-phosphate deacetylase, (e) D-amino acid deacetylase, and (f) renal dipeptidase. The numbers depict the β-strands to which the various metal-ligating side chains belong. Reprinted with permission from ref 122. Copyright 2005 American Chemical Society.

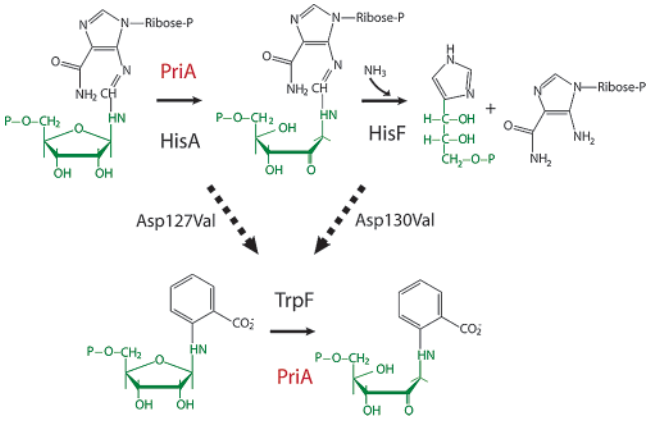

Figure 13. Changing substrate specificities between phosphate-binding ($\beta\alpha$)$_8$-barrel enzymes of histidine and tryptophan biosynthesis. *N*-[(5′-phosphoribosyl)-formimino]-5-aminoimidazole-4-carboxamide-ribonucleotide isomerase (HisA) and phosphoribosyl anthranilate isomerase (TrpF) catalyze Amadori rearrangements of phosphorylated aminoaldoses into the corresponding aminoketoses. Imidazoleglycerol phosphate synthase (HisF) catalyzes a ring closure reaction, along with the cleavage of a carbon−nitrogen double bond. The indicated single amino acid exchanges are sufficient to establish TrpF activity on the HisA and HisF scaffolds.[136,137] PriA has a naturally relaxed substrate specificity that allows it to catalyze both the HisA and the TrpF reaction.[138,139] The phosphorylated aminosugar moieties of the substrates and products are depicted in green. Reprinted with permission from ref 14. Copyright 2005 Elsevier.

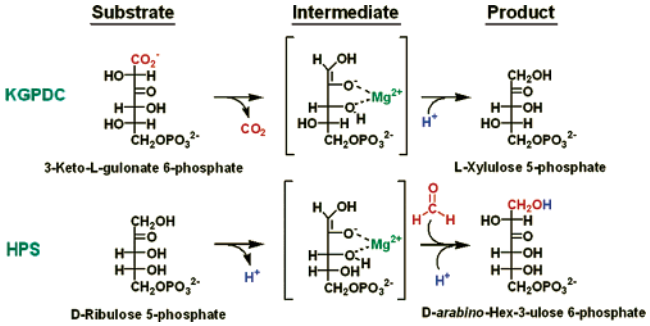

Figure 14. Reactions catalyzed by 3-keto-L-gulonate 6-phosphate decarboxylase (KGPDC) and D-arabino-hex-3-ulose 6-phosphate synthase (HPS). Reprinted with permission from ref 141. Copyright 2005 American Chemical Society.

OMPDC and KGPDC are homodimers with a conserved interface that carries two symmetry-related active sites. An Asp-X-Lys-X-X-Asp motif spans the homodimer so that the first Asp and the Lys are found in one active site of the homodimer and the second aspartate is found in the other active site. Moreover, the substrates of the two enzymes are found at nearly identical positions.[140] Although the X-ray structure of HPS is not yet available, model building suggests that it shares the conserved Asp-X-Lys-X-X-Asp motif at the active site. It was shown that KGPDC and HPS are both naturally promiscuous for the other reaction. Moreover, the HPS activity of KGPDC could be improved by the exchange of three to four conserved active site residues by those conserved in HPS.[141] The X-ray structures of wild-type and altered KGPDC with bound HPS substrate and KGPDC product provide a structural explanation for both the natural promiscuity and the enhanced HPS and the diminished KGPDC reaction catalyzed by the active site variants.[142]

8. Directed Evolution

The ($\beta\alpha$)$_8$-barrel is a stable fold that carries highly efficient and versatile active sites. As a consequence, it has been selected by nature to catalyze an impressive set of diverse reaction mechanisms. For these reasons, ($\beta\alpha$)$_8$-barrels provide an ideal scaffold for directed evolution. This technique mimicks natural evolution by combining random mutagenesis with screening or selection of enzyme variants with de-

sired new properties, which can be useful for applications in the chemical and pharmaceutical industry.[6,7] Some examples of directed evolution have already been discussed in the context of natural evolution in section 7. Here, experiments will be presented that led to increased turnover numbers and changes in substrate specificities and stereoselectivities of $(\beta\alpha)_8$-barrel enzymes.

