



<http://www.diva-portal.org>

## Postprint

This is the accepted version of a paper published in *Journal of the American Chemical Society*. This paper has been peer-reviewed but does not include the final publisher proof-corrections or journal pagination.

Citation for the original published paper (version of record):

Reetz, M., Bocola, M., Wang, L., Sanchis, J., Cronin, A. et al. (2009)

Directed Evolution of an Enantioselective Epoxide Hydrolase: Uncovering the Source of Enantioselectivity at Each Evolutionary Stage.

*Journal of the American Chemical Society*, 131(21): 7334-7343

<http://dx.doi.org/10.1021/ja809673d>

Access to the published version may require subscription.

N.B. When citing this work, cite the original published paper.

Permanent link to this version:

<http://urn.kb.se/resolve?urn=urn:nbn:se:uu:diva-130672>

# Directed Evolution of an Enantioselective Epoxide Hydrolase: Uncovering the Source of Enantioselectivity at Each Evolutionary Stage

Manfred T. Reetz,<sup>\*,†</sup> Marco Bocola,<sup>\*,‡</sup> Li-Wen Wang,<sup>†</sup> Joaquin Sanchis,<sup>†</sup> Annette  
Cronin,<sup>§</sup> Michael Arand,<sup>§</sup> Jinyu Zou,<sup>\*\*</sup> Alain Archelas,<sup>††</sup> Anne-Lise Bottalla,<sup>††</sup>  
Agata Naworyta<sup>‡‡</sup> and Sherry L. Mowbray<sup>\*,‡‡</sup>

*Contribution from the Max-Planck-Institut für Kohlenforschung, Kaiser-Wilhelm-Platz 1,  
D-45470 Mülheim/Ruhr, Germany, Institute of Pharmacology and Toxicology, Universität Zürich  
Winterthurerstrasse 190, CH-8057 Zürich, Switzerland, Department of Cell and Molecular Biology,  
Uppsala University, Box 596, Biomedical Center, S-75124 Uppsala, Sweden, University Paul-  
Cézanne/Faculté Sciences et Techniques de St Jérôme, case 432 Avenue Escadrille Normandie Niemen,  
13397 Marseille Cedex 20, France, and Department of Molecular Biology, Swedish University of  
Agricultural Sciences, Box 590, Biomedical Center, S-75124 Uppsala, Sweden*

Received ....., 2009; E-mail: reetz@mpi-muelheim.mpg.de;

marco.bocola@biologie.uni-regensburg.de; mowbray@xray.bmc.uu.se

---

<sup>†</sup> Max-Planck-Institut für Kohlenforschung

<sup>‡</sup> Present address: Institute of Biophysics and Physical Biochemistry, University of Regensburg, Universitätsstrasse 31,  
93053 Regensburg, Germany

<sup>§</sup> Institute of Pharmacology and Toxicology

<sup>\*\*</sup> Uppsala University

<sup>††</sup> Faculté Sciences et Techniques de St Jérôme

<sup>‡‡</sup> Swedish University of Agricultural Sciences

---

**Abstract:** Directed evolution of enzymes as enantioselective catalysts in organic chemistry is an alternative to traditional asymmetric catalysis using chiral transition metal complexes or organocatalysts, the different approaches often being complementary. Moreover, directed evolution studies allow us to learn more about how enzymes perform mechanistically. The present study concerns a previously evolved highly enantioselective mutant of the epoxide hydrolase from *Aspergillus niger* in the hydrolytic kinetic resolution of racemic glycidyl phenyl ether. Kinetic data, molecular dynamics calculations, molecular modeling, inhibition experiments and X-ray structural work for the wild-type (WT) enzyme and the best mutant reveal the basis of the large increase in enantioselectivity ( $E = 4.6$  versus  $E = 115$ ). The overall structures of the WT and the mutant are essentially identical, but dramatic differences are observed in the active site as revealed by the X-ray structures. All of the experimental and computational results support a model in which productive positioning of the preferred (*S*)-glycidyl phenyl ether, but not the (*R*)-enantiomer, forms the basis of enhanced enantioselectivity. Predictions regarding substrate scope and enantioselectivity of the best mutant are shown to be possible.

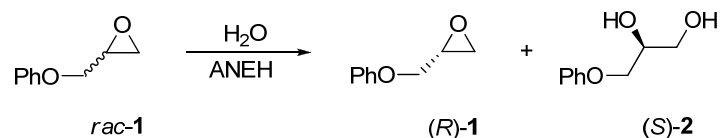
## Introduction

Directed evolution has emerged as a powerful means to engineer the properties of enzymes as catalysts in synthetic organic chemistry and biotechnology.<sup>1</sup> It is based on repeated rounds of mutagenesis, expression and screening (or selection). The properties that have been targeted include enhanced robustness in hostile organic solvents,<sup>2</sup> higher thermostability<sup>3</sup> and increased enantioselectivity.<sup>4</sup> Several challenges accompany these studies. In addition to methodology development necessary for making directed evolution more efficient and less labor-intensive than in the past,<sup>1,4d,5,6</sup> the need to understand the basis of catalyst improvement on a molecular level is likewise important. This kind of knowledge also deepens our understanding of how enzymes function as catalysts. Explanations for improved thermostability of proteins have been reviewed,<sup>3,7</sup> new hydrogen bonds and/or salt bridges on the protein surface often being the crucial factors. In the case of directed

evolution of enantioselectivity, the situation is particularly complex because the improved catalysis may involve very different types of effects.

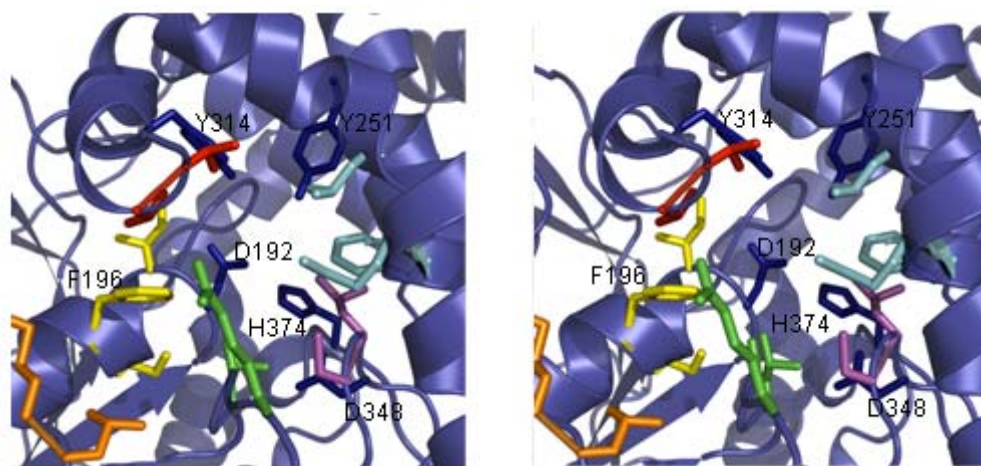
More than a decade ago we reported proof-of-principle regarding directed evolution of enantioselective enzymes, specifically of a lipase in the hydrolytic kinetic resolution of a chiral ester.<sup>8,9</sup> Subsequently an investigation utilizing molecular molecular dynamics (MD) calculations suggested the source of enantioselectivity.<sup>10</sup> Unfortunately, it was not possible to obtain X-ray structural data for any of the enantioselective lipase mutants. Comparison with the available crystal structure of the WT would have provided the possibility to validate the proposed model. The lack of X-ray structural data accompanies all other evolved stereoselective enzyme mutants reported so far.<sup>4,10-14</sup> Of course, X-ray data of enantioselective wild-type (WT) enzymes have been used to interpret the source of enantioselectivity.<sup>15</sup> Wong succeeded in altering the substrate scope of the Neu5Ac aldolase significantly, thereby making diastereoselective reactions possible for enantiomeric aldehydes that are not accepted by the WT.<sup>12</sup> However, while the X-ray structure of one triple mutant was obtained at 2.3 Å resolution, it proved to be essentially identical to the WT, and thus provided no help in interpreting the kinetic results.<sup>12</sup>

Recently we introduced the Combinatorial Active-Site Saturation Test (CAST) as an efficient means to evolve substrate acceptance and/or enantioselectivity of enzymes.<sup>16</sup> In this method sites around the binding pocket are systematically subjected to saturation mutagenesis, a given site being composed of one or more amino acid positions. This approach constitutes a systematization of focused library generation practiced previously in many directed evolution studies.<sup>1,9,17</sup> Thereafter, iterative CASTing was developed, in which the gene of an improved mutant originating from one site is used as a template to perform randomization at another site, the process being repeated until the desired degree of catalyst improvement has been achieved.<sup>18,19</sup> This method was first applied to the hydrolytic kinetic resolution of glycidyl phenyl ether (*rac*-**1**), catalyzed by the epoxide hydrolase from *Aspergillus niger* (ANEH),<sup>18</sup> and has since been applied to other enzymes.<sup>14</sup> The WT ANEH exhibits a selectivity factor of only  $E = 4.6$ , in slight favor of (*S*)-**2**.



### Formula 1

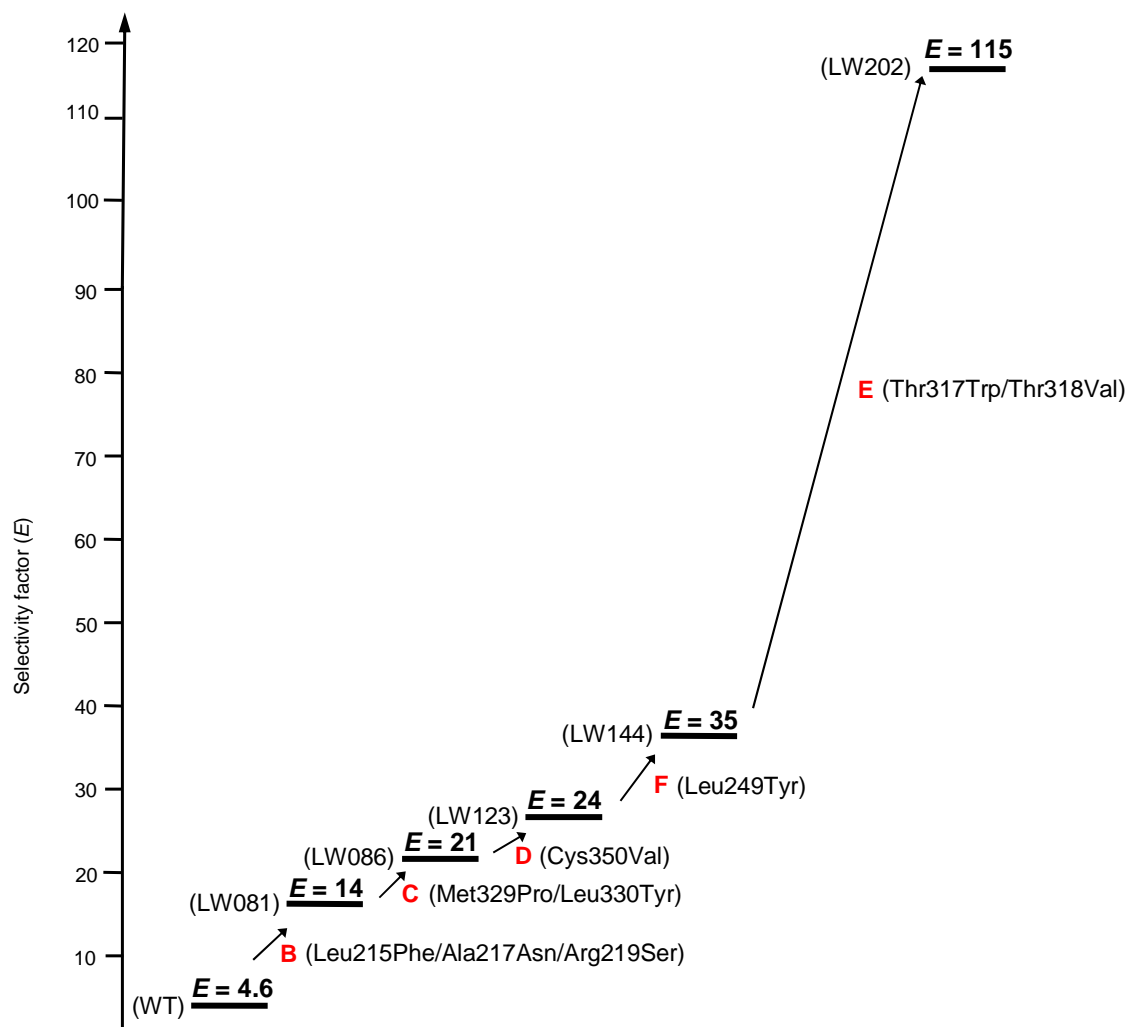
The substrate *rac-1* was first modeled into the narrow tunnel-like binding pocket of the WT ANEH as revealed by its crystal structure,<sup>20</sup> and on this basis six sites were chosen for mutagenesis experiments, namely A (comprising amino acid positions 193/195/196), B (215/217/219), C (329/330), D (349/350), E (317/318) and F (244/245/249) (see stereo diagram in Figure 1).<sup>18</sup>



**Figure 1.** The active site of WT ANEH<sup>20</sup> (stereo diagram). Sites for CASTing<sup>18</sup> were chosen as described in the text, and color coded as follows: A (yellow), B (green), C (orange), D (purple), E (red), F (blue). The nucleophile (Asp192), the key tyrosines that bind and activate the epoxide via hydrogen bonds (Tyr251 and Tyr314), and the residues involved in the hydrolysis of the intermediate ester formed after the ring-opening step (His374 and Asp348) are also labeled. See also the section on Molecular Dynamics Calculations outlining the mechanistic details.

