ACS Catalysis

Viewpoint

pubs.acs.org/acscatalysis

Engineering of Flexible Loops in Enzymes

Bettina M. Nestl and Bernhard Hauer*

Institute of Technical Biochemistry, University of Stuttgart, Allmandring 31, 70569 Stuttgart, Germany

INTRODUCTION

From a chemical point of view, enzymes are by far the most structurally complex and functionally sophisticated molecules known. This is perhaps not surprising, once one realizes that the structure and chemistry of each enzyme has been developed and fine-tuned over billions of years of evolutionary history. Generally, enzymes are characterized by conservation of their functional groups through evolution. However, detailed enzyme studies have revealed the catalytic flexibility of many active sites also discussed as plasticity, in that different functional groups mediate the same mechanistic role. This variability resulted from either hopping of functional groups from one position to another to optimize catalysis, the independent specialization of a low-activity primordial enzyme in different phylogenetic lineages, or the functional convergence after evolutionary divergence or circular permutation events.[1] Upon catalysis, residues in and around the active site must undergo conformational changes associated with the binding and release of substrates and cofactors, the protection of the reaction space from the aqueous environment, the stabilization of reaction intermediates, and the interactions between catalytic residues and binding subsites with various groups of reactants upon completion of catalysis.[2] Further, interactions of the macromolecule with substrate and cofactor require conformational changes either involving the amino acid side chains capable of adopting different rotamers or the external loops located in the periphery of the active site. The particular sequence of an enzyme defines the folding pathway and structure. Secondary structure elements, such as α-helices and β-strands constitute the major building blocks of enzymes. Some important structural elements adopted by enzymes are also loops. Loops are considered a diverse class of secondary structures comprising turns, random coils, and strands which connect the main secondary structures (α-helices and β-strands). Loops exposed on the surface often play a vital role in protein functions, primarily because they have a greater change of interacting with the solvent and ligand molecules. Loop regions belong to the most flexible parts of enzyme structures. The evolution of enzymes involves sequence changes that are frequently localized at loop regions and thus reveal their role in protein evolution and in the diversification of numerous enzyme families and superfamilies.[3]

Over the past few years, a surge of interest from the scientific community has brought the study of loops—their flexibility and impact on enzyme function—to the forefront of biocatalysis. Most of the activity in this field so far has centered on the role of surface and lid loops that cover the active site and their functional role in substrate and cofactor binding, protein−protein interaction, and stability. This Viewpoint aims to present recent selected advances in the research of flexible loops in enzyme catalysis including the generation of new enzymatic functions. Given the wealth of publications that cover this field, we have highlighted a few examples regarding the present advances and possible future impacts. The first section of this Viewpoint addresses the role and flexibility of loops on catalysis. The second section highlights enzyme engineering events by the introduction of point mutations in loop backbones that can induce dramatic changes in the enzyme function and specificity. The final section addresses specific advancements in the improvement of enzyme stability through consideration of loop rigidity that is a prerequisite for high thermostability. We emphasize the role of loop flexibility and engineering in the biological activity and compare with experiments where data exist.

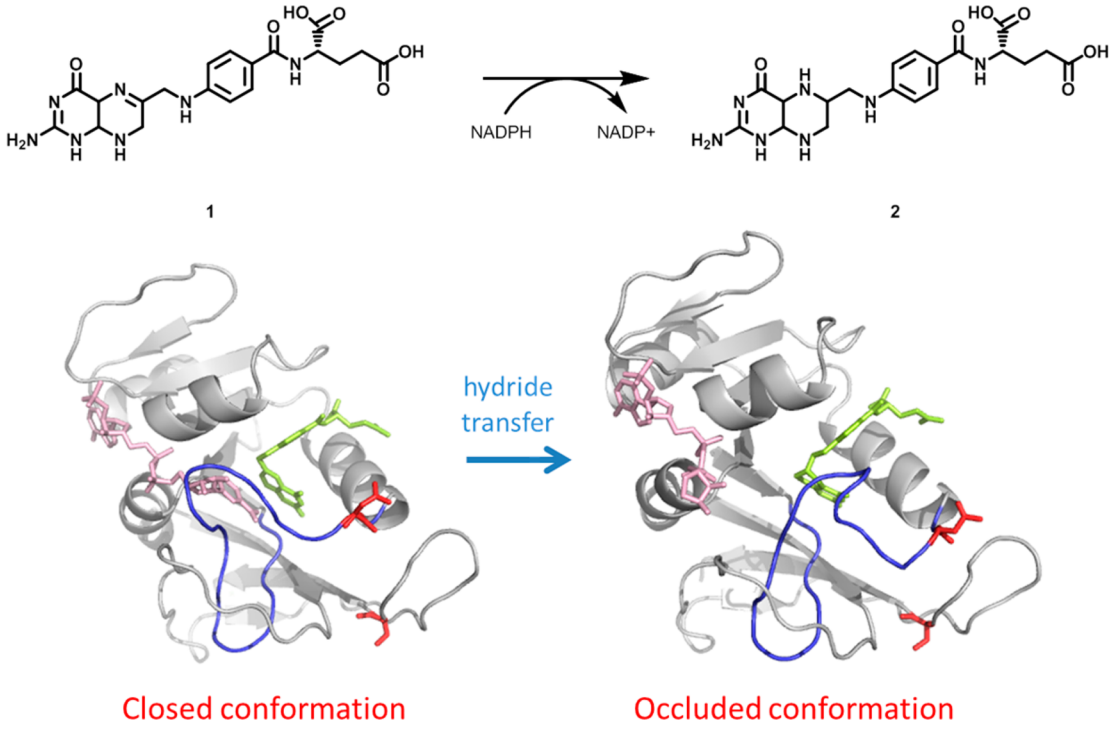
LOOP FLEXIBILITY AND CATALYSIS

Protein−protein and protein−small ligand interactions, signal transduction and assembly of macromolecules, allosteric regulation, and thermal enzymatic adaption are processes which require structural flexibility. In general, enzymes represent attractive models in the study of protein flexibility. Flexibility of the active site is considered as a requirement for reduction of the free energy barrier and acceleration of the enzymatic reaction.[4] Conformational changes are frequently observed as part of enzyme mechanisms, according to which substrate and specific ligand binding is associated with a conformational change. Further, it has been shown that protein motions are critical for enzymatic function; however little is known about their precise role in catalysis. Researchers in the groups of Benkovic and Wright have used a number of experimental and theoretical investigations to suggest that protein motions play an important role in catalysis by the enzyme dihydrofolate reductase DHFR (Scheme 1).[5,6] *Escherichia coli* DHFR (*ec*DHFR) catalyzes the reduction of 7,8-dihydrofolate **1** through stereospecific hydride transfer from reduced NADPH cofactor. It has several flexible loops surrounding the active site that play a functional role in substrate and cofactor binding and in catalysis. The Met20 loop (residues 9−24) serves as a "lid", which closes over the pteridine and nicotinamide rings of the bound folate and cofactor, thereby adopting a closed as well as occluded conformation. The deletion of the central residues in the Met20 loop, which has been achieved by replacement of residues Met16 to Ala19 by glycine, resulted in a striking 500-fold decrease in the rate of hydride transfer.[7] In the closed conformation, the Met20 loop packs tightly against the nicotinamide ring of the cofactor and seals the active site. In the occluded conformation, the loop projects into the active site and sterically occludes the binding site for the nicotinamide-ribose moiety. The occluded conformation is

Received: March 12, 2014
Published: August 8, 2014

   dx.doi.org/10.1021/cs500325p | ACS Catal. 2014, 4, 3201−3211

Scheme 1. Conformational Change of Met20 Loop in the *E. coli* Dihydrofolate Reductase DHFR in the Reduction of Dihydrofolate 1 to Tetrahydrofolate Product 2[a]



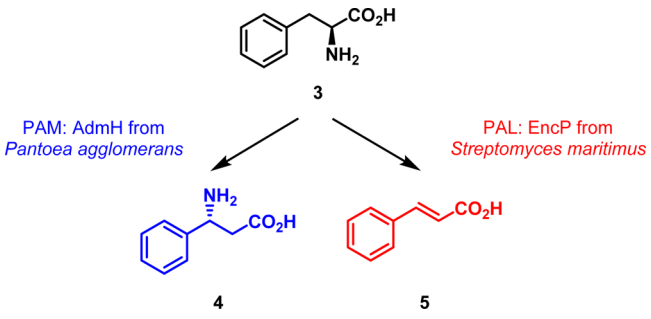
[a]Before hydride transfer, the Met20 loop is in the closed conformation (PDB code 1RX2), in which it packs tightly against the nicotinamide ring of $NADP^+$. After hydride transfer, the loop adopts the occluded conformation (PDB code 1RX4), in which the nicotinamide ring of $NADP^+$ is sterically hindered from binding in the active site. $NADP^+$ undergoes a concurrent conformational change in which the nicotinamide ring is expelled from the binding pocket. Loop Met20 is highlighted in blue. The sites of mutation, Asn23 and Ser148, are shown as sticks in red. $NADP^+$ and NADPH are shown in rose as sticks with dihydrofolate **1** and tetrahydrofolate **2** colored as yellow sticks.