8.1. Improving Catalytic Activities

A monomeric version of homodimeric TIM from *Trypanosoma brucei* was generated by shortening of loop $\beta3\alpha3$.[143] Its catalytic activity was improved by directed evolution,[144] following two different strategies. First, randomization of loop $\beta2\alpha2$ was performed to reinforce intrasubunit interaction, thus compensating for the loss of loop $\beta3\alpha3$; second, random mutagenesis of the entire gene was used. Interestingly, both strategies led to amino acid exchanges at neighboring positions, namely, the replacement of alanine 43 by proline and the substitution of threonine 44 by either alanine or serine. The characterization by steady state enzyme kinetics of the two activated variants revealed an 11-fold increase in k_{cat} and a 4-fold reduction in K_m. This improvement in catalytic efficiency is due to the stabilization of the new loop $\beta2\alpha2$.

The bacterial phosphotriesterase (PTE) degrades highly toxic organophosphates that are commonly used in insecticides and chemical warfare agents. Gene libraries of PTEs were generated by DNA shuffling[145,146] and screened by cell surface display for the improved hydrolysis of methyl parathion. The best variant displayed a 25-fold increased catalytic turnover over the wild-type enzyme.[147] It contains seven amino acid substitutions, one of which (His257Tyr) apparently reduces the size of the binding pocket, thus allowing better accommodation of the small substrate.

Enzymes from hyperthermophilic organisms are often barely active at room temperature, presumably because efficient catalysis is impeded by conformational "freezing".[70] To test the relationship between catalytic activity, conformational flexibility, and stability, a number of hyperthermophilic enzymes have been activated by directed evolution. TrpC variants from the hyperthermophile *S. solfataricus* with up to 4-fold increased turnover numbers at 37 °C were selected by a combination of random mutagenesis and complementation in vivo of an auxotrophic *E. coli* strain.[148] The analysis with steady state and transient kinetics showed that product release is the rate-limiting step in wild-type TrpC. In the activated variants, product release is accelerated, and the overall reaction is now limited by the chemical step (Figure 15).

Similarly, β-glucosidase CelB variants of the hyperthermophile *Pyrococcus furiosus* were isolated[149] that have up to 3-fold increased turnover rates for the hydrolysis of *p*-nitrophenyl-β-D-glucopyranoside. For both TrpC and CelB, the increased catalytic activities seem to be due to increased overall flexibilities, which were, however, achieved at the cost of reduced thermostabilities. The xylose isomerase

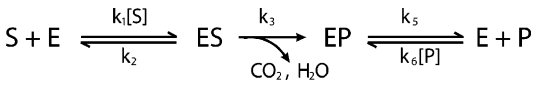

Figure 15. Minimal mechanism of the reaction catalyzed by the indoleglycerol phosphate synthase (TrpC = E). S is 1-(*o*-carboxyphenylamino)-1-deoxyribulose 5-phosphate. P is indoleglycerol phosphate. For the reaction catalyzed by native TrpC from *S. solfataricus*, the release of the product (described by k_5) is rate limiting. For the activated TrpC variants, product release is accelerated and the chemical step of the reaction (described by k_3) becomes rate determining. Reprinted with permission from ref 148. Copyright 2000 American Chemical Society.

(XylA) from *Thermus thermophilus*, whose larger N-terminal domain folds into a $(\beta\alpha)_8$-barrel harboring the active site, was made more active at low temperatures toward the substrates xylose and glucose. The acquired amino acid exchanges were neither affecting the substrate- nor the metal-binding site.[150] Along these lines, it was recently shown that residues remote from the active site can be important for catalysis and are detected more readily by directed evolution than by rational enzyme design or computation.[151] An already improved XylA variant from *Thermotoga neapolitana* was subjected to directed evolution in order to further increase its activity toward glucose.[152] As compared to the parent enzyme, the resulting variants were better catalysts at low temperatures and even more stable at elevated temperatures. This result shows that an increase in catalytic activity does not necessarily compromise stability, supporting earlier findings that the structural determinants of activity and thermostability are not necessarily linked.[70,153]