In the original study, the pathway B→C→D→F→E (with a total of nine amino acid substitutions) led to the highly enantioselective mutant LW202 (*E*-value = 115); see Figure 2.<sup>18</sup> Other pathways were not explored at the time, nor were attempts made to explain the results on a molecular level. The present study addresses the latter question. We report kinetic data for the WT ANEH and mutant LW202,

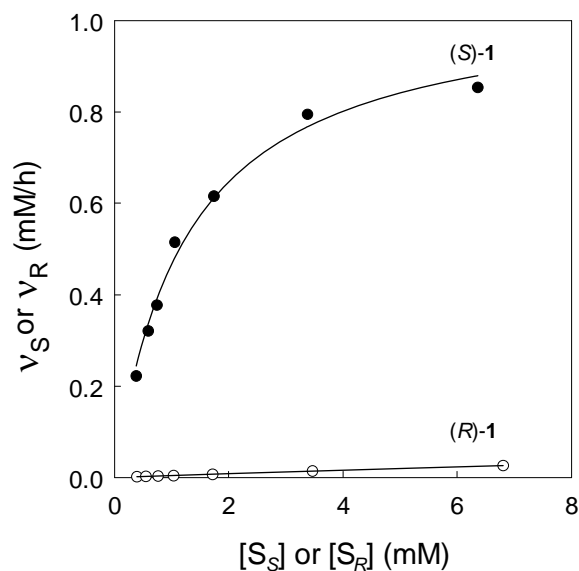
accompanied by MD calculations at all five intermediate stages of the evolutionary pathway. This allows us to propose a model explaining the reasons for enhanced enantioselectivity, which we validated by comparing the X-ray structures of the WT and the best mutant LW202, and by applying docking procedures based on the 3D data.



**Figure 2.** Iterative CASTing in the evolution of enantioselective ANEH mutants as catalysts in the hydrolytic kinetic resolution of *rac-1*.<sup>18</sup>

## Results and Discussion

**Kinetic Data.** Following introduction of a C-terminal His-tag in the best mutant LW202, purification was performed using an affinity column. In a final step the mutant was eluted as a single homogeneous (dimeric) peak on gel filtration. Kinetic studies of the model reaction utilizing (*R*)-**1** and (*S*)-**1** separately were then carried out, and the data fit to the Michaelis-Menten equation. The results with the best mutant LW202 as catalyst are shown in Figure 3, which demonstrate the pronounced preference for (*S*)-**1** in an almost perfect kinetic resolution.



**Figure 3.** Kinetic analysis of mutant LW202 based on the Michaelis-Menten equation, where  $v_S$  and  $v_R$  are the initial rates of hydrolysis of (*S*)- and (*R*)-**1** catalyzed by the enzyme at different substrate concentrations  $[S_S]$  or  $[S_R]$ .

The  $k_{cat}$ - and  $K_M$ -values for both enantiomers of substrate **1** were also determined for WT ANEH. The results summarized in Table 1 reveal two important features. Firstly, the mutational changes in going from the WT to LW202 reduce the  $k_{cat}/K_M$ -values considerably. Indeed, the screening system was devised for optimizing enantioselectivity, not activity.<sup>18</sup> Secondly, comparison of the  $k_{cat}/K_M$  of LW202 for the reaction of the favored (*S*)-substrate with that of the disfavored reaction involving the (*R*)-

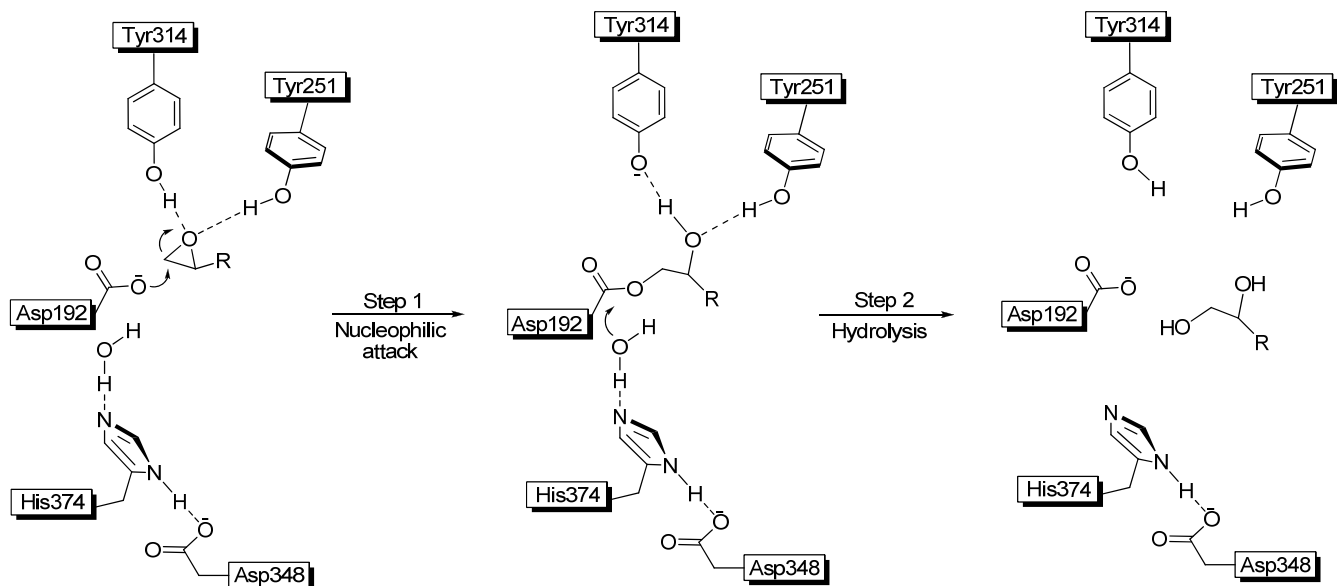
enantiomer again shows the significant degree of preference. The selectivity factor derived from the kinetic data is  $E = 193$ , which is even higher than the previously reported value of 115 based on the Sih equation;<sup>21</sup> the value for the WT is also slightly lower. This difference is probably due to the use of whole cells in the earlier experiments versus isolated enzymes in the present kinetic study.

**Table 1.** Pre-steady state kinetic constants of the WT ANEH and of mutant LW202 (Subscript *S* represents (*S*)-**1** and *R* denotes (*R*)-**1**)

Enzyme	$k_{cat,S}$ [mmol h <sup>-1</sup> g <sup>-1</sup> ]	$K_{M,S}$ [mM]	$(k_{cat}/K_M)_S$ [M <sup>-1</sup> h <sup>-1</sup> ]	$k_{cat,R}$ [mmol h <sup>-1</sup> g <sup>-1</sup> ]	$K_{M,R}$ [mM]	$(k_{cat}/K_M)_R$ [M <sup>-1</sup> h <sup>-1</sup> ]	<i>E</i>
WT	$6.89 \times 10^3$	$8.65 \times 10^{-2}$	$7.97 \times 10^4$	$1.61 \times 10^4$	$6.80 \times 10^{-1}$	$2.64 \times 10^4$	3
LW202	$9.22 \times 10^3$	1.25	$7.35 \times 10^3$	$1.96 \times 10^3$	$5.13 \times 10^1$	$3.8 \times 10^1$	193

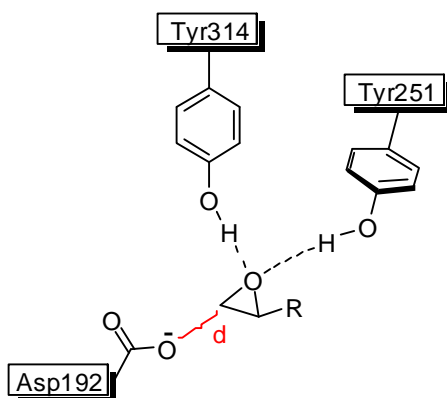
**Molecular Dynamics Calculations.** Hydrolysis of epoxides catalyzed by ANEH follows a mechanism in which a covalent glycol-monoester-enzyme intermediate is formed in the rate-determining step,<sup>20,22</sup> as for many other but not all epoxide hydrolases.<sup>23</sup> Based on mechanistic studies and on the X-ray structure of WT ANEH<sup>20,22</sup>, it was deduced previously that two tyrosines at positions 251 and 314 bind and concomitantly activate the substrate by forming hydrogen bonds at the epoxide O-atom, and that Asp192 then initiates the rate-determining nucleophilic attack at the sterically less hindered C-atom, resulting in ring opening and the formation of a covalently bound ester intermediate (Scheme 1).<sup>20,22</sup> In the second (fast) step the intermediate reacts with water, which is activated and positioned by the spatially close His374. This hydrolysis closes the catalytic cycle.





**Scheme 1.** Reaction mechanism of ANEH.<sup>20,22</sup>

To study how the activated (*R*)- and (*S*)-substrates are positioned in the narrow binding pocket, and to see if any substantial differences occur upon going from WT ANEH to LW202 along the evolutionary pathway, we performed molecular dynamics (MD) simulations. The energetics of binding were expected to depend on several factors, including the hydrogen bonds with the two tyrosines, which also activate the substrate. The through-space distance, *d*, between the attacking O-atom of Asp192 and the epoxide C-atom undergoing  $S_N2$  reaction was thought to be of particular importance (Figure 4).



