IMPROVING ENANTIOSELECTIVITY OF ENZYMES THROUGH DIRECTED EVOLUTION

Reported by Brandon Lange   April 17, 2006

INTRODUCTION

Chirality, the molecular property of being non-superimposable on the mirror image, is an important aspect in organic chemistry. In many pharmaceutical drugs, only one enantiomer is effective in treatment. A classic example is thalidomide, a drug once prescribed for morning sickness. The (*R*)-enantiomer was effective, but the (*S*)-enantiomer caused birth defects. It has become of great interest to the scientific community to have reactions specific in only generating one enantiomer in a cost-effective way. Companies spend billions of dollars annually in order to improve methods of synthesizing drugs.[1]

Developing chiral catalysts to effect enantioselective reactions is a common goal of organic chemistry research. Although many catalytic methods are commonly employed in research laboratories, transition metal catalysts[2] are predominant in process development in the pharmaceutical industry. However, this development process may be rather time-consuming and require the testing of many variations in order to discover a catalyst that can efficiently and effectively produce high quantities of enantiopure material.

In addition to transition metal catalysts, enzyme catalysts[3] can be used. Even though nature has engineered these proteins to catalyze certain reactions, they usually have limited substrate scope. The tertiary structure of the protein allows for certain binding pockets that induce a natural selectivity towards only certain substrates and products. By employing the process of directed evolution to change the structure of the enzyme, researchers are able to increase the natural selectivity for only certain enantiomers to bind the substrate. The aim of this review is to show enzyme catalysis optimized by directed evolution can be a viable alternative for effecting enantioselective organic transformations.

DIRECTED EVOLUTION

Evolution is the result of genetic change over a period of time. Scientists have learned to accelerate the evolutionary process by genetically modifying one or more of the four natural bases in DNA. This is known as directed evolution. Modification of the DNA causes the gene to express a different protein during protein synthesis. After expression the bacterial colonies are isolated by a colony picker and placed in microtiter plates containing nutrient broth. Only one modified species is placed in each well. Cell lysis is usually induced, and a colorimetric or other assay is used to determine which plates contain mutants that display an enhanced enantioselectivity in a given reaction. The active mutants are chosen, and further rounds of directed evolution are performed (Figure 1).
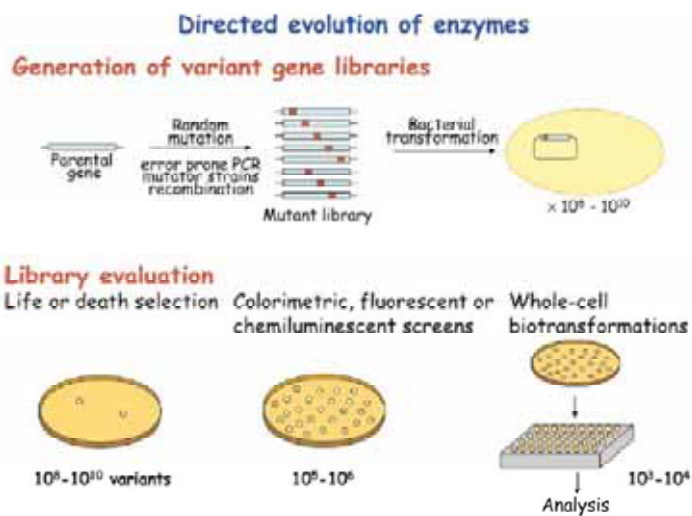
Figure 1. Laboratory representation of how directed evolution is used to generate new mutants