8.2. Changing Substrate Specificities and Stereoselectivities

Microbial hydantoinases, which belong to the superfamily of amidohydrolases[122] (see section 7.2), are widely employed for the commercial production of optically pure D- and L-amino acids. The substrate specificity of the thermostable and highly enantioselective D-hydantoinase from *Bacillus stearothermophilus* is biased toward unsubstituted hydantoin and is therefore not suited for the synthesis of commercially important nonnatural D-amino acids with aromatic side chains. Within D-hydantoinase, amino acids of substrate-determining loops were substituted by a combination of rational design and saturation random mutagenesis.[154] The turnover number for aromatic substrates gradually increased as the size of the amino acid side chain at the randomized position in the loop decreased, apparently due to removal of steric hindrance. Furthermore, saturation random mutagenesis of two additional residues led to a 10-fold increased turnover toward the aromatic hydantoin derivative hydroxy-phenyl-hydantoin, probably due to a reorganization of the active site. The enantioselectivity of L-hydantoinase from *Arthrobacter aurescens* is substrate-dependent, as D-5-methyl-thioethyl-hydantoin is preferred over its L-enantiomer. Single amino acid substitutions identified by directed evolution experiments altered the enantioselectivity of the enzyme in both directions.[155] The

observed effects were rationalized on the basis of the X-ray structure of the wild-type enzyme.[156] Whereas the exchange Val154Ala increases D-enantioselectivity by offering more space for the D-5-methyl-thiolethyl-hydantoin substrate side chain, the substitution Ile95Phe increases L-enantioselectivity by increasing the hydrophobicity close to the side chain of the L-enantiomer.

Aldolases are important ($\beta\alpha$)$_8$-barrel enzymes that catalyze with high enantioselectivity C−C bond forming and C−C breaking reactions but display only a limited substrate spectrum. By randomizing the active site residues of *E. coli* 2-keto-3-deoxy-6-phosphogluconate (KDPG) aldolase, a variant with an altered substrate profile was identified.[157] The change in substrate specificity was caused by a reorganization of the active site, in which the essential lysine was shifted from strand β6 (Lys133) to strand β7 (Lys161). The same enzyme was subjected to multiple rounds of random mutagenesis and DNA shuffling, and the resulting libraries were screened for variants that are more efficient in accepting nonphosphorylated substrates, as well as for variants with altered enantioselectivities regarding D- and L-sugars.[158] The effective substitutions were remote from the active site and did occur in nonconserved sequence regions, so that the mechanistic basis for the observed functional changes remains unclear.[158] The enantioselective cleavage reaction of *N*-acetylneuraminic aldolase from *E. coli* was altered by directed evolution to improve its ability to form L-sialic acid.[159] All amino acid substitutions were again found outside the active site of the ($\beta\alpha$)$_8$-barrel. The X-ray structure of the best second generation variant showed no significant difference to the wild-type enzyme, demonstrating that a small number of substitutions remote from the active site can significantly affect stereoselectivity.[159]

9. Conclusion

The ever increasing amount of information on amino acid sequences and three-dimensional structures of proteins allows us to compare enzyme folds and functions in a comprehensive way. Such comparisons will lead to a deeper understanding of the structural basis of enzyme catalysis, stability, folding, and evolution and certainly will identify new enzymes with novel catalytic activities. The ($\beta\alpha$)$_8$-barrel enzyme fold is particularly suitable to address these questions, because it occurs frequently, is catalytically versatile, and harbors active sites with high functional and structural plasticity. These properties also make ($\beta\alpha$)$_8$-barrels an ideal scaffold for the design of new catalytic activities, either by rational design or by directed laboratory evolution.

10. Acknowledgment

This work was supported by the Deutsche Forschungsgemeinschaft (Grants STE 891/4-1, 4-2). We thank Kasper Kirschner, Rainer Merkl, John Gerlt, and Frank Raushel for thoughtful comments on the manuscript and Marion Strieder for preparing the figures.

