DESIGNING ENZYMES FOR ORGANIC SYNTHESIS: ENZYME MIMICS AND DIRECTED EVOLUTION

Reported by David C. McKinney September 13, 2001

INTRODUCTION

The need for enantiopure synthons in academic and pharmaceutical research has been the motivation behind much of modern synthetic chemistry. The generation of single enantiomers from prochiral compounds requires the use of either chiral catalysts or stoichometric chiral reagents. While there are many small molecule chiral catalysts available, they are not available for all desired transformations. Many of these catalysts also often have major drawbacks, such as low selectivity and toxic or expensive components. Enzymatic catalysts are an interesting alternative to small molecule chiral catalysts, with certain attractive features. High turnover, high enantioselectivity, high specificity, and relatively low toxicity characterize most enzymes, and enzyme specificity often allows for the modification of a specific functional group without requiring the protection of similar groups in the substrate molecule. This specificity, however, can also be a problem. Very often enzymes are available that perform the desired class of reaction on specific substrates, but they lack the generality required to utilize the broader range of substrates that might be of interest to synthetic chemists. Conversely, some enzymes are generic enough to accept unnatural substrates, but do so with poor enantioselectivty. Directed evolution has been applied to improve these and other properties of enzymes, so as to make them more useful for organic synthesis in both laboratory and industrial settings. The development of enzyme mimics is a very different chemical approach to developing chiral catalysts. Mimics are inspired by some of the features of enzymes, but are constituted of non-peptidic fragments. This review covers both enzyme mimics and directed enzyme evolution.

Background

Since it was first recognized that organic reactions could be catalyzed by enzymes, such proteins have been used in many ways for organic synthesis. From simple yeast reductions of prochiral compounds, kinetic resolutions, and desymmetrization of meso compounds, to the more complicated catalysis of aldol condensations, enzymes have found increasing utility in the past years. Enzymes can act either in solution, with cofactors being added in some cases, or, alternatively, whole-cell cultures can be used in the reaction. In this latter case, cells are often designed to secrete all the needed enzymes and cofactors into the reaction media. While in both cases product isolation is not always trivial, in this era of “green” chemistry, aqueous reactions have potential environmental advantages.

In recent years, several modifications to traditional enzymatic reactions have been made to increase their utility greatly. Mutant enzymes that tolerate, or even thrive in, organic solvents, and aqueous/organic solvent mixtures have improved the activity and in some cases increased the selectivity of enzymes for use with non-water-soluble substrates. This is a significant improvement because many substrates of interest are not water soluble, and undesirable synthetic steps are added by the alternate approach of attaching and removing water-solubilizing groups. Cyclodextrins have increased the utility of enzymatic reactions by acting as solubilizing agents. They often behave similarly to phase-transfer catalysts, and help to create high concentrations of substrates near the enzymes, dramatically increasing the observed reaction rates. These benefits have been observed by simply adding the cyclodextrins to the reaction, as well as by covalently attaching them to enzymes.[1]

These modifications to enzyme reactions only take an existing enzyme and make it more useful; they do, however, require that the enzyme already exists. When no enzyme exists that performs the reaction of interest, one may turn to designing an enzyme from existing templates or the more difficult path of *de novo* design of purely chemical systems that function as enzyme mimics. Enzymes are polypeptides composed of complicated tertiary structures that support their selectivity and reactivity. It is known that enzymes contain both active catalytic sites and substrate recognition regions, but they also contain what seems to be unimportant "bulk" with properties that are not fully understood. This suggests that it might be possible to remove much of the bulk of an enzyme to provide a minimal enzyme, which nevertheless retains the active site and the recognition regions. Such enzyme mimics are much simpler than natural enzymes, and they have been useful tools for the understanding enzymatic reactions. To date, however, the utility of enzyme mimics in performing selective, asymmetric organic reactions has been rather limited.