stabilized by a network of hydrogen bonds between Asn23 and the backbone as well as side chain of Ser148. Furthermore, the loop conformational transitions are accompanied by changes in hydrogen bonding to the so-called FG and GH loops, connecting βF and βG strands (residues Ala117−Asp131) and βG and βH strands (residues Asp142−His149), respectively. The role of the active site loop fluctuations in catalysis has been studied by various methodologies using X-ray crystallography, NMR spectroscopy, molecular dynamics simulations, measurements of the temperature-dependence of the intrinsic kinetic isotope effects (KIEs), and other kinetic parameters, as well as Förster resonance energy transfer (FRET).[8−11] The NMR data have revealed that in the apo enzyme the Met20 loop fluctuates between the two conformations at a rate comparable with tetrahydrofolate **2** dissociation, which is rate-determining. High-frequency motions on the nanosecond/picosecond time scale persist in the peptide backbone for residues 16 to 19 of the Met20 loop known to undergo a ligand-dependent conformational change.[12,13] Furthermore, mutants of Asn23 and Ser148 (N23PP, S148A, and N23PP/S148A) have been created and characterized. In this case, the N23PP and N23PP/S148A mutations altered the conversion from the closed to occluded states of the Met20 loop as well as the hydride transfer rate constant by freezing its dynamical motion. The data provide insights into the changes in backbone dynamics during the catalytic cycle and point to an important role of the Met20 loop in controlling access to the active site, according to which the closed conformation of this loop is essential for catalysis. Further experimental and computation studies have proposed a dynamic network in DHFR that includes residues in and remote from the active site. Especially, mutations of two remote residues (Met42Trp and Gly121Val) have demonstrated that these residues are part of the dynamic network that is coupled to the catalyzed chemistry and the loop dynamic fluctuations.[14,15] Nonetheless, some controversy has arisen with respect to the interpretation of these experiments that enzymatic catalysis is due to conformational fluctuations. However, it has been pointed out by the group of Warshel that the changes in the reaction potential surface modify the reorganization free energy (which includes entropic effects), and that such changes in the surface also alter the corresponding motion.[16] What becomes very clear from the extensive work of the groups of Benkovic and Wright and many others on the dihydrofolate reductase is that protein dynamics are at the heart of enzyme catalysis, but identification and analysis of dynamical effects in enzyme-catalyzed reactions have proved very challenging.

In another example, the group of Micklefield has elucidated the thermal bifunctionality of bacterial phenylalanine aminomutase (PAM) and ammonia lyase (PAL) enzymes (Scheme 2).[17] PAMs and PALs are 4-methylidene-imidazol-5-one (MIO)-dependent enzymes that catalyze the isomerization of L-α-phenylalanine to give (*S*)- or (*R*)-β-phenylalanine as well as amination of *trans*-cinnamic acid to L-α-phenylalanine using ammonia, respectively. They discovered a temperature-dependent switch that dictates mutase and lyase activity with an inner lid-like loop implicated in driving these thermal energy differences in activity. The PAM AdmH from *Pantoea agglomerans* and the PAL EncP from the thermotolerant

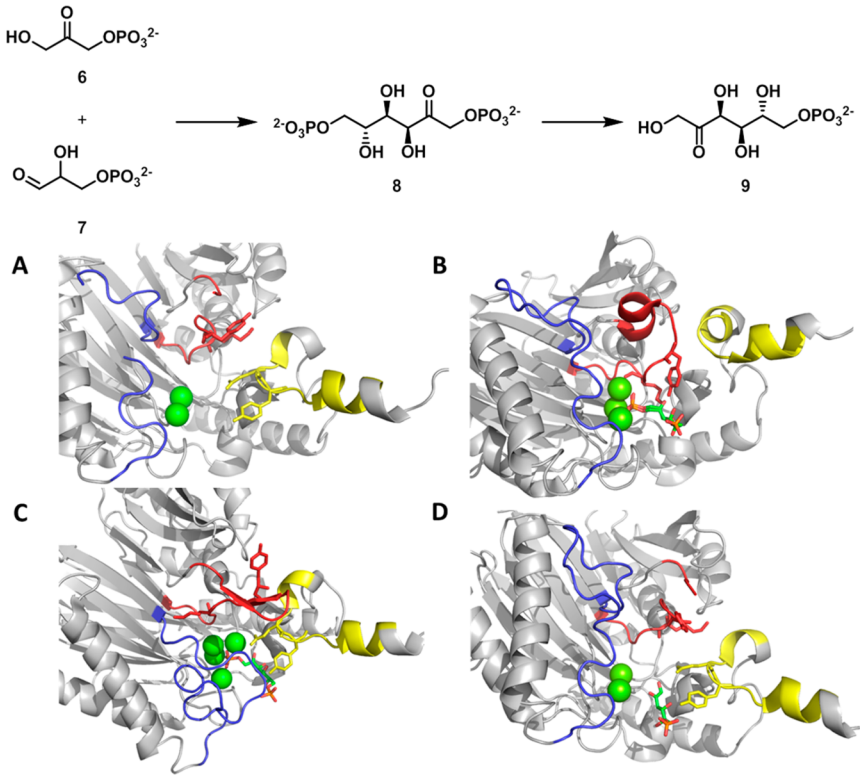
dx.doi.org/10.1021/cs500325p | *ACS Catal.* 2014, 4, 3201−3211

Scheme 2. Thermal Bifunctionality of the Bacterial Phenylalanine Aminomutase AdmH and the Phenylalanine Ammonia Lyase EncP Converting L-α-Phenylalanine 3 to Either (*R*)- or (*S*)-β-Phenylalanine 4 or *trans*-Cinnamic Acid 5



Streptomyces maritimus, sharing significant sequence similarity, were selected as model enzymes. PAM and PAL display thermal bifunctionality with mutase activity predominant at lower temperature (30 °C) and lyase activity preferred at higher temperatures (70−75 °C). Homology models revealed that each enzyme contains two mobile loops flanking the entrance to the active site: (i) the inner loop containing the catalytic base (Tyr78 in AdmH) and (ii) the outer loop from an adjacent monomer (residues 285−320 in AdmH). However, through solvent viscosity and kinetic isotope effect studies, the researchers were able to demonstrate that conformational changes of the inner loop facilitating product release and elimination of ammonia from MIO as well as the cleavage of the C3 hydrogen atom (abstraction of the β-proton) are the rate-limiting steps for PAL and PAM reactions. Abstraction of the C3 hydrogen atom during the PAL reaction is coupled with a conformational change of the inner loop, whereas at lower temperatures of the PAM reaction, no disengagement of the inner loop after abstraction is assumed. On the basis of these observations, they conclude that the inner loop acts as a lid whereby its mobility influences whether immediate release of the lyase product (cinnamic acid) will occur after deamination of the amino acid substrate or whether the release of the mutase product occurs after a relatively longer closed lid structure allows readdition of the amine from the aminated MIO cofactor.