11. References

(1) Walsh, C. *Nature* **2001**, *409*, 226.
(2) Blow, D. *Struct. Fold Des.* **2000**, *8*, R77.
(3) Kraut, J. *Science* **1988**, *242*, 533.
(4) Morett, E.; Korbel, J. O.; Rajan, E.; Saab-Rincon, G.; Olvera, L.; Olvera, M.; Schmidt, S.; Snel, B.; Bork, P. *Nat. Biotechnol.* **2003**, *21*, 790.
(5) Pohl, N. L. *Curr. Opin. Chem. Biol.* **2005**, *9*, 76.
(6) Jaeger, K. E.; Eggert, T. *Curr. Opin. Biotechnol.* **2004**, *15*, 305.
(7) Williams, G. J.; Nelson, A. S.; Berry, A. *Cell Mol. Life Sci.* **2004**, *61*, 3034.
(8) Banner, D. W.; Bloomer, A. C.; Petsko, G. A.; Phillips, D. C.; Pogson, C. I.; Wilson, I. A.; Corran, P. H.; Furth, A. J.; Milman, J. D.; Offord, R. E.; Priddle, J. D.; Waley, S. G. *Nature* **1975**, *255*, 609.
(9) Gerlt, J. A.; Raushel, F. M. *Curr. Opin. Chem. Biol.* **2003**, *7*, 252.
(10) Gerlt, J. A.; Babbitt, P. C.; Rayment, I. *Arch. Biochem. Biophys.* **2005**, *433*, 59.
(11) Nagano, N.; Orengo, C. A.; Thornton, J. M. *J. Mol. Biol.* **2002**, *321*, 741.
(12) Vega, M. C.; Lorentzen, E.; Linden, A.; Wilmanns, M. *Curr. Opin. Chem. Biol.* **2003**, *7*, 694.
(13) Wise, E. L.; Rayment, I. *Acc. Chem. Res.* **2004**, *37*, 149.
(14) Höcker, B. *Biomol. Eng.* **2005**, *22*, 31.
(15) Wierenga, R. K. *FEBS Lett.* **2001**, *492*, 193.
(16) Pujadas, G.; Palau, J. *Biologica (Bratislava)* **1999**, *54*, 231.
(17) Höcker, B.; Jürgens, C.; Wilmanns, M.; Sterner, R. *Curr. Opin. Biotechnol.* **2001**, *12*, 376.
(18) Selvaraj, S.; Gromiha, M. M. *J. Protein Chem.* **1998**, *17*, 407.
(19) Moser, J.; Gerstel, B.; Meyer, J. E.; Chakraborty, T.; Wehland, J.; Heinz, D. W. *J. Mol. Biol.* **1997**, *273*, 269.
(20) Eads, J. C.; Ozturk, D.; Wexler, T. B.; Grubmeyer, C.; Sacchettini, J. C. *Structure* **1997**, *5*, 47.
(21) Rouvinen, J.; Bergfors, T.; Teeri, T.; Knowles, J. K.; Jones, T. A. *Science* **1990**, *249*, 380.
(22) Spezio, M.; Wilson, D. B.; Karplus, P. A. *Biochemistry* **1993**, *32*, 9906.
(23) Moore, S. A.; James, M. N.; O'Kane, D. J.; Lee, J. *EMBO J.* **1993**, *12*, 1767.
(24) Lebioda, L.; Stec, B.; Brewer, J. M. *J. Biol. Chem.* **1989**, *264*, 3685.
(25) Hyde, C. C.; Ahmed, S. A.; Padlan, E. A.; Miles, E. W.; Davies, D. R. *J. Biol. Chem.* **1988**, *263*, 17857.
(26) Miles, E. W. *Chem. Rec.* **2001**, *1*, 140.
(27) Binda, C.; Bossi, R. T.; Wakatsuki, S.; Arzt, S.; Coda, A.; Curti, B.; Vanoni, M. A.; Mattevi, A. *Struct. Fold Des.* **2000**, *8*, 1299.
(28) Chaudhuri, B. N.; Lange, S. C.; Myers, R. S.; Chittur, S. V.; Davisson, V. J.; Smith, J. L. *Structure (Camb)* **2001**, *9*, 987.
(29) Douangamath, A.; Walker, M.; Beismann-Driemeyer, S.; Vega-Fernandez, M. C.; Sterner, R.; Wilmanns, M. *Structure (Camb)* **2002**, *10*, 185.
(30) Hennig, M.; Pfeffer-Hennig, S.; Dauter, Z.; Wilson, K. S.; Schlesier, B.; Nong, V. H. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **1995**, *51*, 177.
(31) Hennig, M.; Jansonius, J. N.; Terwisscha van Scheltinga, A. C.; Dijkstra, B. W.; Schlesier, B. *J. Mol. Biol.* **1995**, *254*, 237.
(32) Hol, W. G.; van Duijnen, P. T.; Berendsen, H. J. *Nature* **1978**, *273*, 443.
(33) Raychaudhuri, S.; Younas, F.; Karplus, P. A.; Faerman, C. H.; Ripoll, D. R. *Protein Sci.* **1997**, *6*, 1849.
(34) La, D.; Sutch, B.; Livesay, D. R. *Proteins* **2005**, *58*, 309.
(35) Livesay, D. R.; La, D. *Protein Sci.* **2005**, *14*, 1158.
(36) Whitlow, M.; Howard, A. J.; Finzel, B. C.; Poulos, T. L.; Winborne, E.; Gilliland, G. L. *Proteins* **1991**, *9*, 153.
(37) Knowles, J. R. *Philos. Trans. R. Soc. London, Ser. B* **1991**, *332*, 115.
(38) Kursula, I.; Wierenga, R. K. *J. Biol. Chem.* **2003**, *278*, 9544.
(39) Guallar, V.; Jacobson, M.; McDermott, A.; Friesner, R. A. *J. Mol. Biol.* **2004**, *337*, 227.
(40) Wade, R. C.; Gabdoulline, R. R.; Luty, B. A. *Proteins* **1998**, *31*, 406.
(41) Xiang, J.; Jung, J. Y.; Sampson, N. S. *Biochemistry* **2004**, *43*, 11436.
(42) Duff, A. P.; Andrews, T. J.; Curmi, P. M. *J. Mol. Biol.* **2000**, *298*, 903.
(43) McMillan, F. M.; Cahoon, M.; White, A.; Hedstrom, L.; Petsko, G. A.; Ringe, D. *Biochemistry* **2000**, *39*, 4533.
(44) Brzovic, P. S.; Hyde, C. C.; Miles, E. W.; Dunn, M. F. *Biochemistry* **1993**, *32*, 10404.
(45) Henn-Sax, M.; Thoma, R.; Schmidt, S.; Hennig, M.; Kirschner, K.; Sterner, R. *Biochemistry* **2002**, *41*, 12032.
(46) Dwyer, M. A.; Looger, L. L.; Hellinga, H. W. *Science* **2004**, *304*, 1967.
(47) Sterner, R.; Schmid, F. X. *Science* **2004**, *304*, 1916.
(48) Radzicka, A.; Wolfenden, R. *Science* **1995**, *267*, 90.