ENZYME MIMICS

Cyclophanes have been used as enzyme mimics because they can be constructed in a modular fashion and as a consequence their geometric and electronic properties can be easily controlled. Diederich and coworkers have designed a cyclophane-based pyruvate oxidase to accomplish the electrochemical oxidation of aromatic aldehydes.[2] Direct electrochemical oxidation of aldehydes is often not possible due to an unfavorable overpotential. Pyruvate oxidase catalyzes this oxidation by the addition of thiazolium **1** into the aldehyde carbonyl of **5**, tautomerization to the enol **3,** subsequent oxidation by a flavin co-enzyme, and finally nucleophilic substitution at the resulting activated acid **4**, releasing the oxidized aldehyde **6** (Scheme 1).[3] By attaching both the thiazole nucleophile and the flavin oxidant to cyclophane **7** (Figure 1), these investigators were able to control the microenvironment

around the substrate so as to promote oxidation of aromatic aldehydes. The flavin is easily re-oxidized by a low voltage applied to the system resulting in a truly catalytic system. This enzyme mimic has allowed the electrochemical oxidation of aldehydes into esters and acids in high yields.

Scheme 1. Pyruvate Oxidase Catalytic Cycle

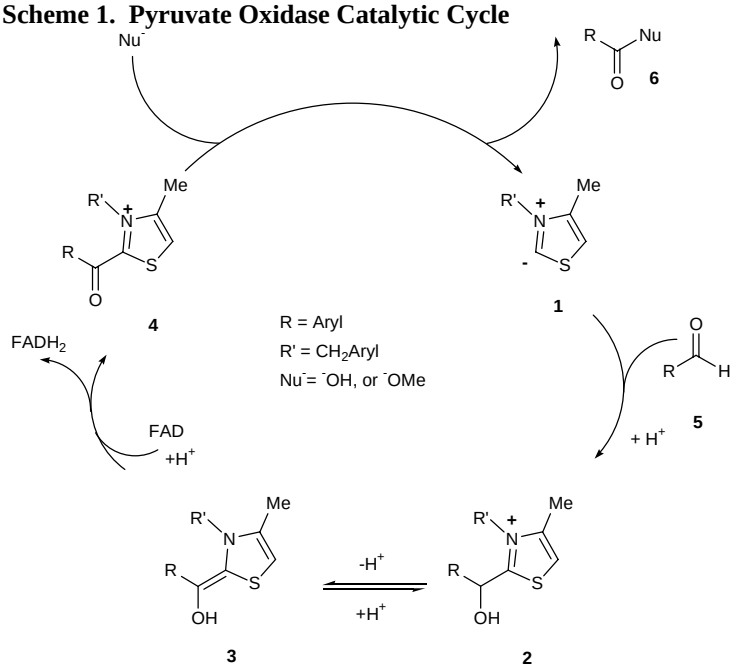
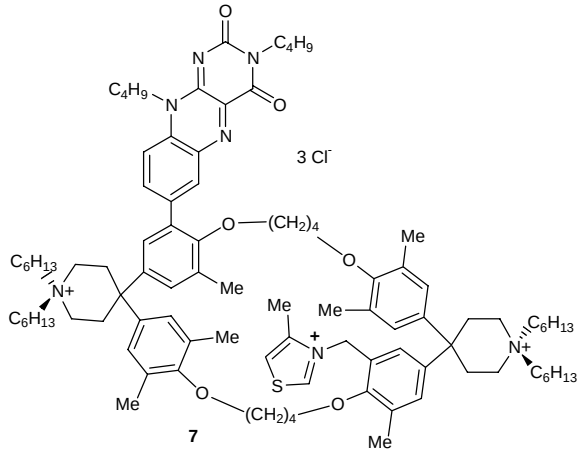
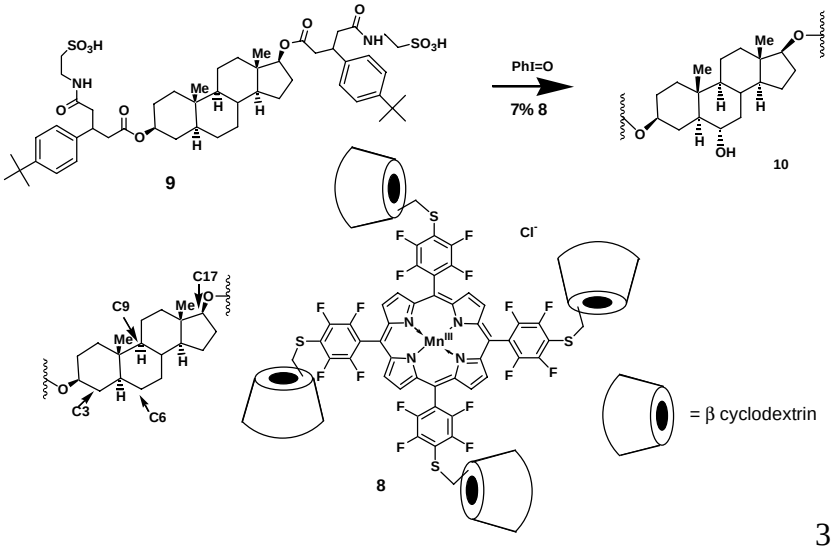