**Figure 4.** Definition of the distance, *d*, in the rate determining step of ANEH-catalyzed reactions.

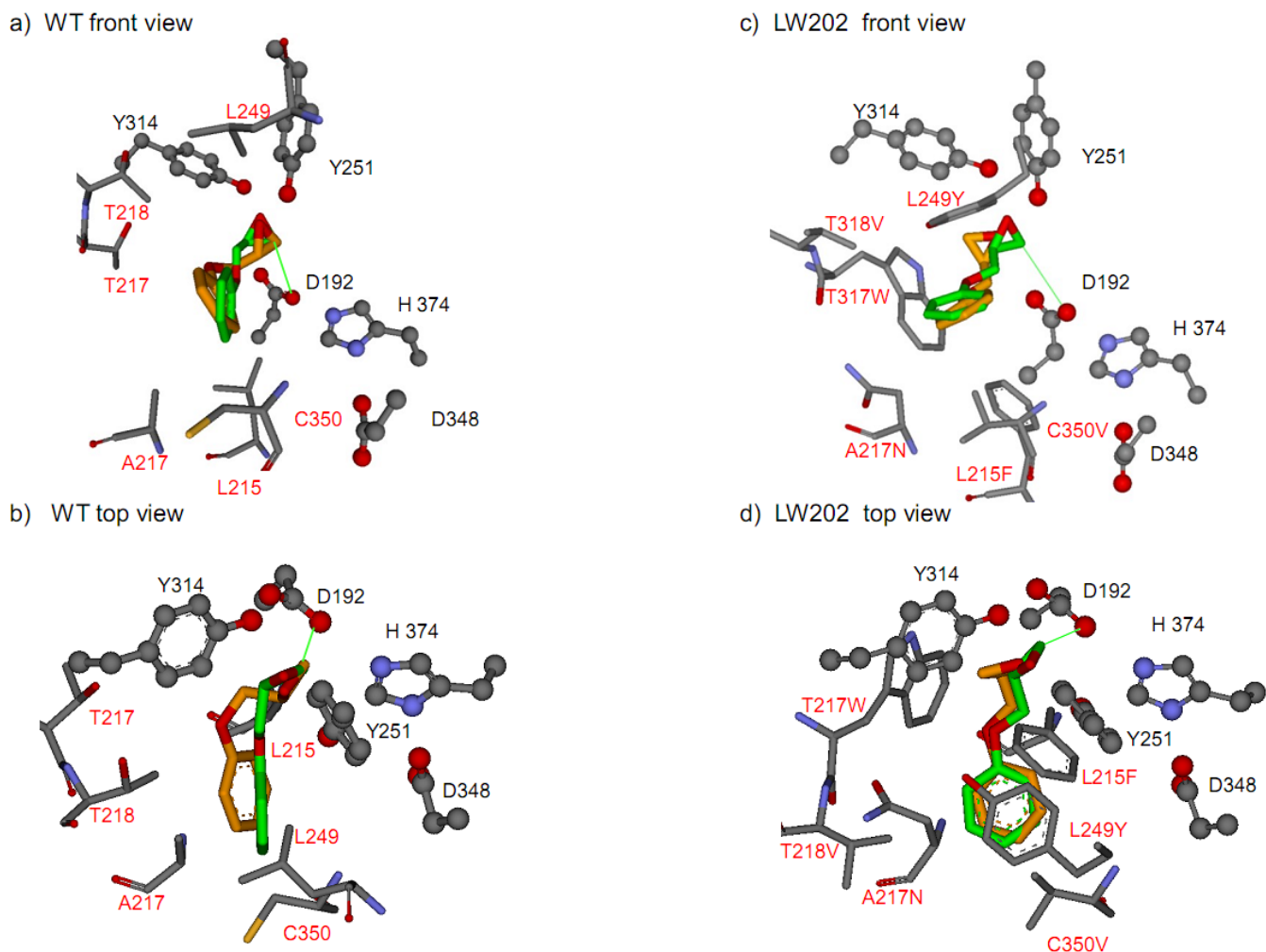
We surmised that a sufficiently small d-value could correspond to a near-attack-conformer as discussed by Bruice in other enzyme-catalyzed reactions,<sup>24</sup> or more generally to a productive position, and that consequently, this distance should be shorter in the case of the preferred enantiomer. MD calculations using the force field MAB/DAB<sup>25</sup> included in the modeling suite MOLOC<sup>26</sup> were performed for (*R*)- and (*S*)-**1** bound in the WT and in all mutants along the evolutionary pathway (Figure 2). Based on the X-ray structure of WT ANEH without bound ligand,<sup>20</sup> we constructed a model of the WT harboring either (*R*)- or (*S*)-**1** in a reactive position prior to the nucleophilic attack of Asp192 (Figure 4); it was assumed that hydrogen bonds originating from both Tyr314 and Tyr251 are present. The doubly protonated His374 in the catalytic triad was maintained in the catalytic H-bond network. All intermediate ANEH mutants were built in the same way until the final mutant LW202 having nine mutations was reached. To test the predictive power of the MD calculations, the X-ray data of LW202 (see below) were not invoked at this point. The analogous procedure was performed for both enantiomers of substrate **1** covalently bound to Asp192 in the form of an ester intermediate (Scheme 1). We analyzed all recorded 1ns trajectories for several properties, including energies, distances and angles (data not shown). The results of these calculations are summarized in Table 2.

**Table 2.** Results of MD calculations (for the definition of d, see Figure 4)

Mutant	$d_R$	$d_S$	$\Delta d_{R-S}$	$E(\text{expl.})$
WT	4.3	3.5	0.8	4.6
LW081	4.8	4.0	0.8	14
LW086	4.9	4.0	0.9	21
LW123	5.1	4.0	1.1	24
LW44	5.1	3.9	1.3	35
LW202	5.4	3.8	1.6	115

The most significant finding is the clear correlation ( $R^2 = 0.86$ ) between the measured  $E$ -values and the differences in the modeled distance,  $\Delta d$  (Figure 4), for the two enantiomers. In the WT the  $\Delta d$ -value is expected to be only 0.8 Å, with the preferred ( $S$ )-enantiomer slightly closer to the attacking Asp192. In the case of the highly enantioselective LW202, the situation is strikingly different; here  $\Delta d$  has increased to 1.6 Å. For the preferred ( $S$ )-**1**, the value for  $d$  does not change substantially when going from WT ANEH to the best mutant LW202, remaining near 3.8 Å, but the activated ( $R$ )-substrate is positioned much further away ( $d = 5.4$  Å), which disfavors ring-opening nucleophilic attack (Table 2). Thus, structural changes in the binding pocket imposed by the evolutionary process are predicted to make it much more difficult for the disfavored enantiomer to experience activation by the two tyrosines *and* at the same time to be positioned close enough to Asp192 for rapid nucleophilic attack to proceed. In summary, our analysis suggests that the enhanced enantioselectivity is due primarily to a reduction in the rate of reaction of ( $R$ )-**1**, which is in agreement with the kinetic study. *Thus, differences in substrate positioning as a function of the absolute configuration of the reacting compound are predicted to increase stepwise as the evolutionary process proceeds.*

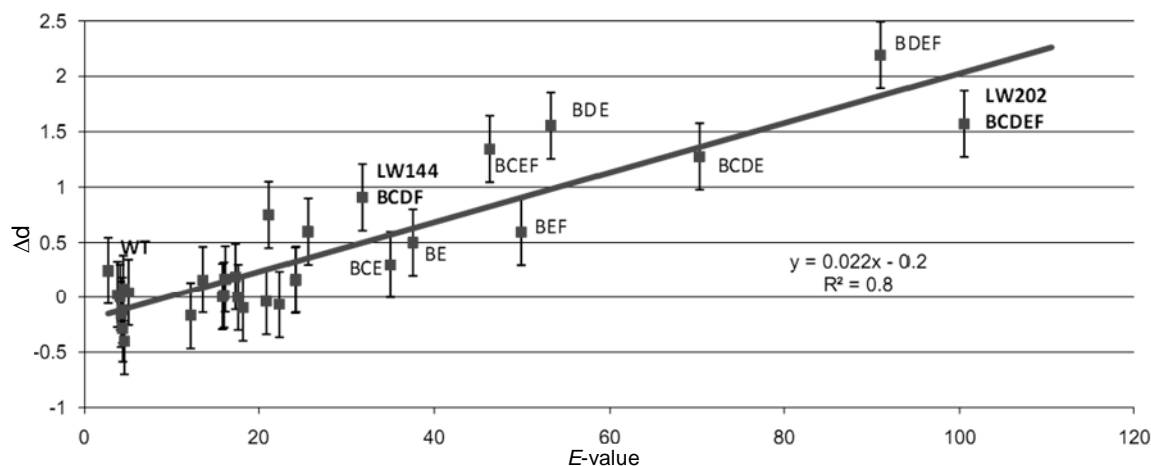
Figure 5 illustrates ( $R$ )- and ( $S$ )-**1** positioned in the binding pocket of WT ANEH and in the binding pocket of the mutant LW202 as predicted by the MD calculations. We see that increased steric clashes at the active site hinder productive binding of ( $R$ )-**1**. In the final mutant LW202, the phenyl moiety of the substrate is predicted to lie in an “aromatic box” built up by Phe215, Tyr 249 and Trp317. This enforces a similar binding mode of the phenyl ring for both enantiomers. In the case of the slow-reacting enantiomer ( $R$ )-**1**, this binding mode places the epoxide ring in an unproductive orientation with C1 pointing away from Asp192. In contrast, the favored enantiomer ( $S$ )-**1** is maintained as a productive, or perhaps even near-attack conformer<sup>24</sup> in the binding pocket of both the WT and the best mutant LW202.



**Figure 5.** Results of MD calculations displaying the predicted positions of the enantiomeric substrates (*R*)- and (*S*)-**1** (green and orange sticks, respectively) in the binding pocket (catalytic residues shown in ball and stick, mutated residues in stick representation) of WT ANEH (a and b), and in the binding pocket of the best mutant LW202 (c and d).

We also predict some flexibility in side-chain conformations that appears significant for epoxide binding. This is most prominent for the residues around the phenyl ring of the substrate, e.g. Leu215, Leu249, Thr317, Thr318 and Cys350 of WT ANEH, which show different conformers for the two bound enantiomers in our MD-simulations. In the final mutant LW202, all are replaced by more hydrophobic (and generally larger) amino acids (Phe215, Tyr249, Trp317, Val318 and Val350). The above mentioned hydrophobic box is maintained, as is the ideal positioning of the favored (*S*)-substrate in the binding pocket.

Our model is also useful in rationalizing further data that were recently acquired concerning LW202 as a catalyst in the hydrolytic kinetic resolution of *rac*-**1**, specifically the results of a deconvolution study.<sup>27</sup> In that investigation the five *sets* of mutations which successively accumulated in the chosen pathway B→C→D→F→E were combined in all permutational ways with the construction of a fitness landscape characterized by the  $5! = 120$  possible trajectories linking the WT with LW202. New mutants were obtained by this procedure, without introducing more amino acid exchanges. In the present study we performed a computational analysis of this deconvolution dataset by calculating  $\Delta d$ -values and comparing them to the experimentally obtained enantioselectivities.<sup>27</sup> Figure 6 shows that our model once again leads to an acceptable correlation ( $R^2 = 0.80$ ).

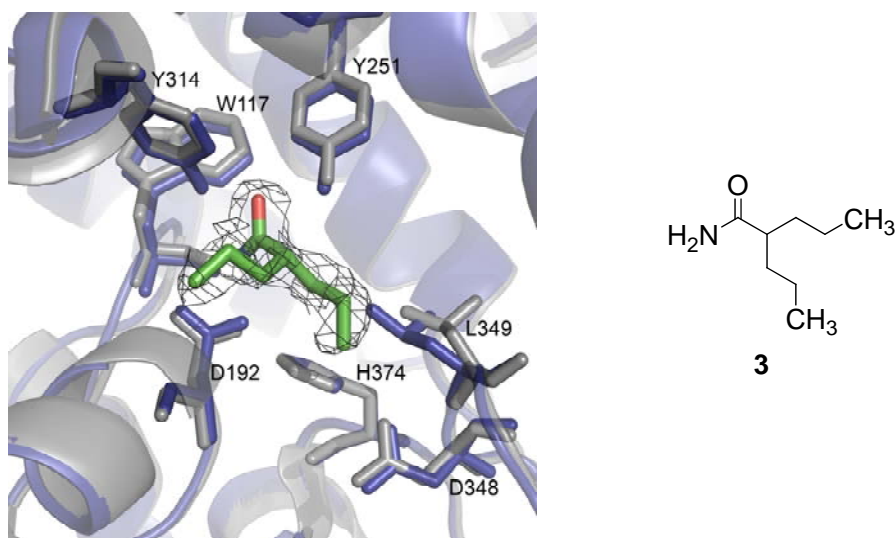


**Figure 6.** Correlation of MD calculations ( $\Delta d$ ) with the experimental *E*-values of the deconvolution dataset<sup>27</sup> considering all combinatorial dissections of the five mutant libraries B to F, the model reaction being the hydrolytic kinetic resolution of *rac*-**1** catalyzed by ANEH mutants.