It may appear that all variants of amino acid exchanges are not possible, owing to the redundancy of the genetic code, in which multiple codons use the same code for protein synthesis. For a protein containing 300 amino acids, a theoretical number of library variants N can be calculated using the equation $N = 19^M * 300!/ [(300-M)! * M!]$, with M being the number of amino acid substitutions per enzyme molecule. With $M = 1$, there will be 5700 variants. Increasing the number of substitutions to two would create over 16 million variants, and to three would be over 60 billion variants. Clearly systematic generation and analysis of all variants is not feasible, even with today's high-throughput methods.[4]
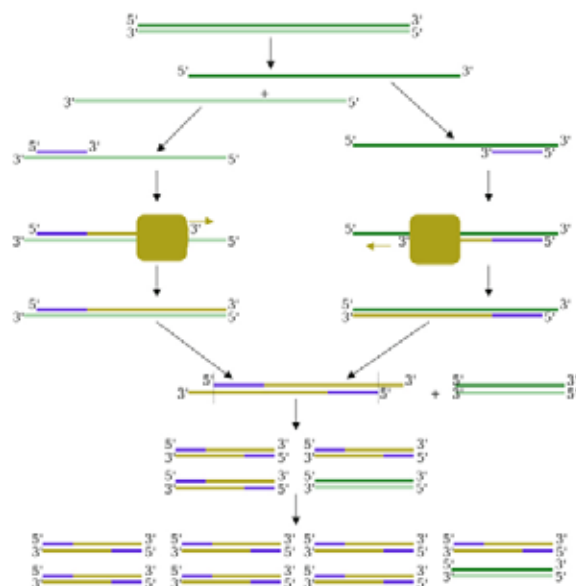
METHODS OF DIRECTED EVOLUTION

Figure 2. Polymerase chain reaction. By introducing MnCl₂, errors are introduced into the polymerase reaction, creating mutants that may have better enantioselectivity.

Many practical methods to accomplish directed evolution have been reported, including error prone polymerase chain reaction (ep-PCR), site-specific mutagenesis, and DNA shuffling. They all have positive and negative attributes. Some are able to access certain substitutions whereas others are not. This can be crucial, especially when trying to effect mutations near the active site of the enzyme.

Perhaps the most widely-used method of creating new enantioselective enzymes is ep-PCR.[5] In this technique, DNA replication (Figure 2) is employed under different conditions. Mutations are incorporated into the DNA fragments by using the low fidelity DNA polymerase *Thermus aquaticus* (*Taq*). *Taq* is used in PCR because it is stable around 95°C, the temperature at which DNA undergoes thermal denaturation. The low fidelity of *Taq* allows more errors to be incorporated during replication.

Leung and coworkers devised five modifications to the normal PCR conditions to make the reaction more error prone. These include increasing the concentration of *Taq* in the reaction mixture,

increasing extension time, increasing the concentration of $MgCl_2$, adding 0.5 mM $MnCl_2$, and increasing the concentrations of dGTP, dCTP, and dTTP to 1 mM in addition to the standard 0.2 mM concentration of dATP. Joyce and coworkers found that standard PCR without modifications leads to only one mutation in 3177 nucleotides that were sequenced, an error rate of <0.14% per position[6]. Using Leung's procedure, they found an overall mutation rate of 1.37% (± 0.29%) per position per PCR, a rate sufficient to generate at least one amino acid modification in an enzyme.
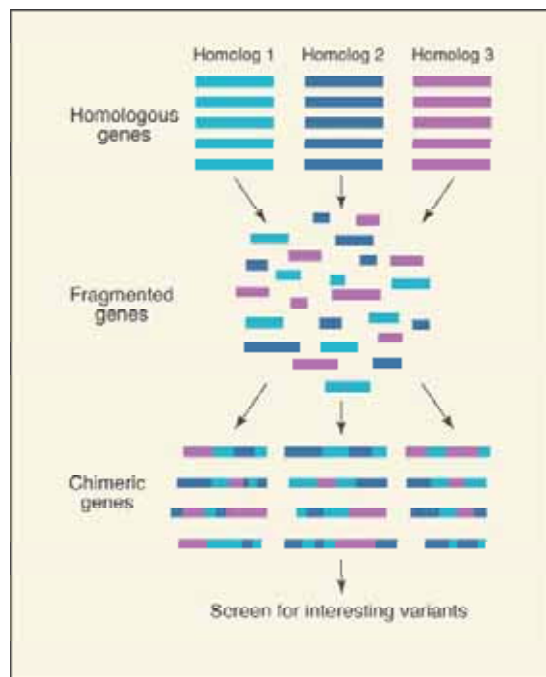

Figure 3. DNA shuffling allows genes of many different mutations to be shuffled to have the potential of having better selective variants.

Another method of directed evolution is site-directed mutagenesis. In this method, specific amino acid substitutions are mapped to any position. For example, if residue 192 on a protein is of particular interest, it could be easily changed to the other 19 amino acids by using the PCR overlap extension method.[7] The site-directed mutagenesis codon consists of an NNS codon (N = a mixture of A + C +T + G; S = a mixture of C and G only). This random assortment allows access to all 20 amino acids and the stop codon. The method is advantageous over ep-PCR because all amino acid substitutions for one particular residue can be generated. A new method, developed by Diversa Corporation, is called Gene Site Saturation Mutagenesis(GSSM).[8] This is similar to the site-directed mutagenesis, but no prior knowledge of the protein tertiary structure is needed, due to its ability to generate all mutations at all positions within the protein.