Figure 1. Cyclophane Pyruvate Oxidase Mimic



A great deal of work has been directed at developing an effective mimic of a cytochrome P-450. These enzymes are the oxidative workhorse of many biotic systems and are capable of hydroxylating aliphatic hydrocarbons. By attaching cyclodextrins as directing groups to the manganese porphyrin **8**, Breslow and co-workers achieved specific hydroxylation at C6 of androstane backbone **9;** this regioselectivity is accomplished by having the C6 fortuitously positioned directly over the active manganese center (Scheme 2).[4] Esterification at C6 with the same directing group found on C3 and C17 allows for selective oxidation at C9 under similar conditions. Unfortunately, several directing groups must be attached and eventually removed for this enzyme mimic to work satisfactorily.[5] Despite these drawbacks the selectivity of this mimic is impressive.

Scheme 2. Cytochrome P-450 Mimic



One problem inherent with enzyme mimics is their impracticality for industrial-scale use; they typically are useful only in laboratory settings, where costly and laboriously made

catalysts are acceptable. While they are very useful tools for the study of active sites, at this point in time, our understanding of structure and function is such that enzyme mimics are not easily adjusted to suit new industrial needs. Fortuitously, Breslow and co-workers were able to selectively oxidize androstane **9**. Other substrates, however, have not shown the same selectivity. These limitations with enzyme mimics have stimulated the investigation of the use of altered enzymes for industrial applications. Industrially, the use of enzymes themselves is much more advantageous, primarily due to the low cost of both purchasing and disposal of the catalyst.

Subtilisin is an aggressive protease that is widely used in the formulation of modern laundry detergents, and over the last twenty years the utility of subtilisin has been greatly improved by random mutations that enhance its theromstability and catalytic activity.[6] With the realization that mutations in the peptide structure can greatly improve stability and affect reactivity in ways that are not easily explained, several groups have decided to take a more radical route and randomly mutate the enzymes in the hopes of improving other properties, such as substrate recognition and enantioselectivity.[7]

DIRECTED EVOLUTION

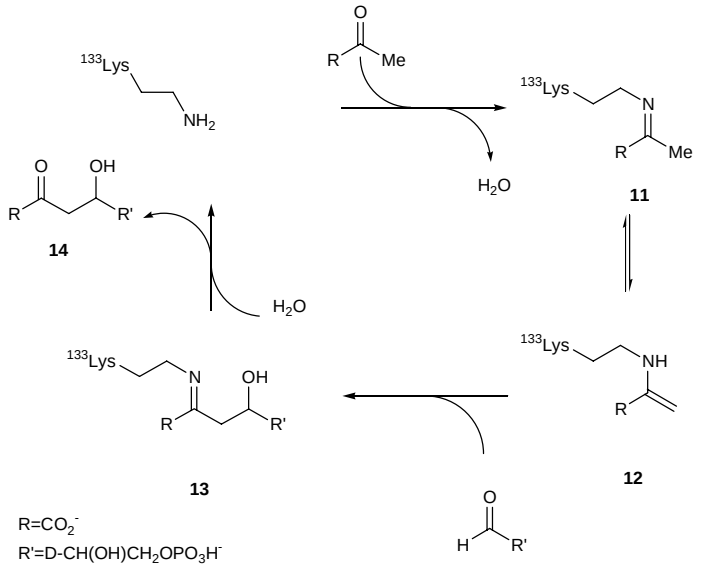
Directed evolution is a process by which mutations are first introduced into the genes that encode an enzyme of interest. Typically these mutations are introduced by the error-prone polymerase chain reaction (EpPCR), DNA shuffling, or a staggered extension process (StEP) such that each enzyme mutant has only one or two amino acid substitutions. Selection pressures for the desired property are then applied, and the positive mutants are carried on into the next round of mutation and selection. UV absorbance, color, IR thermography, ESI-MS, and other methods have been used for the quantification of enantioselecivity, activity, stability and other properties. With an appropriate selection scheme which allows for the evaluation of the library of mutants generated, directed evolution can yield enzymes with desired properties, starting from enzymes that are, synthetically, not particularly useful.