(49) Appleby, T. C.; Kinsland, C.; Begley, T. P.; Ealick, S. E. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 2005.
(50) Traut, T. W.; Temple, B. R. *J. Biol. Chem.* **2000**, *275*, 28675.
(51) Miller, B. G.; Snider, M. J.; Wolfenden, R.; Short, S. A. *J. Biol. Chem.* **2001**, *276*, 15174.
(52) Miller, B. G.; Wolfenden, R. *Annu. Rev. Biochem.* **2002**, *71*, 847.
(53) Gao, J. *Curr. Opin. Struct. Biol.* **2003**, *13*, 184.
(54) Wu, N.; Mo, Y.; Gao, J.; Pai, E. F. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 2017.
(55) Miller, B. G.; Snider, M. J.; Short, S. A.; Wolfenden, R. *Biochemistry* **2000**, *39*, 8113.
(56) Mizohata, E.; Matsumura, H.; Okano, Y.; Kumei, M.; Takuma, H.; Onodera, J.; Kato, K.; Shibata, N.; Inoue, T.; Yokota, A.; Kai, Y. *J. Mol. Biol.* **2002**, *316*, 679.
(57) Andersson, I.; Taylor, T. C. *Arch. Biochem. Biophys.* **2003**, *414*, 130.
(58) Gromiha, M. M.; Pujadas, G.; Magyar, C.; Selvaraj, S.; Simon, I. *Proteins* **2004**, *55*, 316.
(59) Silverman, J. A.; Balakrishnan, R.; Harbury, P. B. *Proc. Natl. Acad. Sci. U.S.A.* **2001**, *98*, 3092.
(60) Wiederstein, M.; Sippl, M. J. *J. Mol. Biol.* **2005**, *345*, 1199.
(61) Urfer, R.; Kirschner, K. *Protein Sci*. **1992**, *1*, 31.
(62) Luger, K.; Hommel, U.; Herold, M.; Hofsteenge, J.; Kirschner, K. *Science* **1989**, *243*, 206.
(63) Hiraga, K.; Yamagishi, A.; Oshima, T. *J. Mol. Biol.* **2004**, *335*, 1093.
(64) Höcker, B.; Beismann-Driemeyer, S.; Hettwer, S.; Lustig, A.; Sterner, R. *Nat. Struct. Biol.* **2001**, *8*, 32.
(65) Wilmanns, M.; Priestle, J. P.; Niermann, T.; Jansonius, J. N. *J. Mol. Biol.* **1992**, *223*, 477.
(66) Sterner, R.; Kleemann, G. R.; Szadkowski, H.; Lustig, A.; Hennig, M.; Kirschner, K. *Protein Sci.* **1996**, *5*, 2000.
(67) Hennig, M.; Sterner, R.; Kirschner, K.; Jansonius, J. N. *Biochemistry* **1997**, *36*, 6009.
(68) Thoma, R.; Hennig, M.; Sterner, R.; Kirschner, K. *Struct. Fold Des.* **2000**, *8*, 265.
(69) Jaenicke, R.; Böhm, G. *Curr. Opin. Struct. Biol.* **1998**, *8*, 738.
(70) Sterner, R.; Liebl, W. *Crit. Rev. Biochem. Mol. Biol.* **2001**, *36*, 39.
(71) Kohlhoff, M.; Dahm, A.; Hensel, R. *FEBS Lett*. **1996**, *383*, 245.
(72) Schramm, A.; Kohlhoff, M.; Hensel, R. *Methods Enzymol*. **2001**, *331*, 62.
(73) Walden, H.; Bell, G. S.; Russell, R. J.; Siebers, B.; Hensel, R.; Taylor, G. L. *J. Mol. Biol.* **2001**, *306*, 745.
(74) Walden, H.; Taylor, G.; Lilie, H.; Knura, T.; Hensel, R. *Biochem. Soc. Trans.* **2004**, *32*, 305.
(75) Walden, H.; Taylor, G. L.; Lorentzen, E.; Pohl, E.; Lilie, H.; Schramm, A.; Knura, T.; Stubbe, K.; Tjaden, B.; Hensel, R. *J. Mol. Biol.* **2004**, *342*, 861.
(76) Hennig, M.; Darimont, B.; Sterner, R.; Kirschner, K.; Jansonius, J. N. *Structure* **1995**, *3*, 1295.
(77) Knöchel, T.; Pappenberger, A.; Jansonius, J. N.; Kirschner, K. *J. Biol. Chem.* **2002**, *277*, 8626.
(78) Merz, A.; Knöchel, T.; Jansonius, J. N.; Kirschner, K. *J. Mol. Biol.* **1999**, *288*, 753.
(79) Szilagyi, A.; Zavodszky, P. *Struct. Fold Des*. **2000**, *8*, 493.
(80) Vieille, C.; Zeikus, G. J. *Microbiol. Mol. Biol. Rev.* **2001**, *65*, 1.
(81) Perl, D.; Jacob, M.; Bano, M.; Stupak, M.; Antalik, M.; Schmid, F. X. *Biophys. Chem.* **2002**, *96*, 173.
(82) Jaenicke, R.; Sterner, R. *Angew Chem., Int. Ed.* **2003**, *42*, 140.