Scheme 3. Remodeling of the Active Site of the Bifunctional Fructose-1,6-Bisphosphate Aldolase/Phosphatase through the Movement of Loop Regions That Create Different Catalytic Functionalities at the Same Location[a]



[a]The enzyme (PDB code 3T2B) mediates the aldol reaction of dihydroxyacetone phosphate **6** and glyceraldehyde 3-phosphate **7** to fructose-1,6-bisposphate **8**, as well as the subsequent irreversible hydrolysis to fructose-6-phosphate **9**. The aldolase, anchor, and phosphatase lid loops are colored in red, yellow, and blue, respectively. Amino acids Lys232 and Tyr229, essential for aldolase activity with Ly232 forming the Schiff base, are highlighted as sticks in yellow. Anchor loop residues Glu357 and Tyr358 important for phosphatase activity are shown as sticks in red. (A) The ligand-free state of the enzyme binding two Mg^{2+} ions (green spheres) is shown. (B) Binding of the substrates **6** and **7** (shown as sticks) as well as the third Mg^{2+} ion (green spheres) redirect the aldolase loop to form a Schiff base with the substrate. The phosphatase and anchor loop are displaced outwards from the active site during aldol reaction. (C) Hydrolysis of the Schiff base releases the loop containing Lys232 from the active site, thereby enabling the enzyme to bind a fourth Mg^{2+} ion (green spheres). The closure of the active site by the phosphatase lid and anchor loops is important for phosphatase reaction due to stabilization of **8** (sticks) in the active site. (D) Hydrolysis of the phosphate ester and formation of the product **9** is shown in sticks.

dx.doi.org/10.1021/cs500325p | *ACS Catal.* 2014, 4, 3201−3211

A recent report by the Einlse and the Wakage lab on the bifunctionality of fructose-1,6-bisphosphate aldolase/phosphatase also reveals conformational changes in the active site architecture in order to exhibit dual activity of this enzyme (Scheme 3).[18,19] Fructose-1,6-bisphosphate aldolase (FBPAP) catalyzes the aldol reaction of dihydroxyacetone phosphate **6** (DHAP) and glyceraldehyde 3-phosphate **7** (GAP) to fructose-1,6-bisphosphate **8** (FBP), as well as the subsequent irreversible dephosphorylation of **8** to fructose-6-phosphate **9** (F6P) and inorganic phosphate. X-ray crystallography is the method of choice for obtaining the molecular structure of proteins at atomic resolution. In fact, crystallography provides a static model of a dynamic molecule. Crystallographers sometimes emphasize the parts of a protein molecule that might not be seen in a crystal structure as evidence of motion or disorder. However, even the static crystal structure contains information for the degree of thermal motion of the atoms. Especially, water molecules oscillating with large amplitude around a conformational energy minimum will be hard to detect, and also molecules with small crystallographic occupancies caused by conformational disorder and exchange with the bulk solvent. In this context, the positional spread of each atom monitored by the mean-squared atomic displacements (B-factors) can be taken as indication for the relative vibrational motion of different parts of the structure. A comparison of them allows for drawing the conclusion regarding the most flexible parts. Atoms with low B-factors belong to a part of the structure that is well ordered, whereas atoms with large B-factors generally belong to a part that is very flexible or labile with respect to binding.[20] Crystal structure analyses of the enzymes from *Sulfolobus tokodaii* (ST0318) and *Thermoproteus neutrophilus* (*Tn*FBPAP) in a ligand-free state as well as in complex with the substrates **6** and **7** showed that three loop regions surrounding the active site of the enzyme substantially change their conformation when substrate, reaction intermediate, or product is bound. Such an reorientation of the loops during catalysis alter the structure and therefore functionality of the active site by bringing amino acid side chains and the catalytically essential Mg^{2+} ions into place. The researchers have organized these rather labile loops as follows: (i) the *aldolase* or *Schiff-base* loop comprising Lys232 and Tyr229 essential for aldolase activity (residues 220−235 in *Tn*FBPAP, 219−233 in ST0318), (ii) the *anchor* or *carboxy C-terminal* loop responsible for fixation of **8** after the aldol condensation (residues 353−364 in *Tn*FBPAP, 346−361 in ST0318), and (iii) the *phosphatase lid* loop fixing intermediate **8** in the binding pocket for phosphatase reaction also undergoing the largest conformational change (residues 89−111 in *Tn*FBPAP, 97−110 in ST0318). In addition, the consecutive binding of four Mg^{2+} ions to the protein is key to switch between aldolase and phosphatase activity. Moreover, Wakagi and co-workers noticed relatively higher average B-factors than that for the overall polypeptide, indicating their flexible nature. Site-directed mutagenesis of five key residues on the three loops has emphasized the tight linkage of the dual activities of the enzyme to distinct amino acid residues and their separate inactivation upon substitution.

Studies have demonstrated that the role of active site flexibility in enzyme catalysis still remains to be explored. In the induced fit hypothesis originally proposed by Koshland, the presence of the substrate induces a conformational change at the active site so as to fit with the structure of the substrate. In particular, the analysis of the *E. coli* dihydrofolate reductase liganded with cofactors and substrates showing the enzyme in different conformational states while complexed with different ligands suggests that the enzyme molecule passes through different conformational states through the whole process of catalysis. All the above indicates that the active site flexibility plays an important role in enzyme catalysis. It is possible that during the catalytic cycle, the enzyme molecule passes through different stages, and each stage requires the molecule to be in a different conformation, especially at the active site. Rapid transition between the different conformational states, and hence the flexibility of the active site, is therefore mandatory for the maximal expression of enzyme activity.

LOOP FLEXIBILITY AND SPECIFICITY

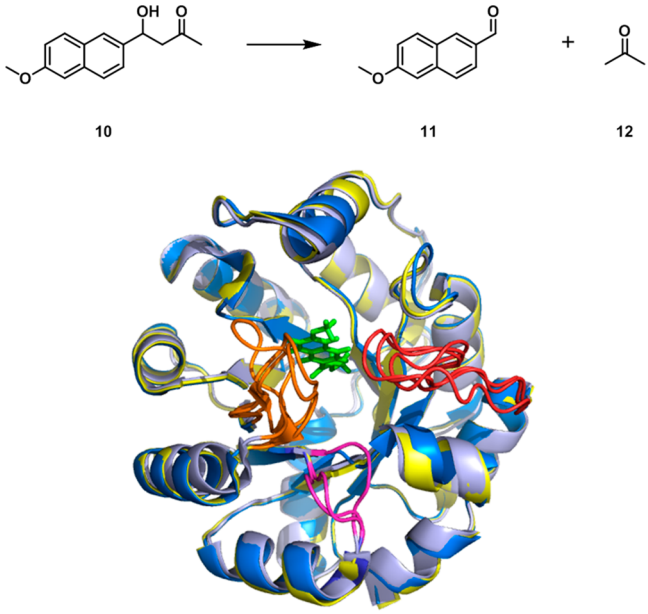
The broad substrate specificity which characterizes some enzymes has been explained under the view of a flexible active site cavity enabling the accommodation of stereochemically diverse substrates. The different substrates may either induce conformational rearrangements to the broader area of the flexible active site or bind to different conformations with different affinity taking the dynamic ensemble of conformations into account. The plasticity as well as local flexibility around the active site and the possible reposition and repacking of sometimes only a few individual side chains causes sufficient space and shape alterations in order to switch from one substrate to another without substantial structural changes.[1] However, in other cases, a more global flexibility is required for promiscuity.[21] The flexibility of loops and their low evolutionary pressure is reflected by the limitation to Ramachandran space and their low sequence conformation.[2] Loops participating in the active sites of enzymes are an exception. These loops often referred to as lids, turns or flaps, and their location at the entry of the active site play a major role in substrate selectivity and recognition and facilitation of substrate binding into the binding cleft. The structural comparison of apo- and holo-states of various enzymes highlights the main structural difference between the two states as the different conformations of active/binding site loops. Upon substrate binding, the disordered, solvent-exposed loop adopts therefore a more compact and ordered conformation-making interaction with subsites of the ligand and/or other residues of the protein. The closing of the cap domain traps the ligand into the cleft and shields fully or partially the ligand molecule from the aqueous environment. Moreover, the access of the active site to other molecules is prevented, and the reaction intermediates are protected and stabilized. Two models are used to explain the transition from the open to the closed conformation: (i) the well-known induced fit model with large conformational changes at a local level and transition to the closed conformation being induced by substrate binding into the active site cleft and (ii) the conformational selection model with the enzyme to pre-exist in two alternative and stable conformations which differ in their ability to bind ligands. It is well understood that the distinction between the induced fit and the original conformational selection models is not absolute.[22,23] The practical advantages of the representation of the ligand as static structure, the notable success attained over the past few decades with such simple models, and the absence of clear guidelines for weighing the advantages against the disadvantages of accounting for flexibility have prompted some investigators to stretch the rigid model beyond its scope.[24] Furthermore, many proteins display functional promiscuity which requires them to be able to interact with multiple partners. Promiscuous proteins that are able to bind to multiple partners through conformational

dx.doi.org/10.1021/cs500325p | *ACS Catal.* 2014, 4, 3201−3211

selection need to explore a larger conformational space than those that bind to only a single partner. More rigid binding sites therefore may have restricted specificity with the benefit of higher binding affinity. Research about the relationship between binding promiscuity, conformational flexibility, and evolvability of proteins has been reviewed by Tokuriki and colleagues.[25] There are a few examples of documented loops that are involved in protein motion related with function. "Triggering" loops were identified, whose conformational changes are required for the catalysis.[26] Molecular dynamics simulations were used to examine the conformational change in the loops essential for 1,4-galactosyltransferase-1, enolase, and lipase function. Such loops are hardly recognizable by the comparison of the apo- and holo-crystal structures. Furthermore, these enzymes belong to different classes, have different folds, and perform different functions. There is no detectable structural or sequence homology between these three enzymes. In particular, loops have mixed sequences and are therefore often difficult to identify from sequence databases. The studies have shown that all three enzymes possess functional loops, whose movements are triggered by a second, interacting loop. This triggering loop is highly conserved in the respective family. Site-directed mutagenesis experiments have shown that modification of the interactions between the functional and the highly conserved triggering loops alter the enzymatic activity. Lids are not the only flexible structural elements around the active sites. Quite often the rearrangement of the active site geometry is accompanied or facilitated by movements and conformational changes of secondary structural elements adjacent to the active site, for examples, helices which frequently precede or follow the lids. A characteristic example is the cap domain of the haloalkane dehalogenase DhlA from *Xanthobacter autotrophicus*. Early work of Janssen and co-workers on DhlA pointed into the direction that short repeats in the N-terminal part of the cap domain play an important role in the evolution of haloalkane dehalogenase specificity. Herein, the cap domain of DhlA acts as an activity-modifying domain that can accommodate spontaneous mutations which lead to alterations of substrate specificity.[27]