It is important to note that EpPCR alone will not generate all possible mutants of a given enzyme. First, only mutations that arise from a single nucleotide change in an existing codon are produced. Second, a complete mutant library for any reasonably sized enzyme would be greater than the mass of the universe. Statistically, however, each amino acid will be changed at least once. A two-staged process has been shown to improve the results of EpPCR. Initially, important residue locations are located by EpPCR, subsequent testing of all amino acids at these sites through saturation mutagenesis results in enzymes containing the optimal residue at these locations.

Evolving Type One Aldolases for Organic Synthesis

The aldol reaction is one of the most common carbon-carbon bond forming reactions found in biochemistry. This reversible process is found in sugar biosynthesis as well as catabolism. *E. coli* D-2-keto-3-deoxy-6-phosphogluconate (KDPG) aldolase catalyzes the addition of pyruvate to D-glyceraldehyde phosphate in a highly diastereoselective manner. While this reaction is useful, phosphorylated aldehydes are not trivial to synthesize, and a generic system that could also make L-sugars is highly desirable. The crystal structure of KDPG aldolase with bound pyruvate has recently been solved, confirming the identity of the active-site lysine in this enzyme, which acts though a well-known mechanism (Scheme 3).[8] Imine **11** is formed by the condensation of pyruvate and the active site lysine at position 133, isomerization results in the enamine donor **12**. Addition into the acceptor glyceraldehyde results in the imine adduct **13**. Hydrolysis gives KDPG **14** and regenerates the enzyme.

Scheme 3. KDPG Aldolase Catalysed Aldol Reaction



Wong and co-workers have used directed evolution to enhance the activity of KDPG aldolase towards non-phoshorylated aldehydes with some success.[9] The wild-type enzyme shows a selectivity of 50,000:1 for phosphorylated glyceraldehyde over the non-phosphorylated analogue. Replacement of an active-site glutamate with a valine resulted in a 20-fold decrease in activity for phosphorylated aldehydes and a modest two-fold increase in activity for non-phosphorylated aldehydes. This glutamate presumably forms a salt bridge with the phosphate group. Replacement of a glycine proximal to the phosphate-binding glutamate with a serine coupled with a distant threonine-to-isoleucine transformation resulted in an enzyme with a selectivity of only 85:1 for phosphorylated substrates. While the enzyme still prefers the phosphorylated analog, it is less selective for the natural substrate by more than two orders of magnitude. It is clear that while "rational" design would target the active site, directed evolution via EpPCR has uncovered several beneficial mutations whose actions are not yet obvious, but have allowed for the use of KDPG aldolase to be extended to non-phosphorylated substrates, as well as unnatural L-aldehydes. In the latter case, L-sugars are produced with a *threo:erythro* ratio of 95:5.

Wymer and co-workers have also broadened the specificity of KDPG aldolase to include not only pyridine carboxaldehyde but also benzaldehyde.[10] Directed evolution generated several mutants

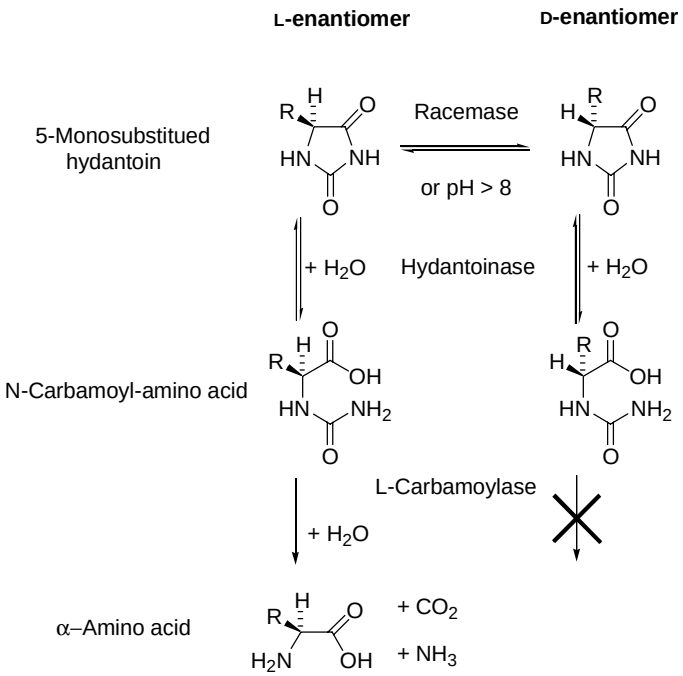
with increased activity towards pyruvate and pyridine carboxaldehyde. One of these contained an asparagine to serine mutation, and its increased activity presumably derives from a change in hydrogen bonding partners. Mutation of a tyrosine to a lysine also resulted in an increased activity. This tyrosine location was also found to be a "hot-spot" by Wong, where an alanine substitution resulted in lower selectivity for the phosphorylated version of the acceptor.[9]