All combinations of sets of mutations obtained in the earlier study<sup>27</sup> can be analyzed on the basis of the proposed model. For example, this leads to an understanding as to why the set of mutations obtained at site B (Leu215Phe/Ala217Asn/Arg219Ser) acts cooperatively with that found at site E (Thr317Trp/Thr318Val) resulting in an *E*-value of 38.<sup>27</sup> The two sites are located on opposite sides of the entrance to the funnel-like binding pocket (Figure 1), and the side-chains of the respective amino acid residues are

expected to interact with the relatively bulky phenyl substituent of substrate *rac*-**1**. In the case of the disfavored (*R*)-substrate, the model predicts steric pressure on the phenyl group from residues of the mutants remote from the actual stereocenter. This effect is further enhanced by other combinations, e.g., with changes at site D (Cys350Val) leading to the best triple combination BDE with an *E*-value of 53.<sup>27</sup> Although all of the mutated sites are next to the binding pocket, they are fairly far away from the actual stereocenter at position C2 in the oxirane ring. This steric stereocontrol-mechanism as part of our model also allows the prediction of substrate acceptance (see Section on Substrate Scope of Mutant LW202).

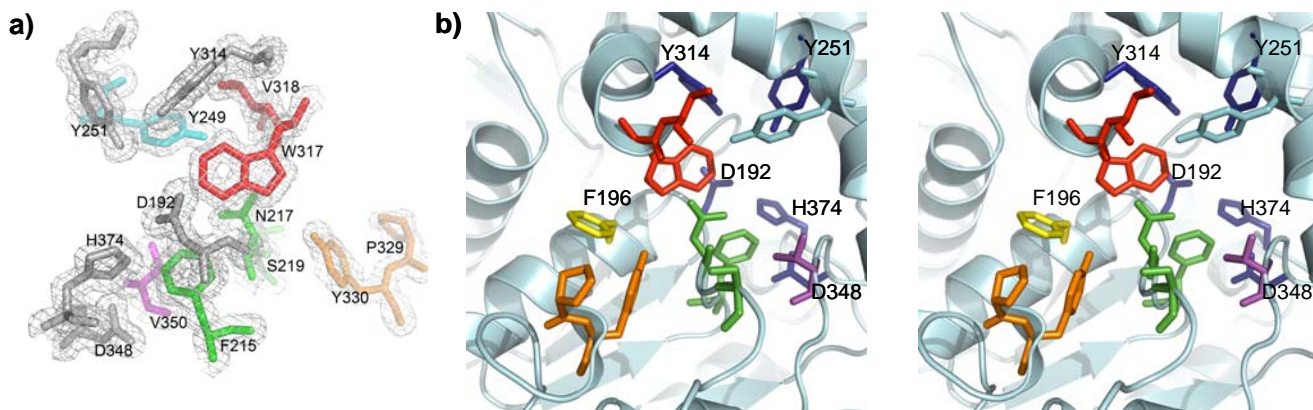
**X-Ray Structural Analyses.** The above model appears plausible, but it lacks direct structural support. As noted above, the X-ray structure of WT ANEH in the absence of a substrate or an inhibitor was already available.<sup>20</sup> In the present study we report the structure of WT ANEH bound to the inhibitor valpromide (**3**)<sup>28</sup> at 2.1 Å resolution (Figure 7). Binding studies indicate a measured  $K_i$  of 250 μM for this compound.



**Figure 7.** X-ray structures of WT ANEH with and without the inhibitor **3**. The active site of the ligand-free enzyme structure is shown in gray, while the structure including **3** is blue. Electron density is shown for bound ligand in the A molecule of the asymmetric unit, using a SIGMAA-weighted  $2|F_o| - |F_c|$  map<sup>29</sup> contoured at  $1\sigma = 0.40 \text{ e}/\text{\AA}^3$ . Residues expected to participate in binding are labeled, together with Leu349, which shows the largest movements in the active-site area.

The amide oxygen of **3** forms short, strong (2.4 to 2.6 Å) hydrogen bonds to Tyr251 and Tyr314. Its amide H<sub>2</sub>N-moiety forms two hydrogen bonds to Asp192, and one to the main-chain carbonyl oxygen of residue 117 (a component of the oxyanion hole). Significant movements of the main- and side-chain (up to 1.7 Å) are observed at Leu349 (a motion not predicted by the MD studies), which are needed to accommodate one of the two aliphatic “arms” of the inhibitor. Otherwise, the structure is very similar to that of WT ANEH lacking an inhibitor or substrate (rms difference of 0.2 Å when all C $\alpha$  atoms are compared). It is also worth noting that the mode of binding of **3** observed here is different from that observed for the potato epoxide hydrolase (19% amino acid sequence identity).<sup>30</sup> Although in both cases the amide H<sub>2</sub>N-moiety interacts with the aspartate nucleophile, the other groups are arranged differently. This arises from changes in the shape of the binding pocket; in the potato enzyme, the aliphatic groups are bound in an extra cavity that is blocked in ANEH.

Obtaining the crystal structure of the best mutant LW202 was considered to be even more important. Efforts to grow crystals with the inhibitor **3** were not successful. However, it was possible to solve the ligand-free structure of LW202 at 1.5 Å resolution (Figure 8), and therefore to compare it to the WT (Figure 1). As for WT ANEH, the asymmetric unit was found to contain two independent copies of the protein, which will be referred to as molecules A and B. Unless stated otherwise, the two molecules can be assumed to be equivalent in the following analysis. The gross structural features of LW202 are the same as in WT ANEH without bound ligand (rms difference 0.3 Å when all C $\alpha$  atoms are compared). *However, comparison of the respective binding pockets reveals significant differences.* The catalytic residues are unperturbed, but other residues of the binding pocket are dramatically different. In addition to the changes that originate directly from the differences of size and character in the residues mutated, the side-chain of Phe196 has a different conformation as a secondary effect of the Thr317Trp mutation. The phenyl group points in a different direction to provide space for Trp317. Interestingly, the conformation of Leu349 is more similar to that in the WT ligand-free enzyme, rather than to that with bound valpromide (**3**). The increased bulk of residues 215 and 317, in particular, in the mutant is expected to prevent binding of compound **3** at the position observed for the WT, due to steric clashes with one of its aliphatic “arms”.



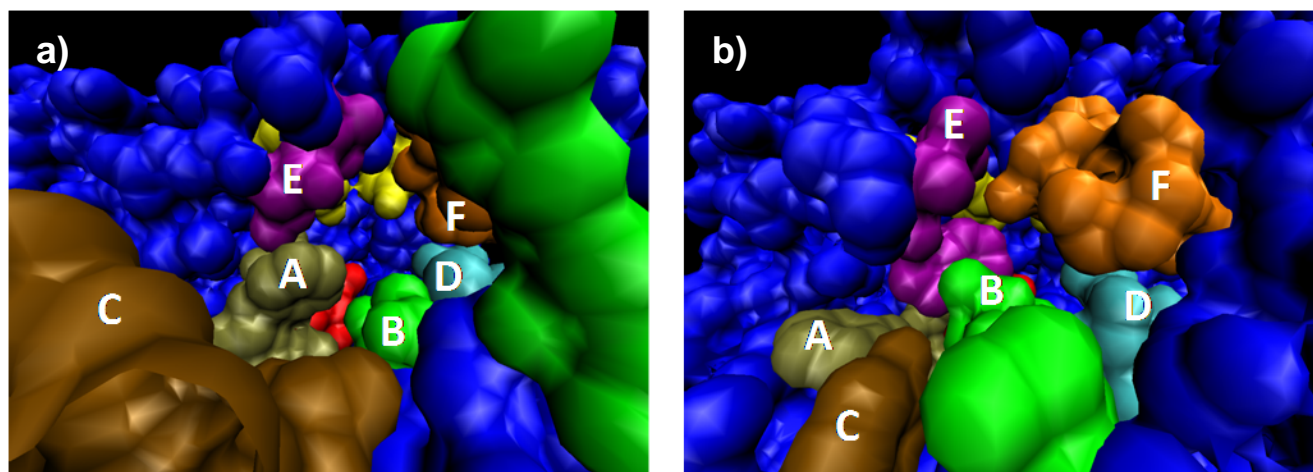
**Figure 8.** X-ray structure of the mutant LW202 in the absence of an inhibitor or a substrate. Panel a): Residues in the active-site of the LW202 mutant (molecule A), together with electron density from a SIGMAA-weighted  $2|F_o|-|F_c|$  map<sup>29</sup> contoured at a level of  $1\sigma = 0.42 \text{ e}/\text{\AA}^3$ . Residues participating in catalysis (Asp192, His374, Asp348, Tyr314 and Tyr251) are shown in gray, as well as those mutated, using the coloring convention introduced in Figure 1. Panel b): Stereo diagram.

As predicted by the MD calculations based solely on the X-ray data of WT ANEH and modeling of the mutations of LW202, the mutant structure exhibits greater flexibility than WT ANEH, although there is not complete agreement concerning which regions are mobile. Many residues have multiple conformations, but side-chain mobility of Ser195 and Val318 is very clear in the active sites of both molecules A and B of the asymmetric unit. When the two molecules of the mutant structure are compared, movements of both main- and side-chains in the 217–222 region are also apparent, as well as those of nearby residue 356. As already noted, the phenyl group of Phe196 assumes a different orientation than was seen in the WT; further, the two molecules of the mutant structure (LW202) show slightly different conformations from each other. Additionally, Tyr330 has somewhat different conformation in molecules A and B of LW202. The relationship to mutations is in most cases clear. Molecule B of the mutant structure LW202 has an extra disordered loop (residues 222–229); otherwise all structures have a single disordered loop at residues 319–329. Given that the usually high degree of



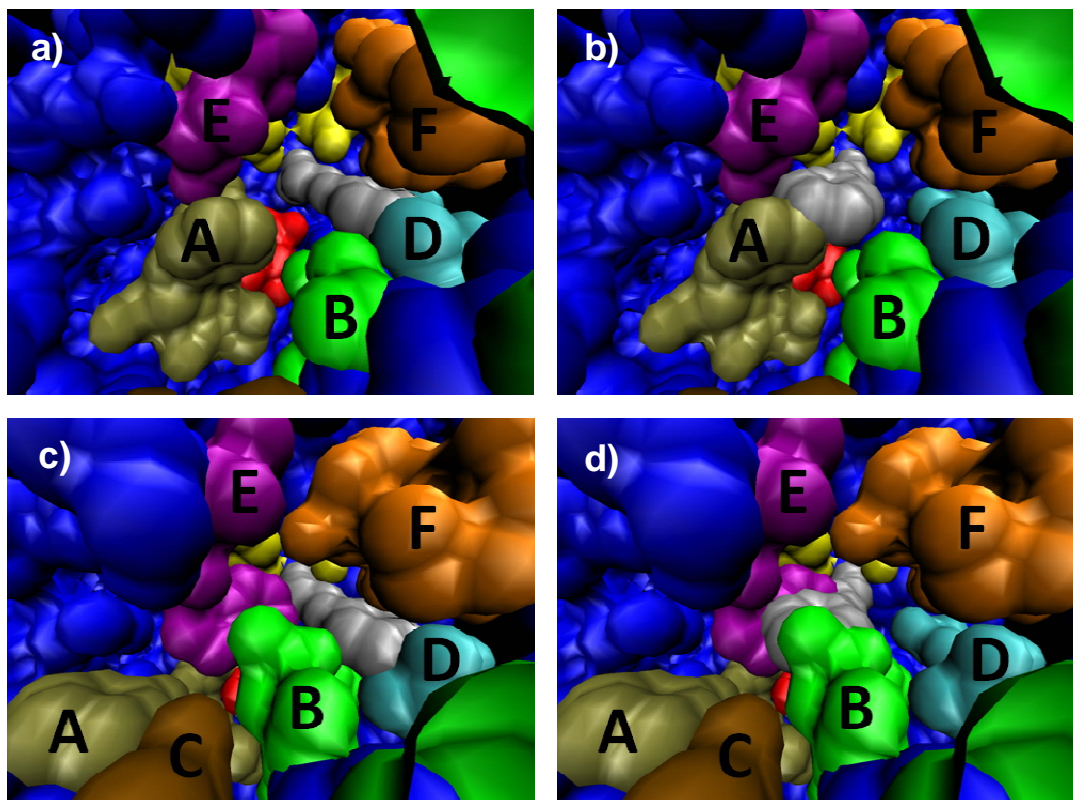
flexibility of the mutant structure could conceivably be linked to a loss of stability, it should be noted that the thermostability of ANEH is not compromised when going from the WT to the mutant LW202.<sup>31</sup> Additional electron density in the active sites of both molecules in the asymmetric unit could be assigned to formic acid originating from the crystallization solutions. The bound formic acid forms strong interactions with the nucleophile Asp192, as well as the two catalytic tyrosines.