Nowadays it is possible to design enzyme active sites de novo. Generally, the designed catalytic residues are introduced into existing protein scaffolds with no initial activity. A recent example is that of a designed retro-aldolase with a remarkable rate enhancement of 15 000 over background. The retro-aldolase catalyzes the cleavage of 4-hydroxy-4-(6-methoxy-2-naphthyl)-2-butanone **10** (methodol) to 6-methoxy-2-naphthaldehyde **11** and acetone **12**. Although the initial activities were not high, they could be significantly increased by laboratory evolution.[28,29] The researchers in the lab of Hilvert used the previously designed retro-aldolase RA95.0 and applied further evolution, which led to a complete active site remodeling and yielded in catalysts achieving efficiencies approaching those of natural enzymes (Scheme 4).[30] The retro-aldolase enzyme RA95.0 belongs to the $(\beta\alpha)_8$-barrel family and contains already 11 active site mutations predicted by the Rosetta software suite. Furthermore, it possesses a reactive lysine at position 210 to form a protonated Schiff base with the substrate and a glutamate at position 53 for a water-mediated stabilization of the β-hydroxy group of **10**. For further enhancements of the activity, iterative cassette mutagenesis of residues at each position at or near the active site has been performed. The best variant RA95.5 contains five active site mutations (Val51Tyr, GLu53Ser, Thr83Lys, Met180Phe, and

Scheme 4. Loop Mobility and Active Site Remodeling[a]

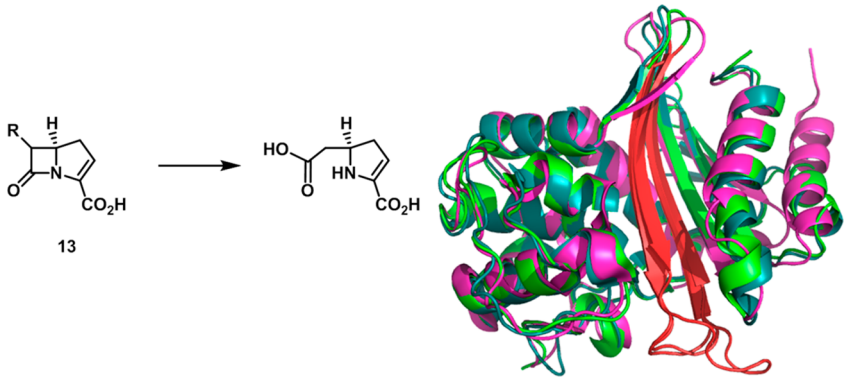


[a]Loop mobility and active site remodeling afforded a retro-aldolase variant with significantly higher catalytic activity. Retro-aldolase is a computationally designed enzyme that utilizes a lysine to convert methodol **10** to the corresponding aldehyde **11** and acetone **12**. Superimposition of three retro-aldolase variants RA95.0 (light blue, PDB code 4A29), RA95.5 (yellow, PDB code 4A2S), and RA95.5-5 (blue, PDB code 4A2R) crystal structures with substrate bound (green sticks) highlight the conformational changes in the loops L1 (red), L7 (pink) and L6 (orange) that occurred during evolutionary optimization.

Arg182Met) of which three induce a conformational change in proximity to the active site. The substitution of active site residue Thr83 resulted in the appearance of a second lysine amino acid residue, which can be found in natural class I aldolases. More surprisingly, the mutagenesis study showed that the newly introduced lysine at position 83 (Thr83Lys) is more effective than Lys210 as a catalytic group. Additional kinetic studies revealed that the more buried Lys83 has a 30-fold advantage over Lys210 under subsaturating conditions, which might be due to better shielding from bulk solvent. The other mutations substantially remodeled the active site by inducing a major conformational change in the L6 loop that creates a deep hydrophobic binding pocket. This was not present in the original computational design demonstrated by the comparison of crystal structures of wild type and variant. Moreover, active site plasticity facilitated generation of this new reaction center through conformational changes of the three loops L1, L6, and L7 bracketing the active site of the retro-aldolase variant. The design could be further improved by a combination of fully random gene diversification using error-prone PCR and DNA shuffling. The most active clone from the eighth round of directed evolution called RA95.5-8 is a remarkably effective retro-aldolase and has a 14-fold higher TON than the commercially available aldolase antibody 38C2. Crystallization of the clone RA95.5-5 from the fifth round of directed evolution in complex with a diketone inhibitor elucidated that conformations of three loops L1, L6, and L7 underline the pronounced deviations from those in the original design model RA95.0. Hilvert and co-workers conclude from their work that conformational plasticity is crucial for the emergence and

dx.doi.org/10.1021/cs500325p | *ACS Catal.* 2014, 4, 3201−3211

Scheme 5. Structural Alignment of Class D Carbapenemases Hydrolyzing Carbapenems 13 Including the Narrow-Spectrum OXA-10 β-Lactamase[a]



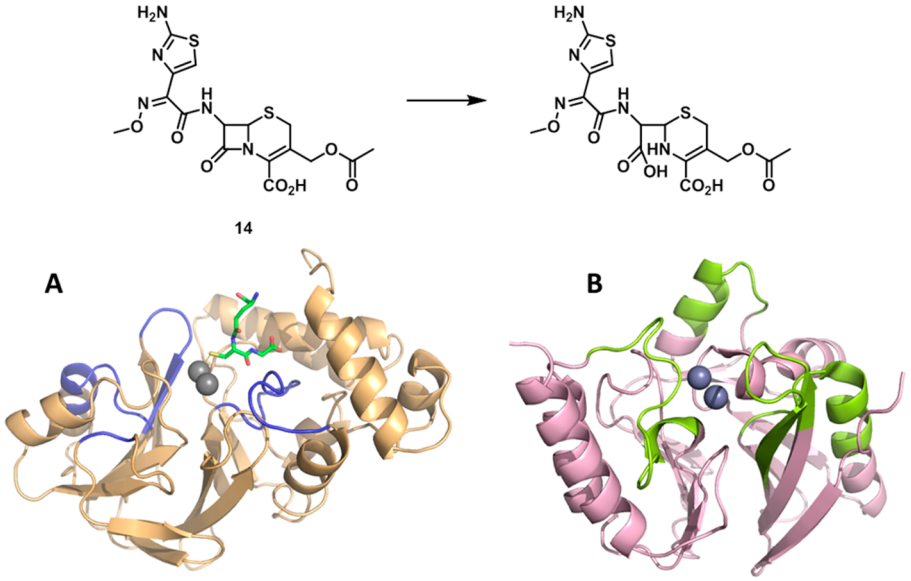
[a]Superimposed structures of native OXA-10 (dark cyan, PDB code 1FOF), OXA-24 (pink, PDB code 2JC7), and OXA-48 (green, PDB code 3HBR) are shown. The location of the catalytically relevant β5-β6 turn connecting the labeled β5 and β6 strands is highlighted in red.

improvement of a new catalytic center. Furthermore, the relocation of reactive lysine in the active site as well as the conformational changes in surface loops have emphasized the prospects to generate catalysts with true enzymatic activities.