Wymer and co-workers noted this lysine was less than 5 Å from the native lysine and presumably acted as an alternate imine-forming base. A double mutant containing both the new lysine and a glutamate in place of the native lysine was prepared. This modified enzyme was found to have increased activity towards pyruvate/benzaldehyde, as well as 2-ketobuturate/benzaldehyde. The increased activity is likely due to the relocation of the imine towards a hydrophobic pocket in the enzyme. While the new enzyme is significantly less enantioselective, giving products with an enantiomeric excess of 32%, it is more tolerant of unnatural substrates. A loss in enantioselectivity is not surprising as this attribute was not selected for during their directed evolution and there is not necessarily any correlation between activity with a new substrate and enantioselectivity of the new product. These results show that a mixture of random mutagenesis-based discovery, coupled with rational modification of the resulting mutants, is an approach that is capable of increasing the scope of available enzymatic reactions, and is perhaps the future trend in directed evolution.

Enantioselectivity by Directed Evolution

Using an *E. coli* expression vector, Arnold and co-workers have transformed the D-selective hydantoinase from *Arthrobacter sp.* DSM 9771 into an L-selective enzyme with increased activity.[11] 5-Monosubstituted hydantoins are common industrial starting materials for both D and L amino acids. Incubation with whole cells and their secreted enzymes results in racemization, hydration, and then selective hydrolysis of the desired enantiomer (Scheme 4). Often a problem is that the hydantoinase enzyme is not L-selective for the desired substrate. For unnatural amino acids, this is not surprising, but what is perhaps difficult to understand is the low selectivity for L-5-(2-methyl-thio-ethyl)hydantoin, the precursor of L-methionine. The native

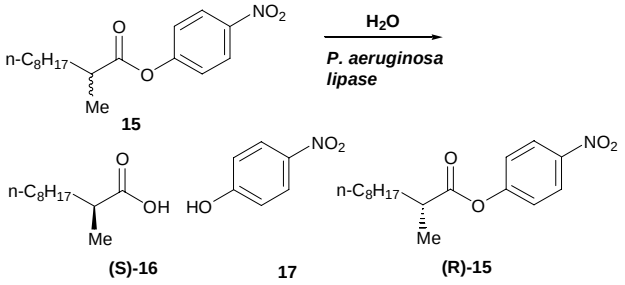
Scheme 4. Amino Acid Synthesis from Hydanoins



hydantoinase is far more selective in hydrating the D-5-monosubstituted hydantoin, which results in a build up of the D-N-carbamoyl-amino acid. The mechanism for this reaction is not well known, and a crystal structure has yet to be solved. Even without knowing the mechanism of action, Arnold and co-workers, used directed evolution to generate several mutants with five-fold increased activity. They found that one conservative valine-to-alanine mutation was responsible for most of this increase. Converting a different valine to alanine increased the D-selectivity by 50%. A less conservative leucine-to-phenylalanine mutation resulted in a slight L-selectivity. This mutated hydantoinase prevents the accumulation of D-carbamoyl amino acids, resulting in an eighty percent reduction in reaction time for the synthesis of L-methionine. What is perhaps more significant than the fact that these mutations enhance both activity and selectivity, is that a single point mutation is sufficient for this. However, this should not be surprising, because a change in enantiomeric excess in the product from 0% to 99% results from changing the free energy of activation by only 3 kcal mol^{-1}, which is on the order of the strength of a single hydrogen bond. This whole-cell example shows that the selectivity of a process can be tremendously affected by single mutations of amino acids, even when the nature of the catalytic site is not known in detail.