**Manual Docking Experiments.** The observed flexibility makes it difficult to do meaningful automated docking studies for the mutant, since it is not reasonable to test all possible sets of relevant conformations in the procedure. We therefore attempted to interpret the pronounced increase in enantioselectivity observed upon going from WT ANEH to the mutant LW202 by manually docking the enantiomeric substrates into the X-ray structures of WT ANEH and LW202, respectively, and compared the results with the predictions made earlier based on MD calculations performed before structural data for the mutant were available (Table 2; Figure 5). Using the X-ray structure of WT ANEH hosting **3** and of the mutant LW202, space-filling representations of the respective empty binding pockets are shown in Figures 9a and 9b. It becomes clear that the binding pockets are geometrically quite different, especially around sites B and E. Some of the differences are due to significant movements of the side-chains of *non-mutated residues*, most notably of Phe196 (part of site A, but not mutagenized in LW202).



**Figure 9.** a) Empty binding pocket of WT-ANEH based on the X-ray structure of the valpromide-bound form; b) Empty binding pocket of mutant LW202 based on the X-ray structure of the ligand-free form. In red is shown the catalytic Asp192, and in yellow, the activating couple Tyr251/Tyr314; A-F indicate the respective randomization sites for CASTing (note that site A was not included in the current evolutionary pathway B→C→D→F→E).

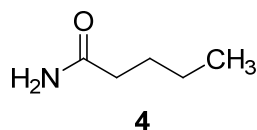
Manual docking of the two enantiomeric substrates (*S*)- and (*R*)-**1** into the binding pockets of WT ANEH and LW202 (molecule A), so that activation by Tyr251/Tyr314 and optimal positioning with respect to the nucleophile Asp192 are ensured, proved to be revealing (Figures 10a–d). In the case of the preferred (*S*)-enantiomer, the substrate fits easily into the WT (Figure 10a) *and* into LW202 (Figure 10c) without any energetically unfavorable steric interactions. In contrast, the disfavored (*R*)-enantiomer fits well into the binding pocket of the WT (Figure 9b), but not into that of the mutant LW202 (Figure 10d). Severe steric clashes between the substrate and the mutated residues, especially at sites B and E, make productive binding essentially impossible in the latter case. This is consistent with the prediction of the MD simulations (see above) that the binding mode of the disfavored (*R*)-enantiomer changes when going from WT ANEH to mutant LW202, resulting in an unproductive substrate orientation. Thus, this structure-based analysis and the MD calculations are both in accord with the kinetic data.



**Figure 10.** a) Binding pocket of WT ANEH (based on the X-ray structure of the valpromide-bound form) in which (*S*)-**1** has been docked by overlapping the oxygen, nitrogen and carbonyl- carbon in valpromide (**3**) with the oxygen, primary (terminal) and secondary carbon in (*S*)-**1** respectively; b) Binding pocket of WT ANEH in which (*R*)-**1** has been docked in the same manner; c) Binding pocket of mutant LW202 (molecule A) in which (*S*)-**1** has been docked; d) Binding pocket of LW202 (molecule A) in which (*R*)-**1** has been docked.

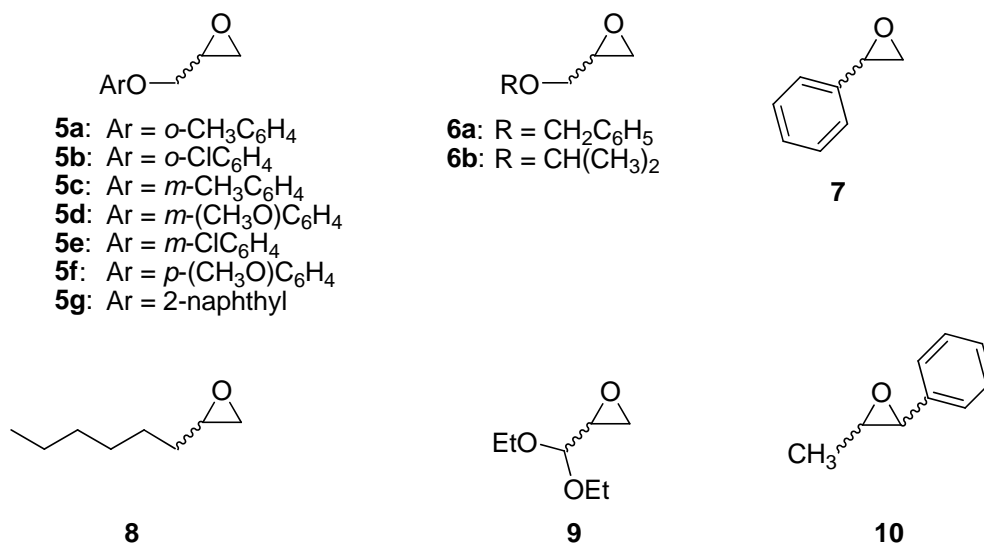
**Inhibition Studies.** At this stage we returned to the structure-based prediction that valpromide (**3**) should not be a good inhibitor of mutant LW202. Experiments in the presence of substrate (*S*)-**1** led to an estimated  $K_i$  of 2.5 mM, i.e., 10-fold weaker binding than was observed for WT ANEH. The compound has two enantiotopic *n*-propyl groups (“arms”), which in a chiral protein environment turn into diastereotopic groups. One of them corresponds to the CH<sub>2</sub>OPh substituent of the favored (*S*)-**1**, and the other to the CH<sub>2</sub>OPh substituent of the disfavored (*R*)-**1**. The *n*-propyl groups, like the CH<sub>2</sub>OPh moiety, have rotatable bonds, but the positions of the first C-atoms are spatially fixed (in the case of **1**, it is the stereogenic center). In a sense, valpromide combines stereochemical features of both enantiomers

of the actual substrate **1**. Thus, in the same way that the mutant LW202 does not (readily) accept (*R*)-**1** in a manner that ensures hydrogen bonding with Tyr251 and Tyr314 and a close distance to the nucleophilic Asp192, compound **3** is a poor inhibitor. This led us to test valeramide (**4**), which to the best of our knowledge, has not been used as an inhibitor of epoxide hydrolases.<sup>28</sup> The compound lacks one of the “arms” present in **3**, and due to the free rotation at the  $\alpha$ -carbon should be capable of rapid repositioning of the alkyl group to suit the steric requirements of the chiral enzyme pocket of mutant LW202, while maintaining the required hydrogen bonds to the two tyrosines (251/314), and Asp192. This is analogous to the accepted (*S*)-**1** and in contrast to its enantiomer (*R*)-**1**. Indeed, very efficient inhibition in the reaction of (*S*)-**1** catalyzed by LW202 was observed ( $K_i = 0.075$  mM). These observations support our previous conclusions regarding the reasons for enhanced enantioselectivity.



## Formula 2

**Substrate Scope of Mutant LW202.** One of the goals of this study was to provide the practicing chemist with guidelines as to the potential use of the mutant LW202 as a catalyst in the hydrolytic kinetic resolution of other epoxides. Our model led us to predict that LW202 should be more enantioselective than WT ANEH in reactions involving mono-substituted substrates, and that 1,2-disubstituted epoxides should react slowly, if at all. We therefore tested the mutant LW202 and WT ANEH using a number of structurally different epoxides<sup>23,33</sup> and compared the experimental data with expectations resulting from our model. A variety of mono-substituted epoxides **5–9** as well as the *trans*-disubstituted substrate **10** were used.



### Formula 3

**Table 3.** Enantioselectivity in the hydrolytic kinetic resolution of epoxides **5–10** using the WT ANEH and mutant LW202 as catalysts. The reactions of **5**, **6** and **9** are (*S*)-selective, those of **7** and **8** are (*R*)-selective (note change in CIP priority).

Substrate	<i>E</i> -value	
	WT ANEH	Mutant LW202
<b>5a</b>	5	54
<b>5b</b>	4	22
<b>5c</b>	4	31
<b>5d</b>	3	60
<b>5e</b>	4	57
<b>5f</b>	5	60
<b>5g</b>	1	42
<b>6a</b>	2.7	65
<b>6b</b>	19	90
<b>7</b>	16	50
<b>8</b>	3.5	30
<b>9</b>	16	30
<b>10</b>	<sup>a</sup> —	<sup>a</sup> —

<sup>a</sup> No conversion under the standard reaction conditions (see Supporting Information)

The results summarized in Table 3 show that a given mutant, specifically evolved for a single model compound, can in fact be used as an effective catalyst in the reaction of other substrates without resorting to additional protein engineering, although the degree of enantioselectivity is not always as high as in the original model reaction using *rac*-**1**. The traditional credo in directed evolution,<sup>1</sup> “You get what you screen for,” can be extended by the corollary, “You may get more than what you originally screened for”.<sup>32</sup> In all cases involving mono-substituted epoxides we observe, as predicted, substantially increased *E*-values relative to the performance of WT ANEH (Table 3). In the case of 1,2-epoxyoctane (**8**), additional MD simulations were carried out (WT ANEH: *d* = 3.7 Å for (*R*)-**8**, and *d* = 3.8 Å for (*S*)-**8**; LW202: *d* = 3.6 Å for (*R*)-**8** and *d* = 4.3 Å for (*S*)-**8**). On this basis enantioselectivities of about *E* = 1 (WT) and *E* = 40 (LW202) can be roughly estimated, which compares well with the experimental results (Table 3). MD simulations were also performed with styrene oxide (**7**), which also led to a prediction that mutant LW202 should be more enantioselective than WT ANEH (*E* = 40 versus *E* = 20, respectively). However, this is a special case because of the activating influence of the phenyl group, which may to some extent guide nucleophilic attack to the benzylic position of the substrate; this requires further experimental and theoretical work for a sound analysis. In the case of the *trans*-disubstituted epoxide **10**, our expectation that this substrate should not react with a reasonable rate was likewise corroborated by MD simulations. Activation by Tyr251 and Tyr314 as well as correct positioning with a *d*-value of about 3.5–4 Å proved not to be possible. Indeed, epoxide **10** was found not to react under the standard reaction conditions (Table 3; see also Supporting Information).

## Conclusions

We have analyzed, on the basis of kinetics, MD calculations, X-ray structures and molecular modeling, the underlying factors responsible for the enhanced enantioselectivity of an epoxide hydrolase mutant generated by directed evolution. The model reaction concerns the hydrolytic kinetic resolution of a chiral epoxide (glycidyl phenyl ether) catalyzed by the epoxide hydrolase from

*Aspergillus niger* (ANEH). The WT has a selectivity factor of only  $E = 4.6$  in slight favor of the (*S*)-enantiomer, while the mutant LW202 previously evolved stepwise in five stages is a highly enantioselective catalyst ( $E = 115$ ),<sup>18</sup> in this particular case matching or surpassing the performance of other epoxide hydrolases<sup>23,34</sup> and synthetic transition metal catalysts.<sup>35</sup> The kinetics reported here show that as a result of directed evolution,  $k_{\text{cat}}/K_M$  of the disfavored (*R*)-glycidyl phenyl ether is reduced much more than that of the favored (*S*)-glycidyl phenyl ether, thereby explaining the observed greatly enhanced enantioselectivity. Extensive MD simulations based on structural data for the WT enzyme identified the underlying reasons for the increasing enantioselectivity at each stage of the evolutionary process, allowing us to propose a viable model for the stereochemical course of the transformation: The preferred (*S*)-substrate is positioned ideally in the binding pocket for the reaction to occur, while the disfavored (*R*)-enantiomer is not. These conclusions were validated by molecular modeling using newly acquired X-ray structural data for both the WT and evolved mutant enzymes.