DNA shuffling makes use of fragmented genes of related species (Figure 3).[9] PCR reassembles these fragments together, creating new genes containing pieces of the old ones. Combinatorial multiple-cassette mutagenesis (CMCM) allows a wild-type gene and cassettes of defined sequences to be randomized, enabling incorporation of many more sequences.[10]

ENANTIOSELECTIVE HYDROLASES FOR ESTERS

The first instance of implementing directed evolution to improve enantioselectivity was reported by Reetz and coworkers who studied a lipase from *Pseudomonas aeruginosa* and the hydrolysis of racemic *p*-nitrophenyl 2-methyldecanoate.[4] The wild type enzyme showed an enantioselectivity of 2% ee for the (*S*) enantiomer. ep-PCR was chosen as the mutagenesis method to enhance enantioselectivity and the 933 base pair gene was mutated. The hydrolysis of e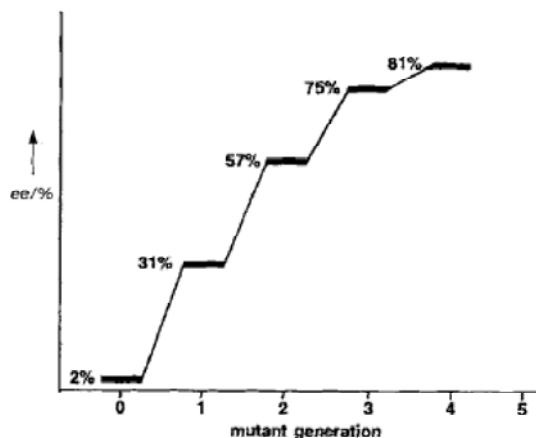ach enantiomer was monitored by measuring the absorption of the *p*-nitrophenolate anion at 410 nm as a function of time. After the first round of mutagenesis, a mutant was found with an improved ee of 31%. The clone that had the greatest selectivity was chosen for the next round of mutagenesis. After four rounds of mutation, the enantioselectivity increased to 81% ee (Figure 4). The selectivity factor *E*, the relative rate of the reactions of the (*S*) and (*R*) substrates, was 11.

Figure 4. Subsequent rounds of ep-PCR led to variants that were more enantioselective. The most enantioselective variant was chosen for further rounds of evolution.

Many other methods of mutagenesis were employed by Reetz and coworkers to improve the enantioselectivity, including saturation mutagenesis at "hot" spots of the improved enzyme variants. CMCM showed the best results by generating a mutant that had an *E* value of 51 (ee >95% at 24% conversion). It is interesting to note that in most cases the Leu162Gly mutation was present. The authors speculated that this side chain could have a large influence because it might directly interact with the methyl group of the (*S*)-substrate. The addition of glycine residues introduces more flexibility into the enzyme, resulting in improved enantioselectivity.[11]

Although the lipase mutants were optimized for the (*S*)-enantiomer, it was noticed that some of the variants were selective for the (*R*)-enantiomer. After further rounds of mutagenesis and DNA shuffling, an (*R*)-selective mutant was obtained with an *E*=30.[12] Thus both enantiomers may be accessible if enough directed evolution is applied to the genes of interest.
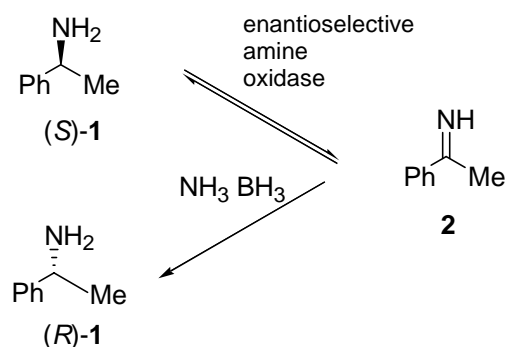
ENANTIOMERICALLY PURE AMINES