Reetz and co-workers used directed evolution of a lipase from *Pseudomonas aeruginosa* to increase its kinetic resolving ability from an enantioselectivity factor of 1.1 to a factor of 11.3.[12] This huge increase was observed after only four generations of mutation. A fifth generation of EpPCR, when coupled with saturation mutagenisis at hotspots, produced an enzyme with an enantioselectivity factor of 25.8, corresponding to an ee of 91% at 18% conversion. Thus, an enzyme derived from the mutagenesis of an enzyme with essentially no enantioselectivity afforded the kinetic resolution of *p*-nitropheny-2-methyl-decanoate **15** (Scheme 5).

Scheme 5. Kinetic Resolution of a *p*-Nitrophenyl Ester



The mechanism by which these mutations increase enantioselectivity is not immediately obvious as only one substitution, that of a serine to a phenylalanine, is located in the active site. This phenylalanine likely interferes with the alcohol portion of the ester by forcing a nearby residue to selectively block one face of the ester, thereby discriminating between enantiomers of the substrate. The remaining mutations are found on the exterior of the enzyme and replace three serines with a valine and two glycines. This removes much of the rigidity in the enzyme structure. Modeling was used to illustrate that this increased flexibility may move a leucine closer into the active site, which could provide the observed enantioselctivity. These subtle changes in the conformation of the active site, such

as the relocation of a leucine that result from these mutations, would be difficult to predict *a prioi* from our knowledge of protein folding. This work also showed that ethyl esters tend to show similar selectivity as *p*-nitrophenyl esters.[13] Kinetic resolution has hereby been afforded for ethyl esters by directed evolution of a natural enzyme that, itself, showed very little selectivity.

CONCLUSION

Enzyme mimics provide an attractive alternative to enzymes due to their relative simplicity. They do however lose much of the selectivity inherent to enzymes and as a result have limited practical utility. Directed evolution has been used to increase substrate specificity towards non-natural substrates, increase activity, and improve the enantioselectivity of natural enzymes. The resulting mutated enzymes have beneficial mutations that are often not located in or near the active site of the enzyme. This highlights the fact that subtle changes in the active site conformation, resulting from substitutions far removed from the active site, can result in dramatic changes in enzyme activity. This is significant because site directed mutagenesis is usually targeted at the active site of enzymes and would not discover these beneficial mutations.

REFERENCES

(1) Harper, J. B.; Easton, C. J.; Lincoln, S. F. *Curr. Org. Chem.* 2000, *4*, 429-454.
(2) Mattei, P.; Diederich, F. *Helv. Chim. Acta* 1997, *80*, 1555-1588.
(3) Tamchang, S. W.; Jimenez, L.; Diederich, F. *Helvetica Chimica Acta* 1993, *76*, 2616-2639.
(4) Yang, J.; Breslow, R. *Angew. Chem.-Int. Edit.* 2000, *39*, 2692-2694.
(5) Breslow, R.; Huang, Y.; Zhang, X. J.; Yang, J. *Proc. Natl. Acad. Sci. U. S. A.* 1997, *94*, 11156-11158.
(6) Bryan, P. N. Biochim. Biophys. Acta-Protein Struct. Molec. Enzym. 2000, 1543, 203-222.
(7) Arnold, F. H.; Volkov, A. A. Current Opinion in Chemical Biology 1999, 3, 54-59.
(8) Machajewski, T. D.; Wong, C. H. *Angew. Chem.-Int. Edit.* 2000, *39*, 1352-1374.
(9) Fong, S.; Machajewski, T. D.; Mak, C. C.; Wong, C.-H. *Chem. Biol.* 2000, *7*, 873-883.
(10) 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*, 1-9.
(11) May, O.; Nguyen, P. T.; Arnold, F. H. *Nat. Biotechnol.* 2000, *18*, 317-320.
(12) Reetz, M. T.; Jaeger, K. E. *CHEMISTRY A European Journal* 2000, *6*, 407-412.
(13) Liebeton, K.; Zonta, A.; Schimossek, K.; Nardini, M.; Lang, D.; Dijkstra, B. W.; Reetz, M. T.; Jaeger, K. E. *Chem. Biol.* 2000, *7*, 709-718.