This study underscores the importance of obtaining structural data for both the WT enzyme and mutants when attempting to rationalize from the results of directed evolution. Indeed, this is the first time that X-ray structural data of an evolved enantioselective enzyme has become available, thereby allowing the kinetic results and the increase in enantioselectivity to be interpreted on a molecular level. It should be noted that the phenomenon responsible for enhanced enantioselectivity in the present case is quite different from the underlying factors identified in the case of an evolved lipase (see Introduction), where the rate of reaction of the preferred enantiomer was increased as a result of mutational changes.<sup>8-10</sup> The contrasting scenarios probably originate in the different nature of the respective screening systems; the assay of the lipase system measures rate and enantioselectivity simultaneously.<sup>8,9</sup> Another interesting observation is that although the evolved enantioselective mutant LW202 shows greater flexibility relative to the WT, thermostability is not compromised.<sup>31</sup>

The lessons learned from this directed evolution study also allows us to make reasonable predictions regarding the substrate scope of this particular evolved variant. Indeed, we have shown experimentally that it is possible to “get more than what you originally screened for”, and just as

importantly, that the model provides a practical basis for anticipating the substrate scope in which useful levels of enantioselectivity can be expected.

Finally, this research demonstrates that enantioselectivity is a sensitive probe when studying the mechanisms of enzymes. In the present case, the basic features of the mechanism of epoxide hydrolases had already been established (Scheme 1),<sup>20,22,23,24b,34</sup> but now the fine details have been revealed. Knowledge of the precise positioning of the reacting substrate in the binding pocket of epoxide hydrolases is crucial for a deeper understanding of how these enzymes function.

**Acknowledgement.** Dedicated to Professor Walter Thiel on the occasion of his 60th birthday. We thank the Deutsche Forschungsgemeinschaft (Schwerpunktprogramm 1170 “Directed Evolution to Optimize and Understand Molecular Biocatalysis”) and the Fonds der Chemischen Industrie generous support. Support from the Swedish Research Council (VR) to SM is also gratefully acknowledged.

**Supporting Information Available:** Details of kinetic and computational studies, inhibition experiments and X-ray structural data as well as standard reaction conditions. The atomic coordinates and structure factors have been deposited at the Protein Data Bank, [www.pdb.org](http://www.pdb.org) [PDB ID codes 3G0I and 3G0J].

## References

- (1) Reviews of directed evolution: (a) Lutz, S.; Bornscheuer, U. T. *Protein Engineering Handbook*; Wiley-VCH: Weinheim, 2009, Vols. 1 and 2. (b) Arndt, K. M.; Müller, K. M. *Protein Engineering Protocols (Methods in Molecular Biology)*; Humana Press: Totowa, New Jersey, 2007; Vol. 352. (c) Arnold, F. H.; Georgiou, G. *Directed Enzyme Evolution: Screening and Selection Methods*; Humana Press: Totowa, New Jersey, 2003; Vol. 230. (d) Jäckel, C.; Kast, P.; Hilvert, D. *Annu. Rev. Biophys.* **2008**, *37*, 153–173. (e) Brakmann, S.; Schwienhorst, A. *Evolutionary Methods in Biotechnology - Clever Tricks for Directed Evolution*; Wiley-VCH:



- Weinheim, Germany, 2004. (f) Hibbert, E. G.; Baganz, F.; Hailes, H. C.; Ward, J. M.; Lye, G. J.; Woodley, J. M.; Dalby, P. A. *Biomol. Eng.* **2005**, *22*, 11–19. (g) Rubin-Pitel, S. B.; Zhao, H. *Comb. Chem. High Throughput Screening* **2006**, *9*, 247–257. (h) Kaur, J.; Sharma, R. *Crit. Rev. Biotechnol.* **2006**, *26*, 165–199. (i) Bershtein, S.; Tawfik, D. S. *Curr. Opin. Chem. Biol.* **2008**, *12*, 151–158.
- (2) (a) Chen, K.; Arnold, F. H. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 5618–5622. (b) Arnold, F. H. *Acc. Chem. Res.* **1998**, *31*, 125–131.
- (3) Reviews of directed evolution of thermostability: (a) Eijsink, V. G. H.; Gåseidnes, S.; Borchert, T. V.; van den Burg, B. *Biomol. Eng.* **2005**, *22*, 21–30. (b) Wintrode, P. L.; Arnold, F. H. *Adv. Protein Chem.* **2001**, *55*, 161–225. (c) O’Fagain, C. *Enzyme Microb. Technol.* **2003**, *33*, 137–149. (d) Bommarius, A. S.; Broering, J. M. *Biocatal. Biotransform.* **2005**, *23*, 125–139.
- (4) Reviews of directed evolution of enantioselectivity: (a) Reetz, M. T. *Proc. Natl. Acad. Sci. U.S.A.* **2004**, *101*, 5716–5722. (b) Reetz, M. T. In *Advances in Catalysis*; Gates, B. C., Knözinger, H., Eds.; Elsevier Academic Press: San Diego, California, 2006; Vol. 49, pp 1–69. (c) Reetz, M. T. In *Asymmetric Organic Synthesis with Enzymes*; Gotor, V., Alfonso, I., García-Urdiales, E., Eds.; Wiley-VCH: Weinheim, 2008; pp 21–63. (d) Reetz, M. T. In *Protein Engineering Handbook*; Lutz, S.; Bornscheuer, U. T., Eds.; Wiley-VCH: Weinheim, 2009; Vol. 2, pp 409–439.
- (5) Lutz, S.; Patrick, W. M. *Curr. Opin. Biotechnol.* **2004**, *15*, 291–297.
- (6) Articles stressing the need for high-quality libraries in directed evolution: (a) Fox, R. J.; Huisman, G. W. *Trends Biotechnol.* **2008**, *26*, 132–138. (b) Herman, A.; Tawfik, D. S. *Prot. Eng., Des. Sel.* **2007**, *20*, 219–226. (c) Bloom, J. D.; Meyer, M. M.; Meinhold, P.; Otey, C. R.; MacMillan, D.; Arnold, F. H. *Curr. Opin. Struct. Biol.* **2005**, *15*, 447–452. (d) Wong, T. S.; Roccatano, D.; Zacharias, M.; Schwaneberg, U. *J. Mol. Biol.* **2006**, *355*, 858–871. (e) Reetz, M. T.; Kahakeaw, D.; Sanchis, J. *Mol. BioSyst.* **2009**, *5*, 115–122.

- (7) (a) Jaenicke, R.; Böhm, G. *Curr. Opin. Struct. Biol.* **1998**, *8*, 738–748. (b) Sterner, R.; Liebl, W. *Crit. Rev. Biochem. Mol. Biol.* **2001**, *36*, 39–106. (c) Buchner, J.; Kiefhaber, T. *Protein Folding Handbook*; Wiley-VCH: Weinheim, 2005. (d) Atomi, H. *Curr. Opin. Biotechnol.* **2005**, *9*, 166–173. (e) Oshima, T. *Curr. Opin. Struct. Biol.* **1994**, *4*, 623–628. (f) Eijssink, V. G. H.; Bjork, A.; Gaseidnes, S.; Sirevag, R.; Synstad, B.; van den Burg, B.; Vriend, G. *J. Biotechnol.* **2004**, *113*, 105–120.
- (8) (a) Reetz, M. T.; Zonta, A.; Schimossek, K.; Liebeton, K.; Jaeger, K.-E. *Angew. Chem.* **1997**, *109*, 2961–2963; *Angew. Chem., Int. Ed. Engl.* **1997**, *36*, 2830–2832. (b) Reetz, M. T. *Pure Appl. Chem.* **1999**, *71*, 1503–1509.
- (9) Reetz, M. T.; Wilensek, S.; Zha, D.; Jaeger, K.-E. *Angew. Chem.* **2001**, *113*, 3701–3703; *Angew. Chem. Int. Ed.* **2001**, *40*, 3589–3591.
- (10) (a) Bocola, M.; Otte, N.; Jaeger, K.-E.; Reetz, M. T.; Thiel, W. *ChemBioChem* **2004**, *5*, 214–223. (b) Reetz, M. T.; Puls, M.; Carballeira, J. D.; Vogel, A.; Jaeger, K.-E.; Eggert, T.; Thiel, W.; Bocola, M.; Otte, N. *ChemBioChem* **2007**, *8*, 106–112.
- (11) Bolt, A.; Berry, A.; Nelson, A. *Arch. Biochem. Biophys.* **2008**, *474*, 318–330.
- (12) (a) Wada, M.; Hsu, C.-C.; Franke, D.; Mitchell, M.; Heine, A.; Wilson, I.; Wong, C.-H. *Bioorg. Med. Chem.* **2003**, *11*, 2091–2098. (b) Hsu, C.-C.; Hong, Z.; Wada, M.; Franke, D.; Wong, C.-H. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 9122–9126.
- (13) For select examples of directed evolution of enantioselective enzymes<sup>4,8-10,14</sup>, see: (a) Hill, C. M.; Li, W.-S.; Thoden, J. B.; Holden, H. M.; Raushel, F. M. *J. Am. Chem. Soc.* **2003**, *125*, 8990–8991. (b) Koga, Y.; Kato, K.; Nakano, H.; Yamane, T. *J. Mol. Biol.* **2003**, *331*, 585–592. (c) May, O.; Nguyen, P. T.; Arnold, F. H. *Nat. Biotechnol.* **2000**, *18*, 317–320. (d) 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. (e) Turner, N. J. *Trends Biotechnol.*

**2003**, *21*, 474–478. (f) Lipovsek, D.; Antipov, E.; Armstrong, K. A.; Olsen, M. J.; Klibanov, A. M.; Tidor, B.; Wittrup, K.D. *Chem. Biol.* **2007**, *14*, 1176–1185. (g) Boersma, Y. L.; Dröge, M. J.; van der Sloot, A. M.; Pijning, T.; Cool, R. H.; Dijkstra, B. W.; Quax, W. J. *ChemBioChem.* **2008**, *9*, 1110–1115. (h) Reetz, M. T.; Höbenreich, H.; Soni, P.; Fernández, L. *Chem. Commun.* **2008**, 5502–5504.

(14) Select examples of CASTing in directed evolution of enantioselective enzymes<sup>16,18,19,27</sup>: (a) Bartsch, S.; Kourist, R.; Bornscheuer, U. T. *Angew. Chem.* **2008**, *120*, 1531–1534; *Angew. Chem. Int. Ed.* **2008**, *47*, 1508–1511. (b) Liang, L.; Zhang, J.; Lin, Z. *Microb. Cell Fact.* **2007**, *6*:36. (c) Reetz, M. T.; Peyralans, J. J.-P.; Maichele, A.; Fu, Y.; Maywald, M. *Chem. Commun. (Cambridge, U. K.)* **2006**, 4318–4320. (d) Reetz, M. T.; Wu, S. *Chem. Commun.* **2008**, 5499–5501.