Despite evident progress, the palette of reactions catalyzed by computationally designed enzymes remains limited. Precise positioning of amino acid side chains are required to interact with the functional groups of substrates to alter the specificity of enzymes. The alteration of enzymatic function through point mutations is a frequently used method; however, major changes in function require major sequence arrangements including insertion, deletion, and recombination. Loop grafting represents an innovative method, in which several active loops (variable regions) of an enzyme are replaced, while preserving its scaffold and key catalytic residues for the generation of novel family members with new reaction specificity. Recently, Baker and co-workers have described the redesign of the specificity of human guanine deaminase (hGDA) by introducing specific cytosine deaminase activity through directed remodeling of loops near the active site.[31] Active site comparisons of the hGDA and the distantly related bacterial cytosine deaminase suggested key interactions with cytosine to be introduced by loop modeling while preserving the residues directly involved in catalysis. A remodeling of the segment between residues 211 and 220 as well as targeted mutagenesis by point mutations increased activity by 2 orders of magnitude. However, activity toward the substrate ammelide was still 7 orders of magnitude lower than the activity of the wild type hGDA for guanine. Furthermore, in order to improve the activity of the computationally designed Diels−Alderase, an undirected remodeling of its protein backbone structure has been reported by the group of Baker.[32] The players of the online game Foldit were asked to remodel the active site loops of the enzyme catalyzing the Diels−Alderase reaction. The insertion of a novel 24-residue helix−loop−helix contributed to the increased catalyst affinity for the diene and dienophile and thus, >18-fold increased enzyme activity. The design of new enzymes by generating variability through point mutations of an existing gene has shown some success. Accordingly, the development of novel methods for the engineering of loops would promote the construction of enzymes with novel functions. The works of Saab-Rincon, Soberon, and colleagues have provided information on protein design through systematic catalytic loop exchange in the $(\beta/\alpha)_8$ fold. Initial studies demonstrated that $(\beta/\alpha)_8$-barrel enzymes constitute good targets for a reshaping of the binding site for new substrates or for grafting of a new catalytic function. Overall 14 loops for the study of the structure−function relationship within the phosphoribosylanthranilate isomerase structure have been exchanged by loops belonging to eight different proteins sharing the same fold but diverse functions by loop grafting and hinge variability experiments.[33,34] A further example for successful loop grafting is the evolution of carbapenemase-hydrolyzing activity in class D β-lactamases (Scheme 5).[35] β-Lactamases play a key role in the development of acquired resistance to β-lactams in many bacterial pathogens. The production of novel β-lactamases capable of hydrolyzing carbapenems (carbepenemase) is one of the major mechanisms of carbapenem resistance. Two structurally and mechanistically unrelated families are known to comprise the most numerous active serine enzymes and the zinc-dependent metallo β-lactamases. Class D (OXA-type) carbapenemases are typically β-lactamases with acquired carbapenem resistance in *Acinetobacter baumannii* and to a lesser extend in *Klebsiella pneumonia*. So far, four class D carbapenemases are known including OXA-23, OXA-24, OXA-48 and OXA-58 with a sequence divergence of 34−63%. The crystal structure of OXA-24 shows the usual tunnel-like entrance to the active site, hypothesized to be essential for the correct orientation of the carbapenem. More recently, the structure of OXA-48 was obtained not showing the same tunnel-like entrance. As a result of structural analysis, molecular dynamics and molecular modeling studies, attempts have been made to demonstrate the role of the residues located in the short β5-β6 loop in determining the behavior of the enzyme toward carbapenem substrates. This turn is in close proximity to the active site. One of which includes the catalytically important conserved motif LysThrGly (KTG) of serine β-lactamases. In order to affirm its central role in carbapenem activity of class D enzymes, the β5-β6 turn of the narrow-spectrum OXA-10 enzyme, which is unable to naturally hydrolyze carbapenems **13**, has been grafted with the corresponding turns of the carbapenemases OXA-23, OXA-24, and OXA-48. Furthermore, the relevance of the β5-β6 turn in the carbapenemase activity was probed by the measurements of minimal inhibitory concentrations, plate-assay-based degradations of the imipenem antibiotic and kinetic analyses. The results obtained indicate that the rationally designed loop mutants are more efficient in hydrolysis than the natural

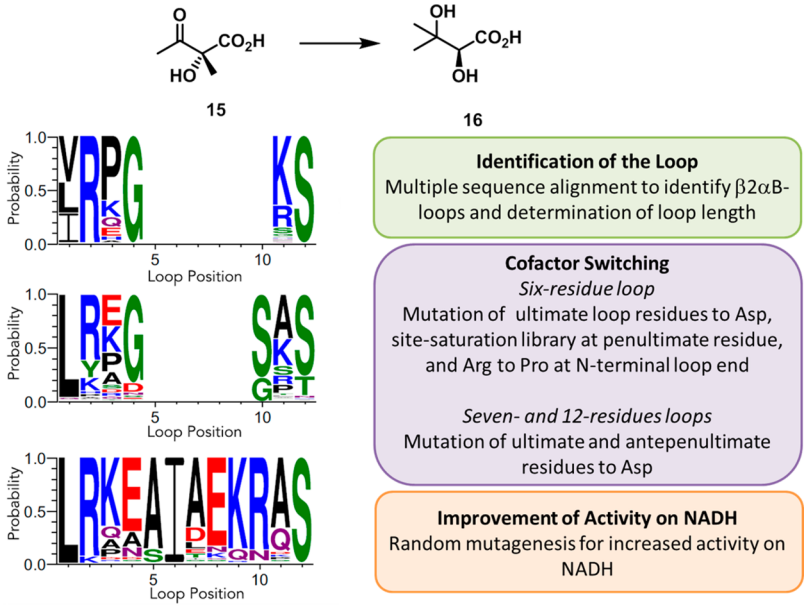
dx.doi.org/10.1021/cs500325p | *ACS Catal.* 2014, 4, 3201−3211

Scheme 6. Structural Comparison of Human Glyoxalase GlyII Colored in Light Orange and IMP-1 Colored in Pink Both Containing $\alpha\beta/\beta\alpha$ Sandwich Structures and Binuclear Metal Ions Essential for Hydrolysis[a]



[a]The natural substrate (*S*)-D-lactoylglutathione of GlyII is shown in green sticks. The replacement of several active site-loops indicated in blue in GlyII (**A**, PDB code 1QH5) and light green in IMP-1 (**B**, PDB code 1DDK) of an enzyme while retaining its scaffold and key catalytic residues including essential Zn^{2+} ions (gray spheres) yielded a family member with new reaction specificity for cefotaxime **14** hydrolysis.

Scheme 7. Guide for Reversing Keto-Acid Reductoisomerase (KARI) Cofactor Dependence from NADPH to NADH[a]



[a]KARI catalyzes the reversible reductive isomerization of 2-acetolactate **15** to 2,3-dihydroisovalerate **16**. The cofactor switching guide comprises the identification of the $\beta2\alpha$B-loop and its length via multiple sequence alignments and generation of loop profiles for six- (top left hand), seven- (middle left hand) and 12-residue loops (bottom left hand). Mutagenesis approaches on loop residues led to reversed cofactor preferences followed by further improvement of activity on NADH by random mutagenesis and screening (adapted from ref 39).

enzymes and that the loop mutants acquired carbapenem-hydrolyzing activity. In addition, the OXA-10 enzyme is not only tolerant to the substitution of its $\beta5$-$\beta6$ turn but also largely maintains its original biochemical properties toward other substrates unaltered. Comparison of the crystal structures of the OXA-10 variants with that of the wild type enzyme did reveal minor differences in the turn leading to the assumption that despite the sequence heterogeneity, conformation is determined by the turn length.

In an attempt to introduce new metallo-β-lactamase function, the redesign of the existing metal-binding site in glyoxalase, known to hydrolyze thioester bonds, has been presented. The insertion, deletion, and substitution of active site loops followed by directed evolution introduced β-lactamase activity into the $\alpha\beta/\beta\alpha$ metallohydrolase scaffold of glyoxylase II (GlyII, Scheme 6).[36] Gly II as well as metallo-β-lactamases like the one from *Pseudomonas aeruginosa* (IMP-1) share the same scaffold, key catalytic residues, and binuclear (typically zinc)

dx.doi.org/10.1021/cs500325p | *ACS Catal.* 2014, 4, 3201−3211

metal ions essential for hydrolysis. Besides zinc, GlyII also contains further metal ions such as iron or manganese in its binuclear site. Comparison of the two protein structures of GlyII and the metallo β-lactamase IMP-1 led to the identification of regions important for the activity switch, which were then modified. To reconstruct the active site of GlyII for the generation of novel activity toward cefotaxime **14**, the binding domain of GlyII has been deleted. Partially randomized functional loops originating from the natural metallo β-lactamase IMP-1 were inserted, and metal-binding residues were changed to those found in the IMP-1 active site. There are no common characteristics to substrate binding between GlyII and IMP-1. The loss of critical substrate-binding features causes the template to lose all natural glyoxylase activity. The design of catalytic and substrate-binding elements for the new active site and thus, the switch in function has been obtained by the incorporation of four active site loops with different lengths and sequences from IMP-1. Noteworthy is loop 6 with considerable variation among lactamase family members, which contains two amino acid residues (K161 and N167) crucial for binding and activation of the β-lactam substrate during catalysis. Thereafter, additional mutations by directed evolution have been introduced for stabilization of Zn-coordination and increased catalytic activities. This resulted in clone *ev*MBL8 showing the highest resistance against cefotaxime **14** but with no traces of the original glyoxalase activity. Overall, 81 out of 198 amino acid residues, except residues responsible for metal coordination, were changed through this design and evolution process. Furthermore, the novel enzyme variant has acquired novel active site architecture with well-defined metal coordination and substrate-binding pocket for new catalytic activity.