Turner and coworkers[13] evolved an amine oxidase from *Aspergillus niger* to generate an enantioselective variant through many rounds of directed evolution. These variants exhibited a higher catalytic activity and have greater substrate specificity toward chiral primary and secondary amines. α-Methylbenzylamine (AMBA) was used as the model substrate to test the ability of directed evolution
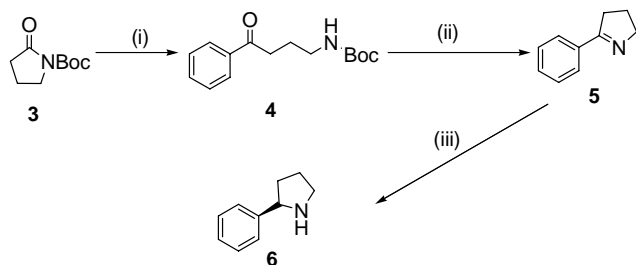
towards this enzyme.

New enzyme variants were assayed by the natural byproduct of the oxidation of amine to the imine, hydrogen peroxide. The use of a peroxidase with 3,3'-diaminobenzidine as its substrate gave rise to a dark pink color that identified active colonies. A library was generated by many cycles of mutagenesis using the *E. coli* XL1-Red mutator strain. These mutants were transferred to an agar plate and grown. There were ~3000 colonies per plate, and this experiment was run using both the (*S*)- and (*R*)-AMBA in order to assay activity for both enantiomers. Of the approximate 150,000 clones, 35 were identified with improved activity towards (*S*)-AMBA compared to that of the wild-type enzyme.

The most selective enzyme contained only a single amino acid exchange from the wild-type enzyme (Asn336Ser). Not only did the mutant possess a significantly higher selectivity (5.8-fold more selective for (*S*) than (*R*)), but it also had a 47-fold increase in catalytic activity as compared to the wild-type enzyme. A de-racemization experiment was conducted to test the enzyme's effectiveness (Scheme 1). An amine-borane complex was chosen as the reducing reagent. Racemic AMBA at 1mM concentration gave (*R*)-AMBA in 93% ee with 77% yield. The optical purity could be increased to 99% ee, but a decrease in yield was observed. An interesting application was the transformation of pure (*S*)-AMBA into (*R*)-AMBA (18% yield, 99% ee). However, when starting with pure (*R*)-AMBA under identical conditions, no (*S*)-AMBA was formed. The synthesis of (*R*)-2-phenylpyrrolidine was carried out in a rather elegant fashion (Scheme 2). Deprotection of the *N*-Boc-ketone and in situ reduction with an ammonia borane complex in the presence of MAO-N-5 variant was completed in 24 hours with 99% ee. This process allows the synthesis of enantiomerically pure secondary amines from amino ketone precursors.[14]

Scheme 1. Deracemization of AMBA to afford only the (*R*)-enantiomer

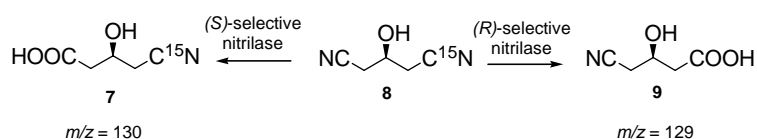**Scheme 2.** Synthesis of Enantiopure Amine

Reagents and Conditions: (i) PhMgBr, THF, -78°C; (ii) TFA, 1.5h; (iii) 10% w/v whole cells (*E. coli*) expressing MAO-N-5 variant, NH₃BH₃ (10 eq), *Ki* buffer (pH 7).

ENANTIOSELECTIVE NITRILASES

In an effort to develop a more efficient synthesis of (*R*)-4-cyano-3-hydroxybutyric acid, a key intermediate for

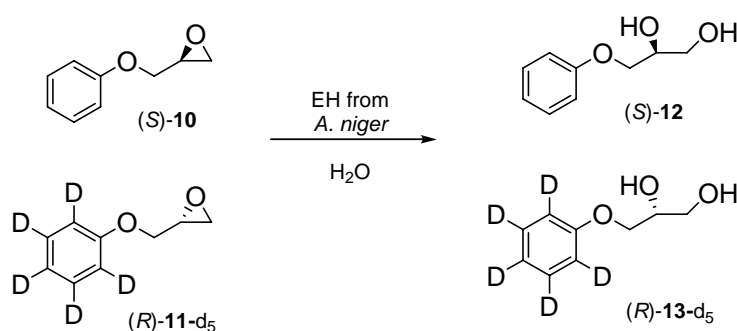**Scheme 3.** Selective Oxidation of Nitriles