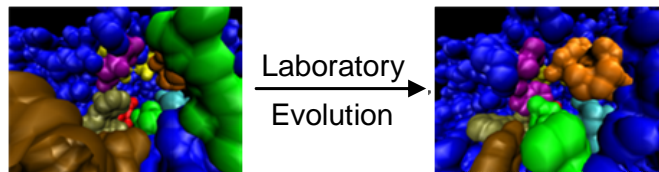
(15) For example, numerous theoretical studies regarding the source of enantioselectivity of WT lipases have been reported, e.g.: (a) Zuegg J.; Höning H.; Schrag J. D.; Cygler M. *J. Mol. Catal. B: Enzym.* **1997**, *3*, 83–98. (b) Scheib H.; Pleiss J.; Stadler P.; Kovac A.; Potthoff A. P.; Haalck L.; Spener F.; Paltauf F.; Schmid R. D.; *Protein Eng.* **1998**, *11*, 675–682. (c) Häffner F.; Norin T.; Hult K. *Biophys. J.* **1998**, *74*, 1251–1262. (d) Schulz, T.; Schmid, R. D.; Pleiss, J. *J. Mol. Model.* **2001**, *7*, 265–270. (e) Tafi A.; van Almsick A.; Corelli F.; Crusco M.; Laumen K. E.; Schneider M. P.; Botta M. *J. Org. Chem.* **2000**, *65*, 3659–3665. (f) Kazlauskas, R. *Science* **2001**, *293*, 2277–2279. (g) Raza S.; Fransson L.; Hult K. *Protein Sci.* **2001**, *1*, 329–338. (h) Tomić, S.; Kojić-Prčić, B. *J. Mol. Graph. Model.* **2002**, *21*, 241–252. (i) Bocola M.; Stubbs M. T.; Sotriffer C.; Hauer B.; Friedrich T.; Dittrich K.; Klebe G. *Protein Eng.* **2003**, *16*, 319–322. (j) Guieysse, D.; Salagnad, C.; Monsan, P.; Remaud-Simeon, M.; Tran, V. **2003**, *14*, 1807–181.

- (16) (a) Reetz, M. T.; Bocola, M.; Carballeira, J. D.; Zha, D.; Vogel, A. *Angew. Chem.* **2005**, *117*, 4264–4268; *Angew. Chem. Int. Ed.* **2005**, *44*, 4192–4196. (b) Reetz, M. T.; Carballeira, J. D.; Peyralans, J. J.-P.; Höbenreich, H.; Maichele, A.; Vogel, A. *Chem.–Eur. J.* **2006**, *12*, 6031–6038.
- (17) Examples of focused libraries in directed evolution:<sup>1,13,14</sup> (a) Horsman, G. P.; Liu, A. M. F.; Henke, E.; Bornscheuer, U. T.; Kazlauskas, R. J. *Chem.–Eur. J.* **2003**, *9*, 1933–1939. (b) Dube, D. K.; Loeb, L. A. *Biochemistry* **1989**, *28*, 5703–5707. (c) Warren, M. S.; Benkovic, S. J. *Protein Eng.* **1997**, *10*, 63–68. (d) Yang, J.; Koga, Y.; Nakano, H.; Yamane, T. *Protein Eng.* **2002**, *15*, 147–152. (e) Rui, L.; Cao, L.; Chen, W.; Reardon, K. F.; Wood, T. K. *J. Biol. Chem.* **2004**, *279*, 46810–46817. (f) Park, S.; Morley, K. L.; Horsman, G. P.; Holmquist, M.; Hult, K.; Kazlauskas, R. J. *Chem. Biol.* **2005**, *12*, 45–54. (g) Nowlan, C.; Li, Y.; Hermann, J. C.; Evans, T.; Carpenter, J.; Ghanem, E.; Shoichet, B. K.; Raushel, F. M. *J. Am. Chem. Soc.* **2006**, *128*, 15892–15902. (h) Miyazaki, K.; Arnold, F. H. *J. Mol. Evol.* **1999**, *49*, 716–720. (i) Valinger, G.; Hermann, M.; Wagner, U. G.; Schwab, H. *J. Biotechnol.* **2007**, *129*, 98–108. (j) Chelliserrykattil, J.; Ellington, A. D. *Nat. Biotechnol.* **2004**, *22*, 1155–1160. (k) Gabor, E. M.; Janssen, D. B. *Prot. Eng., Des. Sel.* **2004**, *17*, 571–579. (l) Fa, M.; Radeghieri, A.; Henry, A. A.; Romesberg, F. E. *J. Am. Chem. Soc.* **2004**, *126*, 1748–1754. (m) Juillerat, A.; Gronemeyer, T.; Keppler, A.; Gendreizig, S.; Pick, H.; Vogel, H.; Johnsson, K. *Chem. Biol.* **2003**, *10*, 313–317. (n) Nair, N. U.; Zhao, H. *ChemBioChem* **2008**, *9*, 1213–1215.
- (18) Reetz, M. T.; Wang, L.-W.; in part Bocola, M. *Angew. Chem.* **2006**, *118*, 1258–1263; *Erratum* 2556; *Angew. Chem. Int. Ed.* **2006**, *45*, 1236–1241; *Erratum* 2494.
- (19) Reetz, M. T.; Kahakeaw, D.; Lohmer, R. *ChemBioChem* **2008**, *9*, 1797–1804.
- (20) (a) Zou, J. Y.; Hallberg, B. M.; Bergfors, T.; Oesch, F.; Arand, M.; Mowbray, S. L.; Jones, T. A. *Structure (London)* **2000**, *8*, 111–122. (b) <http://www.pdb.ccdc.cam.ac.uk/pdb>.

- (21) (a) Chen, C.-S.; Fujimoto, Y.; Girdaukas, G.; Sih, C. J. *J. Am. Chem. Soc.* **1982**, *104*, 7294–7299.  
(b) Faber, K. *Biotransformations in Organic Chemistry*; 5th ed.; Springer: Berlin, 2004.
- (22) (a) Morisseau, C.; Archelas, A.; Guitton, C.; Faucher, D.; Furstoss, R.; Baratti, J. C. *Eur. J. Biochem.* **1999**, *263*, 386–395. (b) Arand, M.; Hemmer, H.; Baratti, J.; Archelas, A.; Furstoss, R.; Oesch, F. *Biochem. J.* **1999**, *344*, 273–280.
- (23) Reviews of epoxide hydrolases: (a) Archelas, A.; Furstoss, R. *Curr. Opin. Chem. Biol.* **2001**, *5*, 112–119. (b) Morisseau, C.; Hammock, B. D. *Annu. Rev. Pharmacol. Toxicol.* **2005**, *45*, 311–333.  
(c) Faber, K.; Orru, R. V. A. In *Enzyme Catalysis in Organic Synthesis*; Drauz, K.; Waldmann, H., Eds.; 2nd ed.; Wiley-VCH: Weinheim, 2002; pp 579–608.
- (24) For a review of the theory of near-attack-conformers in enzyme-catalyzed reactions, see: (a) Bruice, T. C. *Acc. Chem. Res.* **2002**, *35*, 139–148. For an MD-study of epoxide hydrolases, see: (b) Schiøtt, B.; Bruice, T. C. *J. Am. Chem. Soc.* **2002**, *124*, 14558–14570.
- (25) Gerber P. R. *J. Comput.-Aided Mol. Des.* **1998**, *12*, 37–51.
- (26) Gerber P. R., Müller K. *J. Comput.-Aided Mol. Des.* **1995**, *9*, 251–268.
- (27) Reetz, M. T.; Sanchis, J. *ChemBioChem* **2008**, *9*, 2260–2267.
- (28) Zollner, H.; *Handbook of Enzyme Inhibitors*; 3rd ed.; VCH: Weinheim, 1999; pp 554–557.
- (29) Read, R. J. *Acta Crystallogr., Sect. A: Found. Crystallogr.* **1986**, *A42*, 140–149.
- (30) Mowbray, S. L.; Elfström, L. T.; Ahlgren, K. M.; Andersson, C. E.; Widersten, M. *Protein Sci.* **2006**, *15*, 1628–1637.
- (31) Gumulya, Y.; Reetz, M. T. unpublished results.
- (32) For a report describing wide substrate acceptance and high enantioselectivity of a Baeyer-Villigerase mutant, originally evolved for a single substrate, see: Mihovilovic, M. D.; Rudroff, F.; Winninger, A.; Schneider, T.; Schulz, F.; Reetz, M. T. *Org. Lett.* **2006**, *8*, 1221–1224.
- (33) Manoj, K. M.; Archelas, A.; Baratti, J.; Furstoss, R. *Tetrahedron* **2001**, *57*, 695–701.

- (34) For studies concerning directed evolution or site-specific mutagenesis of other epoxide hydrolases, see: (a) van Loo, B.; Spelberg, J. H. L.; Kingma, J.; Sonke, T.; Wubbolts, M. G.; Janssen, D. B. *Chem. Biol.* **2004**, *11*, 981–990. (b) Rui, L.; Cao, L.; Reardon, K. F.; Wood, T. K. *Appl. Environ. Microbiol.* **2005**, *71*, 3995–4003. (c) Kotik, M.; Štěpánek, V.; Kyslík, P.; Marešová, H. *J. Biotechnol.* **2007**, *132*, 8–15. (d) Lutje Spelberg, J. H.; Rink, R.; Archelas, A.; Furstoss, R.; Janssen, D. B. *Adv. Synth. Catal.* **2002**, *344*, 980–985. (e) Lee, E. Y.; Shuler, M. L. *Biotechnol. Bioeng.* **2007**, *98*, 318–327.
- (35) (a) Jacobsen, E. N. *Acc. Chem. Res.* **2000**, *33*, 421–431. (b) Schaus, S. E.; Brandes, B. D.; Larrow, J. F.; Tokunaga, M.; Hansen, K. B.; Gould, A. E.; Furrow, M. E.; Jacobsen, E. N. *J. Am. Chem. Soc.* **2002**, *124*, 1307–1315. (c) Kawthekar, R. B.; Kim, G.-J. *Helv. Chim. Acta* **2008**, *91*, 317–332.

## Table of Contents Graphic



## Supporting Information

# Directed Evolution of an Enantioselective Epoxide Hydrolase: Uncovering the Source of Enantioselectivity at Each Evolutionary Stage

Manfred T. Reetz,<sup>\*,†</sup> Marco Bocola,<sup>\*,‡</sup> Li-Wen Wang,<sup>†</sup> Joaquin Sanchis,<sup>†</sup>  
Annette Cronin,<sup>§</sup> Michael Arand,<sup>§</sup> Jinyu Zou,<sup>\*\*</sup> Alain Archelas,<sup>††</sup> Anne-Lise Bottalla,<sup>††</sup>  
Agata Naworyta<sup>‡‡</sup> and Sherry L. Mowbray<sup>\*,‡‡</sup>

<sup>†</sup> Contribution from the Max-Planck-Institut für Kohlenforschung, Kaiser-Wilhelm-Platz 1, 45470 Mülheim/Ruhr, Germany

<sup>‡</sup> Present address: Institute of Biophysics and Physical Biochemistry, University of Regensburg, Universitätsstrasse 31, 93053 Regensburg, Germany

<sup>§</sup> Institute of Pharmacology and Toxicology, Universität Zürich Winterthurerstrasse 190, 8057 Zürich, Switzerland

<sup>\*\*</sup> Department of Cell and Molecular Biology, Uppsala University, Box 596, Biomedical Center, 75124 Uppsala, Sweden

<sup>††</sup> University Paul-Cézanne/Faculté Sciences et Techniques de St Jérôme, case 432 Avenue Escadrille Normandie Niemen, 13397 Marseille Cedex 20, France

<sup>‡‡</sup> Department of Molecular Biology, Swedish University of Agricultural Sciences, Box 590, Biomedical Center, 75124 Uppsala, Sweden

### Cloning

*Aspergillus niger* strain LCP521 (EMBL Accession Number AJ238460) was the source of enzyme (ANEH) used in the current study, as described in Zou et al, 2000.<sup>1</sup> For the mutant proteins, the stop codon of pQEEH was mutated to GGG (glycine) to allow transcription to continue into the hexa-His coding sequence of the pQE60 vector. Directed mutagenesis was performed according to the QuikChange Site-Directed Mutagenesis® method using the mutagenic primers GGAGAAATTAACATGTCCGCTCTCATGAATC and GGTGATGAGATCTACCCTTCTGCCACACC (Invitrogen, Karlsruhe, Germany).