In this light, a recent directed evolution strategy involving random circular permutation was successfully applied to increase the ability of TEM-1 β-lactamase to provide resistance to the antibiotic cefotaxime **14** highlighting the importance of loops for substrate specificity. Two circularly permuted variants were identified that conferred elevated resistance to cefotaxime but decreased resistance to other antibiotics. Very few studies have explored random circular permutation, the intramolecular relocation of the N- and C-termini of a protein, as a diversity-generating step for directed evolution. These variants were circularly permuted in the Ω-loop proximal to the active site. Remarkably, one variant was circularly permuted such that the key catalytic residue Glu166 was located at the N-terminus of the mature protein. This finding indicates that circular permutation altered the substrate specificity of the enzyme.[37]

Ketol-acid reductoisomerases (KARIs) belong to the family of NADPH-dependent oxidoreductases and catalyze the reversible reductive isomerization (alkyl migration) of 2-acetolactate **15** to 2,3-dihydroxyisovalerate **16** (Scheme 7). The laboratory evolution of enzymes having specificity for the NADH cofactor as opposed to NADPH is known for some enzymes and is commonly referred to as "cofactor switching". Recent work on the anaerobic 2-methylpropan-1-ol production in *E. coli* generated a KARI variant containing four mutations showing reversed cofactor specificity for NADH over NADPH.[38] Structural analysis of wild type showed that three of the four mutations (Ala71Ser, Arg76Asp, and Ser78Asp) are located in a loop connecting the β2 sheet and αB helix of the Rossmann fold (β2αB-loop) and are thus, responsible for the distinction between NADPH and NADH-dependent members.[23] These findings have been the starting point for further studies aimed at understanding the key cofactor specificity determinants and the construction of KARIs with reversed cofactor preference. A sequence alignment of 643 KARI sequences revealed significant differences with respect to chain length and amino acid sequence within this loop: the six-residue, the seven-residue and 12-residue loops, respectively. Further loop profile developments and conservation pattern analysis provided the basis for cofactor switching experimental work. As a result, two mutations were necessary to switch cofactor specificity from NADPH to NADH in KARIs with seven-residue and 12-residue β2αB-loops. Amino acid substitutions to Asp resulted in variants with 7300-fold reduced catalytic efficiency for NADPH, whereas the catalytic efficiency for NADH remained the same. However, this was not transferable to the β2αB-loops with six residues. In this light, single- and dual-site saturation mutagenesis libraries at each of the six loop amino acid residues have been created. Two dual variants were obtained with increased specificity for NADH over NADPH, also demonstrating six-residue KARIs as a little more challenging to engineer. Four residues of the six were replaced including the conserved Arg at the beginning of the loop, which forms strong interactions with the 2′-phosphate. In the mutant structure, the β2αB-loop is moved slightly closer toward the cofactor, prereserving the cation-π stacking of the adenine moiety and readjusting the catalytically active nicotinamide moiety of NADH to take on a more favorable position for electron transfer. Finally, variants have been randomly mutated and led to activities comparable to wild type using NADH as cofactor.

Flexible loops in the proximity of active sites frequently play various roles in catalysis, and it is widely accepted that their pronounced conformational flexibility may contribute to their functions. Particular attention is paid to methods such as loop grafting, enzyme engineering, as well as structural and sequential analysis that can reveal more about the mobility of them. This ensemble view is more suitable to shed light on the potential selective role of a loop, as the scale of conformations available to the loop under native conditions may elucidate the ability of the loop to take part in diverse molecular interactions.

■ PROTEIN STABILITY

Evolution has modified the strength and number of stabilizing interactions in enzymes to achieve the optimal balance of stability and flexibility at a given temperature. Different regions of the protein have manifested different changes in mobility in response to temperature adaption. As a result, surface loops showing the largest differences in mobility can undergo large displacements from their native conformations. Such large loop displacements may expose the hydrophobic core of the protein to water penetration, leading to unfolding.[40] As a working hypothesis, it has been assumed that rigidity is a prerequisite for high protein thermostability.[41] This is further supported by flexibility comparisons in mesophilic and thermophilic proteins. Protein stability engineering should therefore not be directed at the protein core but at protein surface loop residues that are typically involved in less intramolecular interactions than internal residues. The engineering of protein stability continues to be of major importance for the use of enzymes in a wide range of industrial applications. While efficient strategies for engineering thermostability by directed evolution[42,43] and rational design[44] techniques have been established over the past decades, it is nevertheless difficult to increase the stability of an enzyme while maintaining or increasing the catalytic

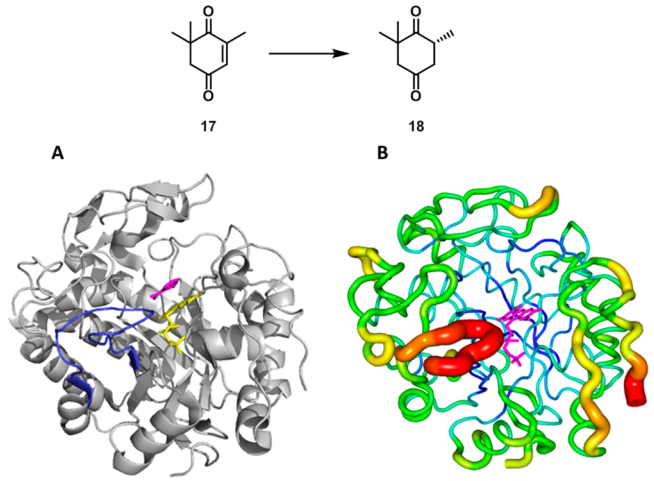
dx.doi.org/10.1021/cs500325p | *ACS Catal.* 2014, 4, 3201−3211

activity.[45] Early work from the group of Frances Arnold demonstrated the generation of a thermostable subtilisin E enzyme through directed evolution. Hereby a relatively small number of eight amino acid substitutions was sufficient to convert the mesophilic *Bacillus subtilis* subtilisin E into an enzyme as stable as its thermophilic counterpart thermitase from *Thermoactinomyces vulgaris*. Most of the thermostabilizing mutations were located on the protein surface in loops connecting elements of secondary structure.[46]

Molecular dynamics simulations including solvation have been employed to study how protein solution structure and dynamics are affected by adaption to high temperature. Therefore, the group of Goddard carried out simulations on the *para*-nitrobenzyl esterase and two thermostable variants that were generated by laboratory evolution.[40] Simulations were validated by examining the fluctuations of W102 and comparing with W102 phosphorescence lifetimes that had been previously measured. The variants display up to 18 °C higher melting temperatures as well as high sequence identity compared to the wild type. The most dramatic changes between the thermophilic mutant 8G8 and wild type have been seen in the regions composed of residues 315−323 undergoing fluctuations of up to 6 Å in 8G8, while it shows little movement in wild type. These fluctuations contribute to entropy and hence to the stability of the native state. The largest changes in localized dynamics occur in surface loops, while other regions, particularly the active site residues, remain essentially unchanged. Several mutations, most notably L313F and H322Y in variant 8G8, are in the region showing the largest increase in fluctuations, suggesting that these mutations confer more flexibility to the loops. The authors conclude that certain modes of motion, particularly those that may initiate unfolding, must be reduced if a protein is to achieve increased stability at elevated temperatures. Furthermore, it has been shown in pioneering work of Nagi and Regan that the loop length is important to successful protein design. Herein, the role of loop length in protein folding and stability with respect to the increase of the loop length of the four-helix bundle protein Rop led to a decreased stability toward thermal and chemical denaturation.[47] The Gilbert group recently reported a directed protein evolution approach to generate hyperthermostability.[48] In this work, Gene Site Saturation Mutagenesis (GSSM) in combination with GeneReassembly technology is employed in the generation of a hyperthermophilic variant of a thermophilic GH11 xylanase. Applying this strategy resulted in an enzyme variant *Ev*Xyn11^{TS}, that was 25 °C more thermostable than the wild-type xylanase. The comparison of structural data for the wild-type and hyperthermostable variant showed that the *Ev*Xyn11^{TS} contains seven mutations (N14H, S9P, T13F, Y18F, Q34L, S71T, S35E). These mutations are all in the N-terminal region of the protein, located on β-strands 2, 3, and 4 and on the loops connecting β-strands 1 and 2 as well as 5 and 6. In particular, one mutation of serine at position 9 to proline in the loop connecting β-strands 1 and 2 resulted in a substantial increase in the T_m (4.7 °C) of the enzyme. From these results, the authors assume that the increased stability displayed by *Ev*Xyn11^{TS} is through additive effects such as more van der Waals contacts from the introduced amino acids. Furthermore, subtle changes in protein structure by the introduction of amino acids can have a substantial impact on thermostability. The mutations did not introduce ion pairs, significantly increased direct hydrogen bonding or enhanced the hydrophobic interactions within the xylanase. The stable variant *Ev*Xyn11^{TS} retained full activity at 90 °C for 60 min using cell lysate.

