

# Prediction of the Solvent Dependence of Enzymatic Prochiral Selectivity by Means of Structure-Based Thermodynamic Calculations

Tao Ke, Charles R. Wescott,<sup>†</sup> and Alexander M. Klibanov\*

Contribution from the Department of Chemistry, Massachusetts Institute of Technology, Cambridge, Massachusetts 02139

Received August 7, 1995<sup>⊗</sup>

**Abstract:** A new, quantitative model is elaborated to rationalize the solvent dependence of enzymatic selectivity solely on the basis of the thermodynamics of substrate solvation. The model predicts that any type of the selectivity (defined as the ratio of  $k_{\text{cat}}/K_{\text{M}}$  values) should be proportional to the ratio of the thermodynamic activity coefficients of the desolvated portions of the substrate(s) in the relevant transition state of the enzymatic reaction. The latter ratio is calculated by (i) determining the desolvated portion of the substrate in the transition state using molecular modeling based on the crystal structure of the enzyme, (ii) approximating this desolvated portion of the substrate by a distinct model compound, and (iii) calculating the activity coefficient of this model compound using the UNIFAC computer algorithm. In this study, the developed general model has been applied to, and verified with, prochiral selectivity of enzymes. Crystals (lightly cross-linked with glutaraldehyde) of  $\gamma$ -chymotrypsin or subtilisin Carlsberg used as asymmetric catalysts in organic solvents almost quantitatively adhere to the predictions of the model in the acetylation of 2-substituted 1,3-propanediols. In contrast, little agreement between the predicted and observed solvent dependences of the prochiral selectivity has been obtained with lyophilized or acetone-precipitated chymotrypsin, thus confirming that the enzyme structure in such preparations (but not in crystals) is non-native.

## Introduction

One of the most profound revelations arisen from nonaqueous enzymology<sup>1</sup> is the discovery that the specificity of an enzyme strongly depends on the solvent.<sup>2</sup> Of all the types of enzyme specificity found to be controlled by the solvent<sup>2</sup>—enantioselectivity, prochiral selectivity, substrate specificity, regioselectivity, and chemoselectivity—the first two are particularly important for synthetic applications. Indeed, if generalized and understood, solvent control of enzymatic stereoselectivity should enhance the utility of biocatalysis in organic chemistry by allowing the rational manipulation of the stereochemical outcome of asymmetric transformations simply by altering the reaction medium. The ultimate challenge in this regard is to learn how to predict enzyme selectivity as a function of the solvent.

As a first step toward this goal, we have recently elaborated a thermodynamic model which explains the substrate specificity of the protease subtilisin Carlsberg in organic solvents on the basis of solvent-to-water partition coefficients of the substrates.<sup>3</sup> These partition coefficients can be either measured experimentally or calculated using the UNIFAC computer algorithm.<sup>4</sup> An explicit assumption of our analysis is that the substrates are fully desolvated, i.e., inaccessible to the solvent, in the enzyme-bound

transition state. This assumption precludes the extension of the proposed model to enantioselectivity, since the partition coefficients for different enantiomers of the same compound are identical. Likewise, prochiral, regio-, and chemo-selectivities cannot be analyzed either, because in all these instances the same substrate molecule (just different parts of it) reacts with the enzyme.

In the present study, we further develop and broaden our thermodynamic treatment to eliminate the aforementioned limitations. The resultant model, tested herein with prochiral selectivity, takes into account variations in substrate desolvation in the transition states for *pro-R* and *pro-S* orientations. Thermodynamic activity coefficients of the desolvated portions of the substrate, calculated using computer-generated, transition-state structures and UNIFAC, correctly predict the solvent dependence of prochiral selectivity of crystalline chymotrypsin and subtilisin.

## Theory

The solvent may influence enzymatic selectivity through several distinct mechanisms. For instance, it could change the

(4) (a) Fredenslund, A.; Gmehling, J.; Rasmussen, P. *Vapor-Liquid Equilibria Using UNIFAC*; Elsevier: New York, 1977. (b) Steen, S.-J.; Bärbel, K.; Gmehling, J.; Rasmussen, P. *Ind. Eng. Chem. Process Des. Dev.* **1979**, *18*, 714. (c) Rasmussen, P.; Fredenslund, A. *Ind. Eng. Chem. Process Des. Dev.* **1982**, *21*, 118. (d) Macedo, E. A.; Weidlich, U.; Gmehling, J.; Rasmussen, P. *Ind. Eng. Chem. Process Des. Dev.* **1983**, *22*, 678. (e) Teigs, D.; Gmehling, J.; Rasmussen, P.; Fredenslund, A. *Ind. Eng. Chem. Res.* **1987**, *26*, 159. (f) Hansen, H. K.; Rasmussen, P.; Schiller, M.; Gmehling, J. *Ind. Eng. Chem. Res.* **1991**, *30*, 2355. Recently, the use of UNIFAC for estimating thermodynamic activity coefficients has been called into question by van Tol et al. (van Tol, J. B. A.; Stevens, R. M. M.; Veldhuizen, W. J.; Jongejan, J. A.; Duine, J. A. *Biotechnol. Bioeng.* **1995**, *47*, 71) on the basis of a small number of experiments. However, references (a)–(f) above provide ample validation for UNIFAC for most solutes/solvents examined. Also, van Tol et al. used partition coefficients combined with solubility measurements (plagued with problems outlined in ref 3a) to determine the activity coefficients, while a rigorous treatment requires the use of vapor–liquid equilibrium measurements.<sup>33</sup>