(83) Williams, J. C.; Zeelen, J. P.; Neubauer, G.; Vriend, G.; Backmann, J.; Michels, P. A.; Lambeir, A. M.; Wierenga, R. K. *Protein Eng.* **1999**, *12*, 243.
(84) Ivens, A.; Mayans, O.; Szadkowski, H.; Jürgens, C.; Wilmanns, M.; Kirschner, K. *Eur. J. Biochem.* **2002**, *269*, 1145.
(85) Gunasekaran, K.; Eyles, S. J.; Hagler, A. T.; Gierasch, L. M. *Curr. Opin. Struct. Biol.* **2001**, *11*, 83.
(86) Forsyth, W. R.; Matthews, C. R. *J. Mol. Biol.* **2002**, *320*, 1119.
(87) Eder, J.; Kirschner, K. *Biochemistry* **1992**, *31*, 3617.
(88) Jasanoff, A.; Davis, B.; Fersht, A. R. *Biochemistry* **1994**, *33*, 6350.
(89) Soberon, X.; Fuentes-Gallego, P.; Saab-Rincon, G. *FEBS Lett*. **2004**, *560*, 167.
(90) Bertolaet, B. L.; Knowles, J. R. *Biochemistry* **1995**, *34*, 5736.
(91) Pan, H.; Raza, A. S.; Smith, D. L. *J. Mol. Biol.* **2004**, *336*, 1251.
(92) Silverman, J. A.; Harbury, P. B. *J. Mol. Biol.* **2002**, *324*, 1031.
(93) Silverman, J. A.; Harbury, P. B. *J. Biol. Chem.* **2002**, *277*, 30968.
(94) Zitzewitz, J. A.; Matthews, C. R. *Biochemistry* **1999**, *38*, 10205.
(95) Rojsajjakul, T.; Wintrode, P.; Vadrevu, R.; Matthews, C. R.; Smith, D. L. *J. Mol. Biol.* **2004**, *341*, 241.
(96) Wintrode, P. L.; Rojsajjakul, T.; Vadrevu, R.; Matthews, C. R.; Smith, D. L. *J. Mol. Biol.* **2005**, *347*, 911.
(97) Wu, Y.; Matthews, C. R. *J. Mol. Biol.* **2003**, *330*, 1131.
(98) Finke, J. M.; Onuchic, J. N. *Biophys. J.* **2005**, *89*, 488.
(99) Rudolph, R.; Siebendritt, R.; Kiefhaber, T. *Protein Sci.* **1992**, *1*, 654.
(100) Wilmanns, M.; Hyde, C. C.; Davies, D. R.; Kirschner, K.; Jansonius, J. N. *Biochemistry* **1991**, *30*, 9161.
(101) Shukla, A.; Guptasarma, P. *Proteins* **2004**, *55*, 548.
(102) Wodak, S. J.; Lasters, I.; Pio, F.; Claessens, M. *Biochem. Soc. Symp.* **1990**, *57*, 99.
(103) Beauregard, M.; Goraj, K.; Goffin, V.; Heremans, K.; Goormaghtigh, E.; Ruysschaert, J. M.; Martial, J. A. *Protein Eng*. **1991**, *4*, 745.
(104) Tanaka, T.; Hayashi, M.; Kimura, H.; Oobatake, M.; Nakamura, H. *Biophys. Chem.* **1994**, *50*, 47.
(105) Tanaka, T.; Kimura, H.; Hayashi, M.; Fujiyoshi, Y.; Fukuhara, K.; Nakamura, H. *Protein Sci.* **1994**, *3*, 419.
(106) Tanaka, T.; Kuroda, Y.; Kimura, H.; Kidokoro, S.; Nakamura, H. *Protein Eng.* **1994**, *7*, 969.
(107) Houbrechts, A.; Moreau, B.; Abagyan, R.; Mainfroid, V.; Preaux, G.; Lamproye, A.; Poncin, A.; Goormaghtigh, E.; Ruysschaert, J. M.; Martial, J. A.; et al. *Protein Eng.* **1995**, *8*, 249.
(108) Offredi, F.; Dubail, F.; Kischel, P.; Sarinski, K.; Stern, A. S.; Van de Weerdt, C.; Hoch, J. C.; Prosperi, C.; Francois, J. M.; Mayo, S. L.; Martial, J. A. *J. Mol. Biol.* **2003**, *325*, 163.
(109) Scheerlinck, J. P.; Lasters, I.; Claessens, M.; De Maeyer, M.; Pio, F.; Delhaise, P.; Wodak, S. J. *Proteins* **1992**, *12*, 299.
(110) Dahiyat, B. I.; Sarisky, C. A.; Mayo, S. L. *J. Mol. Biol.* **1997**, *273*, 789.
(111) Lang, D.; Thoma, R.; Henn-Sax, M.; Sterner, R.; Wilmanns, M. *Science* **2000**, *289*, 1546.
(112) Thoma, R.; Schwander, M.; Liebl, W.; Kirschner, K.; Sterner, R. *Extremophiles* **1998**, *2*, 379.
(113) Höcker, B.; Claren, J.; Sterner, R. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 16448.
(114) Gerlt, J. A.; Babbitt, P. C. *Nat. Struct. Biol.* **2001**, *8*, 5.
(115) Heinz, D. W.; Essen, L. O.; Williams, R. L. *J. Mol. Biol.* **1998**, *275*, 635.
(116) Höcker, B.; Schmidt, S.; Sterner, R. *FEBS Lett.* **2002**, *510*, 133.