the synthesis of the cholesterol lowering drug Lipitor, Burk et al applied GSSM to a nitrilase.[15] This nitrilase effectively transforms 3-hydroxyglutaronitrile into the *R* acid of interest. However, the wild-type enzyme was only efficient at low substrate concentration (100 mM). As concentration increased from 100 mM to 0.5, 1, 2, and 3 M, the enantiomeric excess decreased to 92.1, 90.7, 89.2, and 87.6% respectively. GSSM created 31,584 clones that were screened to identify the mutants with improved enantioselectivity at higher substrate concentrations. A novel high-throughput assay was developed using isotopically labeled (*R*)-3-hydroxyglutaronitrile in order to screen this large library (Scheme 3). An *S* specific nitrilase will convert the non-labeled group to the carboxylic acid, while an *R* specific nitrilase will convert the labeled group to the carboxylic acid with loss of the ^{15}N label. The one mass unit difference is detectable by mass spectrometry.

Extensive screening led to many variants having enhanced enantioselectivity. Residues Ala190 and Phe191 were considered "hot spots" as multiple mutants were seen with variations at these residues. The most effective was the Ala190His mutant, which yielded the (*R*)-4-cyano-3-hydroxybutyric acid in 98% ee after 15 h at 2.25 M concentration of starting material. This was a significant improvement over the wild type enzyme, which yielded the product in only 88% ee after 24 h at 2.25M. No rationale for the change was given; however it is important to note that a mutation of an alanine residue to a histidine residue is statistically impossible with ep-PCR and other methods because a change of two codons in the codon triplet would be required. Thus, in this respect, GSSM has a significant advantage over other methods.

ENANTIOSELECTIVE EPOXIDE HYDROLYSIS

Aspergillus niger also produces an epoxide hydrolase that transforms many epoxides into diols with enantioselectivity.[16] The wild-type hydrolase had an *E*=4.6 for (*S*)-phenyl glycidyl ether **10** (Scheme 4). A single round of ep-PCR generated many new variants with higher selectivity factors. The 20,000 clones were prescreened for activity by using racemic epoxide **10** with 4-*p*-nitrobenzylpyridine to form a blue dye. The absence of blue color indicated the enzyme was active towards hydrolysis.

Scheme 4. Selective Hydrolysis of Epoxides

A similar high-throughput assay was employed using mass spectrometry in order to test the activity of the enzymes. A mixture of the (*S*)-enantiomer of glycidyl phenyl ether and its deuterated pseudo enantiomer was used as substrate. The five mass unit difference can be seen easily by electron

spray ionization mass spectrometry. The most efficient enzyme had a selectivity factor of 10.8, more than double that of the wild-type enzyme. Molecular modeling of the wild-type enzyme by using its crystal structure with both enantiomers of the glycidyl phenyl ether showed that residue 217 is in van der Walls contact with the phenyl ring of the substrate. The improved variant had a modification of Ala217Val, which implied that the larger valine side chain disfavors the (*R*) enantiomer due to steric interaction.

Further rounds of directed evolution were performed on this gene in subsequent experiments. However, ep-PCR was not used because it has no advantage over other methods such as DNA shuffling or GSSM. A practical method was developed and coined combinatorial active site saturation test (CAST).[17] This mutation technique focuses on the active site of the enzyme, on the assumption that modifications near the active site are more likely to affect activity and selectivity. Randomizations of the enzyme at many positions led to the most selective epoxide hydrolase to date, with *E*=115. This enzyme variant contains nine mutations from the wild-type enzyme and is 25 times more enantioselective.

OTHER EXAMPLES

Many other groups have used directed evolution to enhance product selectivity. Wong and coworkers applied ep-PCR as a means to evolve an *N*-acetylneuraminic acid aldolase to improve its catalytic activity towards enantiomerically pure substrates.[18] Three rounds of ep-PCR and saturation mutagenesis produced variants with better reactivity, and a complete reversal of enantioselectivity was observed in one variant.

Reetz and coworkers also developed the directed evolution of a cyclohexanone monooxygenase for the catalysis of an enantioselective Baeyer-Villiger reaction.[19] Using ep-PCR, they found variants that led to both enantiomers of the hydroxyl lactone from 4-hydroxycyclohexanone in significantly higher ee than the wild type enzyme. More interestingly, this same enzyme was used for the oxidation of prochiral sulfides to their respective sulfoxides.[20] The best variant had 99.8% ee, with little or no over-oxidation to the sulfone.