Recently, the Iwasaki lab reported rational design based on the deletion of a dynamic surface loop of a phospholipase D.[49] This work builds on existing work of the same group, who previously disclosed the identification of amino acid residues affecting the thermostability of the phospholipase D from *Streptomyces antibioticus* (SaPLD) by site-directed combinatorial mutagenesis based on the residue's B-factor. The thermostable variant DYR contains two mutations (Asp40His, Thr291Tyr) of which the residue Asp40 is part of a flexible, highly exposed nine-residue surface loop of the N-terminal half domain.[50] In this study, the trimming of the loop residues (Asp37-Gly45) resulted in a more thermostable variant Δ38−46DYR with an average increase in the enzyme's half-life at 70 °C compared to the DYR. Recently, we also used rational protein engineering to generate thermo- and solvent stable variants of the nicotinamide-dependent cyclohexen-2-one reductase NCR from *Zymomonas mobilis* (Scheme 8).[51] NCR

Scheme 8. Mutagenesis of a Specific Surface Loop Segment near the Active Site of the *Zymomonas mobilis* Ene Reductase NCR Showing an Increased Tolerance to Organic Solvents along with an Enhanced Thermostability in the Reduction of Ketoisophorone 17 to Enantiopure *R*-Levodione 18[a]



[a](A) Loop region L4 (blue, PDB code 4A3U) shown in blue in the native NCR enzyme has been shortened by six amino acids. (B) Prediction of functional flexibility and rigidity of NCR is indicated by the colors: the more red, the higher the B-value (flexible); the more blue, the lower (rigid). The most flexible residues in NCR are shown in red and are located in loop 4 which has been mutated for enhancement in stability. Flavin prosthetic group is shown in yellow (left) and pink (right) sticks, respectively. Ligand nicotinamide is shown in structure A in pink sticks.

ene reductase has been shown to catalyze efficiently the NADH-linked reduction of various α,β-unsaturated alkenes. The reduction of these substrates proceeds in a stepwise fashion. The first step involves hydride transfer from the enzyme-reduced flavin to carbon 2 of the ketoisophorone **17** substrate. The protonation at carbon 1 to form the final reduced product **18** also is catalyzed by the enzyme and involves a tyrosine as an active site proton donor. By combining a systematic analysis of mesophilic and thermophilic reductase sequences and structures and following loop-directed mutagenesis of two selected surface loops adjacent to the active site,

dx.doi.org/10.1021/cs500325p | *ACS Catal.* 2014, 4, 3201−3211

we have identified a surface loop variant which demonstrated improved stability. In addition, highly flexible sites identified on the basis of B-factors have been used in the quest to improve thermostability. In this case, Reetz and co-workers have successfully demonstrated the use of B-factors, atomic displacement parameters available from X-ray data as the basis for decision making (the B-FIT method).[52] The most stable loop variant L4_Short, lacking the insertion region (Glu230-Asp236), showed clearly improved thermostability and robustness toward organic solvents compared with wild-type enzyme. The enzymatic activity of the thermostable L4_Short at 45 °C was 2-fold increased. Biotransformations using varying amounts of acetone and isopropanol in buffer resulted in up to 2-fold higher product yields.

Loops and N and C termini are usually the regions with the highest mobility when crystallizing a protein, and most likely points where denaturation starts. Previously researchers often disregarded loops and termini in having any influence on protein stability. However, through their inherent flexibility, loops can be considered to be potential initiation points for thermal denaturation, and therefore a reduction in their size and thus flexibility may contribute to the protein's kinetic stability. This flexibility may also be reduced by additional stabilizing interactions as well as bonds. Loop anchoring can be achieved in multiple ways, through ion pairing, H bonding, or hydrophobic interactions. From the point of view of designing a thermophilic protein, these studies suggest inclusion of loop regions in the engineering strategy through a better packing, deletion or shortening of loops. Additionally, it indicates including amino acids substitutions within loop segments of enzymes in order to withstand high temperatures.

SUMMARY AND OUTLOOK

In this Viewpoint, we have briefly emphasized the consideration of flexible loops and their role in catalysis through several representative examples of engineering and systematic analysis. Our understanding of the molecular and evolutionary basis for functional innovation is developing. The observation of natural enzymatic activity evolution as for example in the α/β class protein enolase and lipase superfamily helped us recognize loops as modular components of variability.[53] We should take advantage of the knowledge that is stored in different members of enzyme superfamilies that have common ancestors but sometimes rather different catalytic machineries. The alteration and/or improvement of enzyme properties and the generation of new functionalities would benefit from the consideration and insertion of loops as modular components of variability besides point mutations. The importance of loops for enzyme function and stability shows that some researchers in the scientific community have taken notice and begun to realize the potential of loops for catalysis. In this way, it has been possible to generate novel enzymatic function and to alter enzyme specificity as well as stability.

Although, we demonstrated that the consideration of loops in evolutionary approaches might provide a reasonable starting point for the selection of new enzymatic activities, in practice, this methodology is still rarely exploited. In fact, we rely on the expertise and intuition of researchers for conducting, developing, and implementing both more experimental as well as computational and theoretical studies before clear "rules" can be formulated. A growing variety of methods was used and developed to study protein flexibility and mobility; however, most of the studies are presented from a structural point of view with an emphasis on X-ray crystallography results.

Despite some developments recently achieved in this field, the understanding of the role of loops in enzyme catalysis and flexibility as well as the development of new and more powerful enzyme catalysts is still in the early stages. Future developments will require not only a more systematic characterization but also a better understanding of the relationship between loop mobility and its size or exposure in order to design and generate promiscuous and novel enzyme catalysts. We are even tempted to say that loops are enzyme modifying elements according to which nature has probably used loops to evolve different functions. It is clear, however, that the study and understanding of flexible loops associated with catalysis has many opportunities for discovery.

AUTHOR INFORMATION

Corresponding Author

*E-mail: bernhard.hauer@itb.uni-stuttgart.de.

Notes

The authors declare no competing financial interest.

ACKNOWLEDGMENTS

The research leading to these results has received funding from the European Union's Seventh Framework Programme FP7/2007-2013 under grant agreement no. 266025 as well as the Innovative Medicines Initiative Joint Undertaking under grant agreement no. 115360, resources of which are composed of financial contribution from the European Union's Seventh Framework Programme (FP7/2007-2013) and EFPIA companies in kind contribution. This research is part of project CHEM21 (https://www.chem21.eu/). Stephan C. Hammer and Martin Weissenborn are acknowledged for their comments and suggestions for corrections.

REFERENCES

(1) Todd, A. E.; Orengo, C. A.; Thornton, J. M. *Trends Biochem. Sci.* **2002**, *27*, 419−426.
(2) Kokkinidis, M.; Glykos, N. M.; Fadouloglou, V. E. In *Advances in Protein Chemistry and Structural Biology*; Christov, C., Karabencheva-Christova, T., Eds.; Academic Press: Waltham, MA, 2012; Vol. *87*, pp 181−218.
(3) Furnham, N.; Sillitoe, I.; Holliday, G. L.; Cuff, A. L.; Laskowski, R. A.; Orengo, C. A.; Thornton, J. M. *PLoS Comput. Biol.* **2012**, *8*, e1002403.
(4) Hammes-Schiffer, S.; Benkovic, S. J. *Annu. Rev. Biochem.* **2006**, *75*, 519−541.
(5) Osborne, M. J.; Schnell, J.; Benkovic, S. J.; Dyson, H. J.; Wright, P. E. *Biochemistry* **2001**, *40*, 9846−9859.
(6) Bhabha, G.; Lee, J.; Ekiert, D. C.; Gam, J.; Wilson, I. A.; Dyson, H. J.; Benkovic, S. J.; Wright, P. E. *Science* **2011**, *332*, 234−238.
(7) Li, L.; Wright, P. E.; Benkovic, S. J.; Falzone, C. J. *Biochemistry* **1992**, No. 34, 7826−7833.
(8) Venkitakrishnan, R. P.; Zaborowski, E.; McElheny, D.; Benkovic, S. J.; Dyson, H. J.; Wright, P. E. *Biochemistry* **2005**, *44*, 5948−5948.
(9) Chen, S.; Fahmi, N. E.; Wang, L.; Bhattacharya, C.; Benkovic, S. J.; Hecht, S. M. *J. Am. Chem. Soc.* **2013**, *135*, 12924−12927.
(10) Singh, P.; Sen, A.; Francis, K.; Kohen, A. *J. Am. Chem. Soc.* **2014**, *136*, 2575−2582.
(11) Wang, Z.; Singh, P.; Czekster, C. M.; Kohen, A.; Schramm, V. L. *J. Am. Chem. Soc.* **2014**, *136*, 8333−8341.
(12) Falzone, C. J.; Wright, P. E.; Benkovic, S. J. *Biochemistry* **1994**, *33*, 439−442.
(13) Epstein, D. M.; Benkovic, S. J.; Wright, P. E. *Biochemistry* **1995**, *34*, 11037−11048.