(117) Nicolet, Y.; Drennan, C. L. *Nucleic Acids Res.* **2004**, *32*, 4015.
(118) Farber, G. K.; Petsko, G. A. *Trends Biochem. Sci.* **1990**, *15*, 228.
(119) Babbitt, P. C.; Gerlt, J. A. *Adv. Protein Chem.* **2000**, *55*, 1.
(120) Henn-Sax, M.; Höcker, B.; Wilmanns, M.; Sterner, R. *Biol. Chem.* **2001**, *382*, 1315.
(121) Copley, R. R.; Bork, P. *J. Mol. Biol.* **2000**, *303*, 627.
(122) Seibert, C. M.; Raushel, F. M. *Biochemistry* **2005**, *44*, 6383.
(123) Gerlt, J. A.; Babbitt, P. C. *Annu. Rev. Biochem.* **2001**, *70*, 209.
(124) Palmer, D. R.; Wieczorek, S. W.; Hubbard, B. K.; Mrachko, G. T.; Gerlt, J. A. *J. Am. Chem. Soc.* **1997**, *119*, 9580.
(125) Schmidt, D. M.; Mundorff, E. C.; Dojka, M.; Bermudez, E.; Ness, J. E.; Govindarajan, S.; Babbitt, P. C.; Minshull, J.; Gerlt, J. A. *Biochemistry* **2003**, *42*, 8387.
(126) Palmer, D. R.; Garrett, J. B.; Sharma, V.; Meganathan, R.; Babbitt, P. C.; Gerlt, J. A. *Biochemistry* **1999**, *38*, 4252.
(127) Thoden, J. B.; Taylor Ringia, E. A.; Garrett, J. B.; Gerlt, J. A.; Holden, H. M.; Rayment, I. *Biochemistry* **2004**, *43*, 5716.
(128) Taylor Ringia, E. A.; Garrett, J. B.; Thoden, J. B.; Holden, H. M.; Rayment, I.; Gerlt, J. A. *Biochemistry* **2004**, *43*, 224.
(129) Chen-Goodspeed, M.; Sogorb, M. A.; Wu, F.; Hong, S. B.; Raushel, F. M. *Biochemistry* **2001**, *40*, 1325.
(130) Chen-Goodspeed, M.; Sogorb, M. A.; Wu, F.; Raushel, F. M. *Biochemistry* **2001**, *40*, 1332.
(131) Vincent, F.; Yates, D.; Garman, E.; Davies, G. J.; Brannigan, J. A. *J. Biol. Chem.* **2004**, *279*, 2809.
(132) Liaw, S. H.; Chen, S. J.; Ko, T. P.; Hsu, C. S.; Chen, C. J.; Wang, A. H.; Tsai, Y. C. *J. Biol. Chem.* **2003**, *278*, 4957.
(133) Ireton, G. C.; McDermott, G.; Black, M. E.; Stoddard, B. L. *J. Mol. Biol.* **2002**, *315*, 687.
(134) Nitanai, Y.; Satow, Y.; Adachi, H.; Tsujimoto, M. *J. Mol. Biol.* **2002**, *321*, 177.
(135) Beismann-Driemeyer, S.; Sterner, R. *J. Biol. Chem.* **2001**, *276*, 20387.
(136) Jürgens, C.; Strom, A.; Wegener, D.; Hettwer, S.; Wilmanns, M.; Sterner, R. *Proc. Natl. Acad. Sci. U.S.A.* **2000**, *97*, 9925.
(137) Leopoldseder, S.; Claren, J.; Jürgens, C.; Sterner, R. *J. Mol. Biol.* **2004**, *337*, 871.
(138) Barona-Gomez, F.; Hodgson, D. A. *EMBO Rep*. **2003**, *4*, 296.
(139) Kuper, J.; Doenges, C.; Wilmanns, M. *EMBO Rep.* **2005**, *6*, 1.
(140) Wise, E.; Yew, W. S.; Babbitt, P. C.; Gerlt, J. A.; Rayment, I. *Biochemistry* **2002**, *41*, 3861.
(141) Yew, W. S.; Akana, J.; Wise, E. L.; Rayment, I.; Gerlt, J. A. *Biochemistry* **2005**, *44*, 1807.
(142) Wise, E. L.; Yew, W. S.; Akana, J.; Gerlt, J. A.; Rayment, I. *Biochemistry* **2005**, *44*, 1816.
(143) Borchert, T. V.; Abagyan, R.; Jaenicke, R.; Wierenga, R. K. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 1515.
(144) Saab-Rincon, G.; Juarez, V. R.; Osuna, J.; Sanchez, F.; Soberon, X. *Protein Eng.* **2001**, *14*, 149.
(145) Stemmer, W. P. *Proc. Natl. Acad. Sci. U.S.A.* **1994**, *91*, 10747.
(146) Stemmer, W. P. *Nature* **1994**, *370*, 389.
(147) Cho, C. M.; Mulchandani, A.; Chen, W. *Appl. Environ. Microbiol.* **2002**, *68*, 2026.
(148) Merz, A.; Yee, M. C.; Szadkowski, H.; Pappenberger, G.; Crameri, A.; Stemmer, W. P.; Yanofsky, C.; Kirschner, K. *Biochemistry* **2000**, *39*, 880.
(149) Lebbink, J. H.; Kaper, T.; Bron, P.; van der Oost, J.; de Vos, W. M. *Biochemistry* **2000**, *39*, 3656.