CONCLUSION

Directed evolution affords an effective way to change the properties of a naturally occurring enzyme rapidly to one of specific interest. The ability to enhance enantioselectivity is of great interest to organic chemists, particularly in the areas of asymmetric synthesis and pharmaceutical development. The major advantage of directed evolution is the ease of generating many catalytically active variants, which can be screened for the most selective variants. As well as having the benefits of

enantioselectivity, the mutants will allow scientists to understand better the catalytic mechanisms of enzymes. However, improvements in technique and selectivity are needed. More will be learned about directed evolution in the future; it is entirely possible that enzyme catalysts will be used in conjunction with conventional modes of asymmetric catalysts in order to produce enantiomerically enriched organic compounds and pharmaceuticals.

REFERENCES

(1) Stinson, S. C. *Chem. Eng. News* **2000***, 78*, 101-120.
(2) Noyori, R. *Angew. Chem. Int. Ed.* **2002***, 41*, 2008-2022.
(3) Brakmann, S.; Johnsson, K., Eds.; In *Directed Molecular Evolution of Proteins;* Wiley-VCH Verglab GmbH, Weinheim: 2002; , pp 257-279.
(4) Reetz, M. T.; Zonta, A.; Schimossek, K.; Liebeton, K.; Jaeger, K. E. *Angew. Chem. Int. Ed.* **1997***, 36*, 2830-2832.
(5) Leung, D. W.; Chen, E.; Goeddel, D. V. *Technique (Philadelphia)* **1989***, 1*, 11-15.
(6) Cadwell, R. C.; Joyce, G. F. *PCR Meth. App.* **1992***, 2*, 28-33.
(7) Warren, M. S.; Benkovic, S. J. *Protein Engineering* **1997***, 10*, 63-68.
(8) Kretz, K. A.; Richardson, T. H.; Gray, K. A.; Robertson, D. E.; Tan, X.; Short, J. M. *Methods Enzymol.* **2004***, 338*, 3-11.
(9) Cohen, J. *Science* **2001***, 293*, 237-239.
(10) A. Crameri, W.P.C. Stemmer *BioTechniques* **1995***, 18*, 194-196.
(11) 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.
(12) Zha, D. X.; Wilensek, S.; Hermes, M.; Jaeger, K. E.; Reetz, M. T. *Chem. Commun.* **2001**, 2664-2665.
(13) Alexeeva, M.; Enright, A.; Dawson, M. J.; Mahmoudian, M.; Turner, N. J. *Angew. Chem. Int. Ed.* **2002***, 41*, 3177-3180.
(14) Dunsmore, C. J.; Carr, R.; Fleming, T.; Turner, N. J. *J. Am. Chem. Soc.* **2006***, 128*, 2224-2225.
(15) DeSantis, G.; Wong, K.; Farwell, B.; Chatman, K.; Zhu, Z.; Tomlinson, G.; Huang, H.; Tan, X.; Bibbs, L.; Chen, P.; Kretz, K.; Burk, M. J. *J. Am. Chem. Soc.* **2003***, 125*, 11476-11477.
(16) Reetz, M. T.; Torre, C.; Eipper, A.; Lohmer, R.; Hermes, M.; Brunner, B.; Maichele, A.; Bocola, M.; Arand, M.; Cronin, A.; Genzel, Y.; Archelas, A.; Furstoss, R. *Org. Lett.* **2004***, 6*, 177-180.
(17) Reetz, M. T.; Wang, L.; Bocola, M. *Angew. Chem. Int. Ed.* **2006***, 45*, 1236-1241.
(18) Wada, M.; Hsu, C. C.; Franke, D.; Mitchell, M.; Heine, A.; Wilson, I.; Wong, C. H. *Bioorg. Med. Chem.* **2003***, 11*, 2091-2098.
(19) Reetz, M. T.; Brunner, B.; Schneider, T.; Schulz, F.; *Angew. Chem. Int. Ed.* **2004***, 43*, 4075-4078.
(20) Reetz, M. T.; Daligault, F.; Brunner, B.; Hinrichs, H.; Deege, A. *Angew. Chem. Int. Ed.* **2004***, 43*, 4078-4081.