(14) Boehr, D. D.; Schnell, J. R.; McElheny, D.; Bae, S.-H.; Duggan, B. M.; Benkovic, S. J.; Dyson, H. J.; Wright, P. E. *Biochemistry* **2013**, *52*, 4605−4619.
(15) Wang, L.; Goodey, N. M.; Benkovic, S. J.; Kohen, A. *Proc. Natl. Acad. Sci. U.S.A.* **2006**, *103*, 15753−15758.
(16) Adamczyk, A. J.; Cao, J.; Kamerlin, S. C. L.; Warshel, A. *Proc. Natl. Acad. Sci. U.S.A.* **2011**, *108*, 14115−14120.
(17) Chesters, C.; Wilding, M.; Goodall, M.; Micklefield, J. *Angew. Chem., Int. Ed.* **2012**, *51*, 4344−4348.
(18) Du, J.; Say, R. F.; Lu, W.; Fuchs, G.; Einsle, O. *Nature* **2011**, *478*, 534−537.
(19) Fushinobu, S.; Nishimasu, H.; Hattori, D.; Song, H.-J.; Wakagi, T. *Nature* **2011**, *478*, 538−541.
(20) Yuan, Z.; Bailey, T. L.; Teasdale, R. D. *Proteins Struct. Funct. Bioinf.* **2005**, *58*, 905−912.
(21) Zimmermann, M. T.; Jernigan, R. L. *Entropy* **2012**, *14*, 687−700.
(22) Grünberg, R.; Leckner, J.; Nilges, M. *Structure* **2004**, *12*, 2125−2136.
(23) Wlodarski, T.; Zagrovic, B. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 19346−19351.
(24) Alvarez-Garcia, D.; Barril, X. *J. Chem. Theory Comput.* **2014**, *10*, 2608−2614.
(25) Tokuriki, N.; Tawfik, D. S. *Science* **2009**, *324*, 203−207.
(26) Gunasekaran, K.; Ma, B.; Nussinov, R. *J. Mol. Biol.* **2003**, *332*, 143−159.
(27) Pries, F.; van den Wijngaard, A. J.; Bos, R.; Pentenga, M.; Janssen, D. B. *J. Biol. Chem.* **1994**, *269*, 17490−4.
(28) Althoff, E. A.; Wang, L.; Jiang, L.; Giger, L.; Lassila, J. K.; Wang, Z.; Smith, M.; Hari, S.; Kast, P.; Herschlag, D.; Hilvert, D.; Baker, D. *Protein Sci.* **2012**, *21*, 717−726.
(29) Wang, L.; Althoff, E. A.; Bolduc, J.; Jiang, L.; Moody, J.; Lassila, J. K.; Giger, L.; Hilvert, D.; Stoddard, B.; Baker, D. *J. Mol. Biol.* **2012**, *415*, 615−625.
(30) Giger, L.; Caner, S.; Obexer, R.; Kast, P.; Baker, D.; Ban, N.; Hilvert, D. *Nature Chem. Biol.* **2013**, *9*, 494−498.
(31) Murphy, P. M.; Bolduc, J. M.; Gallaher, J. L.; Stoddard, B. L.; Baker, D. *Proc. Natl. Acad. Sci. U.S.A.* **2009**, *106*, 9215−9220.
(32) Eiben, C. B.; Siegel, J. B.; Bale, J. B.; Cooper, S.; Khatib, F.; Shen, B. W.; Players, F.; Stoddard, B. L.; Popovic, Z.; Baker, D. *Nat. Biotechnol.* **2012**, *30*, 190−192.
(33) Ochoa-Leyva, A.; Soberon, X.; Sanchez, F.; Argaello, M.; Montero-Moran, G.; Saab-Rincon, G. *J. Mol. Biol.* **2009**, *387*, 949−964.
(34) Ochoa-Leyva, A.; Barona-Gomez, F.; Saab-Rincon, G.; Verdel-Aranda, K.; Sanchez, F.; Soberon, X. *J. Mol. Biol.* **2011**, *411*, 143−157.
(35) De Luca, F.; Benvenuti, M.; Carboni, F.; Pozzi, C.; Rossolini, G. M.; Mangani, S.; Docquier, J.-D. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *108*, 18424−18429.
(36) Park, H.-S.; Nam, S.-H.; Lee, J. K.; Yoon, C. N.; Mannervik, B.; Benkovic, S. J.; Kim, H.-S. *Science* **2006**, *311*, 535−538.
(37) Guntas, G.; Kanwar, M.; Ostermeier, M. *PLoS One* **2012**, *7*, e35998.
(38) Bastian, S.; Liu, X.; Meyerowitz, J. T.; Snow, C. D.; Chen, M. M. Y.; Arnold, F. H. *Metab. Eng.* **2011**, *13*, 345−352.
(39) Brinkman-Chen, S.; Flock, T.; Cahn, J. K. B.; Snow, C. D.; Brustad, E. M.; McIntosh, J. A.; Meinhold, P.; Zhang, L.; Arnold, F. H. *Proc. Natl. Acad. Sci. U.S.A.* **2013**, *110*, 10946−10951.
(40) Wintrode, P. L.; Zhang, D.; Vaidehi, N.; Arnold, F. H.; Goddard, W. A., III *J. Mol. Biol.* **2003**, *327*, 745−757.
(41) Vieille, C.; Zeikus, G. J. *Microbiol. Mol. Biol. Rev.* **2001**, *65*, 1−43.
(42) Eijsink, V. G. H.; Gaseidnes, S.; Borchert, T. V.; van den Burg, B. *Biomol. Eng.* **2005**, *22*, 21−30.
(43) Farinas, E. T.; Bulter, T.; Arnold, F. H. *Curr. Opin. Biotechnol.* **2001**, *12*, 545−551.
(44) Bloom, J. D.; Wilke, C. O.; Arnold, F. H.; Adami, C. *Biophys. J.* **2004**, *86*, 2758−2764.
(45) Polizzi, K. M.; Bommarius, A. S.; Broering, J. M.; Chaparro-Riggers, J. F. *Curr. Opin. Chem. Biol.* **2007**, *11*, 220−225.
(46) Zhao, H.; Arnold, F. H. *Protein Eng.* **1999**, *12*, 47−53.
(47) Nagi, A. D.; Regan, L. *Folding Des.* **1997**, *2*, 67−75.
(48) Dumon, C.; Varvak, A.; Wall, M. A.; Flint, J. E.; Lewis, R. J.; Lakey, J. H.; Morland, C.; Luginbühl, P.; Healey, S.; Todaro, T.; DeSantis, G.; Sun, M.; Parra-Gessert, L.; Tan, X.; Weiner, D. P.; Gilbert, H. J. *J. Biol. Chem.* **2008**, *283*, 22557−22564.
(49) Damnjanovic, J.; Nakano, H.; Iwasaki, Y. *Biotechnol. Bioeng.* **2013**, *111*, 674−682.
(50) Damnjanovic, J.; Takahashi, R.; Suzuki, A.; Nakano, H.; Iwasaki, Y. *Protein Eng. Des. Sel.* **2012**, *25*, 415−424.
(51) Reich, S.; Kress, N.; Nestl, B. M.; Hauer, B. *J. Struct. Biol.* **2013**, *185*, 228−233.
(52) Reetz, M. T.; Carballeira, J. D.; Vogel, A. *Angew. Chem., Int. Ed.* **2006**, *45*, 7745−7751.
(53) Gerlt, J. A.; Babbitt, P. C. *Curr. Opin. Chem. Biol.* **1998**, *2*, 607−612.