(150) Lönn, A.; Gardonyi, M.; van Zyl, W.; Hahn-Hagerdal, B.; Otero, R. C. *Eur. J. Biochem.* **2002**, *269*, 157.
(151) Bocola, M.; Otte, N.; Jaeger, K. E.; Reetz, M. T.; Thiel, W. *Chembiochem.* **2004**, *5*, 214.
(152) Sriprapundh, D.; Vieille, C.; Zeikus, J. G. *Protein Eng.* **2003**, *16*, 683.
(153) Giver, L.; Gershenson, A.; Freskgard, P. O.; Arnold, F. H. *Proc. Natl. Acad. Sci. U.S.A.* **1998**, *95*, 12809.
(154) Cheon, Y. H.; Park, H. S.; Kim, J. H.; Kim, Y.; Kim, H. S. *Biochemistry* **2004**, *43*, 7413.
(155) May, O.; Nguyen, P. T.; Arnold, F. H. *Nat. Biotechnol.* **2000**, *18*, 317.
(156) Abendroth, J.; Niefind, K.; May, O.; Siemann, M.; Syldatk, C.; Schomburg, D. *Biochemistry* **2002**, *41*, 8589.
(157) Wymer, N.; Buchanan, L. V.; Henderson, D.; Mehta, N.; Botting, C. H.; Pocivavsek, L.; Fierke, C. A.; Toone, E. J.; Naismith, J. H. *Structure* **2001**, *9*, 1.
(158) Fong, S.; Machajewski, T. D.; Mak, C. C.; Wong, C. *Chem. Biol.* **2000**, *7*, 873.
(159) Wada, M.; Hsu, C. C.; Franke, D.; Mitchell, M.; Heine, A.; Wilson, I.; Wong, C. H. *Bioorg. Med. Chem.* **2003**, *11*, 2091.

CR030191Z