## Purification of His tagged ANEH

*E. coli* SG13009 (Qiagen, Les Ulis, France) was transformed with plasmid extracted from a single colony culture of DH5a. Selected transformants were grown and the enzymes produced as described in the Qiaexpress™ System manual (Qiagen). Overnight cultures were harvested and cells suspended in 20 mM sodium phosphate, pH 7.4, 20% glycerol and 500 mM NaCl. Bacterial lysis was performed using a French press. Clarified lysate was first filtered using a 0.22 mm filter and then loaded on 1 mL HisTrap™ HP affinity column (GE Healthcare, Uppsala, Sweden). His-tagged enzyme was eluted with an imidazole gradient using HPLC and the purity checked on SDS/PAGE. The resulting solution was further passed over a Superdex™ 200 10/300 GL (GE Healthcare) gel-filtration column previously equilibrated with 10 mM Tris-HCl, 20 mM NaCl and 1 mM EDTA buffer with the final pH 7.4. The *A. niger* epoxide hydrolase was eluted as a single homogeneous peak. The purified protein was analyzed by SDS/PAGE, followed by Coomassie Blue staining and immunoblotting.

## Typical procedure for hydrolytic kinetic resolution

In a 50 mL reactor, the appropriate amount of epoxide hydrolase (WT ANEH (0.1–1.25U against styrene oxide)\* or mutant LW202 (0.5–7U against styrene oxide)\*) was dissolved in a phosphate buffer (10 mL, 50 mM, pH7). Pure racemic epoxide (e.g., **5**, **6** or **9**), dissolved in ethanol (500µL), was then added (final epoxide concentration about 10 mM), and the mixture was stirred at 25 °C for about 1–2 hours. The kinetic resolution was followed by withdrawing aliquots (500 µL) from the reaction medium at different time intervals. These were saturated with NaCl and extracted with ethyl acetate (500 µL). The enantiomeric excess of the remaining epoxide and the formed diol were directly determined by chiral GC or HPLC analysis, from which the *E*-values were determined. For the resolution of **7** and **8**, the same procedure was used except that the substrate concentration was lower (final epoxide concentration 4–6 mM). Epoxide **10** did not show any turnover under these conditions.

\*) Mateo, C.; Archelas, A.; Furstoss, R. *Anal. Biochem.* **2003**, *314*, 135–141.

## Kinetic analysis

A closed reaction vessel containing 4 mL sodium phosphate buffer (20 mM, pH 7.2) and 0.75 mM (*S*)-**1** was incubated 30 min at 30 °C with stirring. The purified enzyme was added to a concentration that allowed the pre-steady state of the reaction to be measurable, i.e. conversion of a maximum of 10% of the substrate. The reaction was monitored by periodically removing samples. The 100  $\mu$ L samples were added to the same volume of the internal standard solutions ((*R*)-(+)-1-phenyl-1-butanol, 6.66 mM in methanol). Enantioselectivity was measured by HPLC using a Chiralcel OD-R chiral column from Daicel Chemical Industries (Tokyo, Japan) with methanol:water (7:3, flow rate 0.5 mL/min). Retention time: (*R*)-**2**,  $t_R = 8.6$  min; (*S*)-**2**,  $t_S = 9.8$  min; (*R*)-**1**,  $t_R = 20.9$  min; (*S*)-**1**,  $t_S = 24.4$  min. Similar experiments were carried out for substrate concentrations varying from 0.25 to 5 mM. The entire process was repeated with (*R*)-**1** and *rac*-**1**. From the time-course variation of the conversions, the initial rates for each enantiomer were estimated and used for the kinetic analysis.

### Inhibition assays

Enzymatic activity was measured with pure enzyme ( $10^{-3}$  mg/mL for (*R*)-**1**;  $10^{-4}$  mg/mL for (*S*)-**1**) at 30 °C in sodium phosphate buffer (20 mM, pH = 7.2, 3.7 mL) with acetonitrile (7.15%, 0.3 mL) and a defined concentration of substrate, either (*S*)-**1** or (*R*)-**1** (1.25 mM). Experiments were run in the presence of inhibitor, either **3** or **4** (at concentrations of 0, 0.3, 0.6, 1.25, 6.25 and 12.5 mM). Samples (taken after 0, 15, 30, 60, 90, 120, 180 and 240 min) were measured by HPLC as described above. The results are summarized in Figure 1. Initial reaction rates were used to calculate the  $I_{50}$ . In the case of (*S*)-**1**, where  $[(S)\text{-}1] = 1.25 \text{ mM} = K_m$ , then  $I_{50} = 2 K_i$ , according to Brandt.<sup>2</sup> Inhibition with **3**: LW202 and (*S*)-**1**,  $I_{50} \approx 5 \text{ mM}$ ,  $K_i \approx 2.5 \text{ mM}$ ; LW202 and (*R*)-**1**,  $I_{50} \approx 5.5 \text{ mM}$ . Inhibition with **4**: LW202 and (*S*)-**1**,  $I_{50} \approx 0.15 \text{ mM}$ ,  $K_i \approx 0.07 \text{ mM}$ .

### X-ray crystallographic studies

Crystals of WT ANEH were obtained by sitting drop vapor diffusion at room temperature, as follows: 2  $\mu$ L of protein solution (17 mg/ml in 10 mM Tris-HCl, pH 7.4, 1 mM EDTA, 20 mM NaCl and 0.02% sodium azide) were mixed with 2  $\mu$ L of the reservoir (20% PEG6000, 0.1 M MES, pH 6.0, and 0.1 M unbuffered sodium acetate). Crystals ( $0.3 \times 0.4 \times 0.02 \text{ mm}^3$ ) grew within one week. These crystals were soaked for 4 days in a cryoprotectant solution containing 10 mM valpromide, before flash-

cooling in liquid nitrogen. Data were collected at beamline ID14:4 at the European Synchrotron Radiation Facility (ESRF), Grenoble, and indexed and scaled with MOSFLM<sup>3</sup> and SCALA.<sup>4</sup> The protein model from the previously reported structure (PDB entry code 1QO7)<sup>5,1</sup> was used as the starting point for alternating cycles of rebuilding<sup>6</sup> and refinement.<sup>7</sup> Statistics for data processing and the final model are reported in Table 1. The model and experimental diffraction data are deposited at the PDB with entry code 3G0I.

The LW202 mutant protein was crystallized by vapor diffusion in hanging drops of 2  $\mu$ L protein (10 mg/mL in the final purification buffer) and 2  $\mu$ L reservoir solution (15% PEG3350, 0.25 M ammonium formate) at room temperature. Crystals ( $\sim 0.3 \times 0.4 \times 0.8$  mm<sup>3</sup>) appeared within 5 days. Before being flash-cooled in liquid nitrogen, crystals were transferred to a cryoprotectant composed of 20% glycerol, 15% PEG3350, and 0.25 M ammonium formate. X-ray data were collected at beamline ID14:2 at the ESRF. Diffraction data were indexed using as described above. The structure was solved by molecular replacement with MOLREP,<sup>8</sup> using a search model including all protein atoms from the earlier wild-type structure.<sup>1</sup> Two molecules were found in the asymmetric unit. The structure was refined by alternating cycles of REFMAC refinement and model rebuilding in O. Statistics for data processing and the final model are reported in Table 1. The model and experimental diffraction data are deposited at the PDB with entry code 3G02.

### Preparation of figures

Structure figures were prepared using PyMOL (<http://www.pymol.org>). Figures 8 and 9 were prepared with VMD.<sup>9</sup>

### References:

- (1) Zou, J. Y.; Hallberg, B. M.; Bergfors, T.; Oesch, F.; Arand, M.; Mowbray, S. L.; Jones, T. A. *Structure* **2000**, *8*, 111–122.
- (2) Brandt, R. B.; Laux, J. E.; Yates, S. W. *Biochem. Med. Metab. Biol.* **1987**, *37*, 344–349.
- (3) Leslie, A. G. W. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **1999**, *D55*, 1696–1702.
- (4) Evans, P. R. In *Proceedings of CCP4 Study Weekend on Data Collection & Processing*; Sawyer, L.; Isaac, N.; Nailey, S., Eds.; SERC Daresbury Laboratory: Warrington, 1993; pp 114–122.

- (5) Berman, H. M.; Westbrook, J.; Feng, Z.; Gilliland, G.; Bhat, T. N.; Weissig, H.; Shindyalov, I. N.; Bourne, P. E. *Nucl. Acids Res.* **2000**, *28*, 235–242.
- (6) Jones, T. A.; Zou, J. Y.; Cowan, S. W.; Kjeldgaard, M. *Acta Crystallogr., Sect. A: Found. Crystallogr.* **1991**, *A47*, 110–119.
- (7) Murshudov, G. N.; Vagin, A. A.; Lebedev, A.; Wilson, K. S.; Dodson, E. J. *Acta Crystallogr., Sect. D: Biol. Crystallogr.* **1999**, *55*, 247–255.
- (8) Vagin, A. A.; Teplyakov, A. *J. Appl. Crystallogr.* **1997**, *30*, 1022–1025.
- (9) Humphrey, W.; Dalke, A.; Schulten, K. *J. Mol. Graphics Modell.* **1996**, *14*, 33–38.
- (10) Kleywegt, G. J.; Jones, T. A. *Structure* **1996**, *4*, 1395–1400.
- (11) Engh, R. A.; Huber, R. *Acta Crystallogr., Sect. A: Found. Crystallogr.* **1991**, *A47*, 392–400.

**Table S1.** *Data Collection and Refinement Statistics.* The space group in each case is P2<sub>1</sub>. Information in parentheses refers to the highest resolution shell

<b>Data collection statistics</b>	LW202 mutant	WT-VLP
Data collection beamline/detector	ESRF ID14:2/ADSC scanner	ESRF ID14:4/ADSC scanner
Cell axial lengths (Å)	61.8, 89.7, 75.3 (β = 104.8°)	63.0, 89.7, 75.8 (β = 105.3°)
Resolution range (Å)	30.54 – 1.5 (1.58 – 1.5)	40.0 – 2.1 (2.14 – 2.1)
Number of reflections measured	209,428	171,672
Number of unique reflections	120,696	47,576
Average multiplicity	1.7 (1.7)	3.6 (2.7)
Completeness (%)	95.5 (98.2)	99.5 (92.8)
R <sub>meas</sub> (%)	2.7 (12.8)	6.6 (21.3)
<<I>/<I> <sub>0</sub> >	15.8 (7.2)	13.7 (4.3)
<b>Refinement statistics</b>		
Resolution range (Å)	30.0 – 1.5	38.2 – 2.1
No. of reflections used in working set	114,575	46,082
No. of reflections for R <sub>free</sub> calculation	5728	2304
R-value, R <sub>free</sub> (%)	18.3, 20.9	17.5, 22.1
No. of non-hydrogen atoms	7102	6696
No. of solvent waters	707	457
Mean B-factor, protein atoms, A and B molecules (Å <sup>2</sup> )	13.7, 13.9	13.6, 13.9
Mean B-factor, solvent atoms (Å <sup>2</sup> )	28.0	18.6
Mean B-factor, valpromide atoms (Å <sup>2</sup> )	-	17.6
Mean B-factor, formic acid (Å <sup>2</sup> )	11.4	-
Ramachandran plot outliers (%) <sup>a</sup>	1.6	1.5
r.m.s. deviation from ideal bond length (Å) <sup>b</sup>	0.006	0.010
r.m.s. deviation from ideal bond angle (°)	1.01	1.19

<sup>a</sup> Calculated using a strict-boundary Ramachandran plot<sup>10</sup>

<sup>b</sup> Using the parameters of Engh and Huber<sup>11</sup>

**Figure S1.** Determination of  $I_{50}$ s. a) Inhibition of LW202 by **3** in the hydrolysis of (*R*)-**1** at 30 °C and pH = 7.2.  $[E] = 10^{-3}$  mg/mL,  $[(R)\text{-1}]_0 = 1.25$  mM. b) Inhibition of LW202 by **3** in the hydrolysis of (*S*)-**1** at 30 °C and pH = 7.2.  $[E] = 10^{-4}$  mg/mL,  $[(S)\text{-1}]_0 = 1.25$  mM. c) Inhibition of LW202 by **4** in the hydrolysis of (*S*)-**1** at 30 °C and pH = 7.2.  $[E] = 10^{-4}$  mg/mL,  $[(S)\text{-1}]_0 = 1.25$  mM. d) An enlargement of panel c used for the graphical determination of the  $I_{50}$  value.