<sup>†</sup> An NIH Biotechnology Predoctoral Trainee.

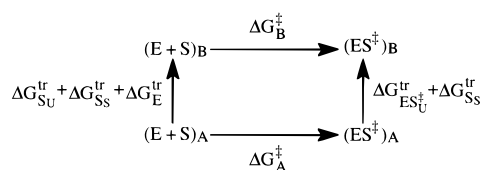
<sup>⊗</sup> Abstract published in *Advance ACS Abstracts*, March 15, 1996.

(1) (a) Klibanov, A. M. *Trends Biochem. Sci.* **1989**, *14*, 141. (b) Chen, C.-S.; Sih, C. J. *Angew. Chem., Int. Ed. Engl.* **1989**, *28*, 695. (c) Dordick, J. S. *Enzyme Microb. Technol.* **1989**, *11*, 194. (d) Klibanov, A. M. *Acc. Chem. Res.* **1990**, *23*, 114. (e) Gupta, M. N. *Eur. J. Biochem.* **1992**, *203*, 25. (f) Faber, K.; Riva, S. *Synthesis* **1992**, 895. (g) Halling, P. J. *Enzyme Microb. Technol.* **1994**, *16*, 178. (h) *Enzymatic Reactions in Organic Media*; Koskinen, A. M. P.; Klibanov, A. M., Eds.; Blackie: London, 1996.

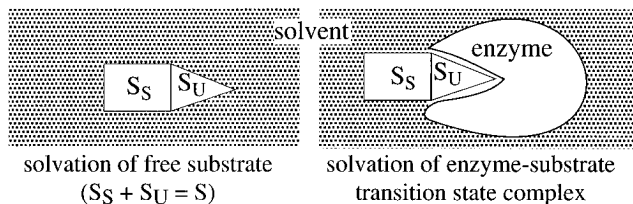
(2) For a review, see (a) Wescott, C. R.; Klibanov, A. M. *Biochim. Biophys. Acta* **1994**, *1206*, 1. (b) Carrea, G.; Ottolina, G.; Riva, S. *Trends Biotechnol.* **1995**, *13*, 63.

(3) (a) Wescott, C. R.; Klibanov, A. M. *J. Am. Chem. Soc.* **1993**, *115*, 1629. (b) Wescott, C. R.; Klibanov, A. M. *J. Am. Chem. Soc.* **1993**, *115*, 10362.

## Scheme 1



## Scheme 2



enzyme conformation and thus affect the selectivity of the reaction by altering enzyme-substrate interactions.<sup>5,6</sup> Alternatively, solvent molecules could bind within the enzyme active site and block the normal binding mode of the substrate.<sup>7,8</sup> While these two possible mechanisms do not necessarily influence selectivity, a third, driven by the energetics of substrate solvation, must do so, regardless of the presence of other mechanisms. Indeed, it has been demonstrated that the energetics of substrate solvation is the dominant means by which the solvent influences the substrate specificity of subtilisin.<sup>3</sup>

The contribution of solvation energies to the solvent dependence of enzyme kinetics is demonstrated by the thermodynamic cycle in Scheme 1. The lower horizontal arrow represents the enzyme (E) and the substrate (S) reacting in solvent A to form the transition state  $(ES^\ddagger)_A$  with an activation free energy of  $\Delta G^\ddagger_A$ . This transition state spontaneously decomposes to ultimately form the free enzyme and products. An alternative, hypothetical path exists leading to the same transition state complex. In this path, the substrate and enzyme are separately transferred from solvent A to solvent B, where they react to form the transition state  $(ES^\ddagger)_B$ , which is subsequently transferred back to solvent A. For thermodynamic purposes, the substrate can be represented as the sum of two portions (see Scheme 2), that which is solvated in the transition state ( $S_S$ ) and that which is enveloped by the enzyme and is thus

(5) This hypothesis has been proposed to explain solvent-induced variations in enzymatic selectivity: (a) Wu, H.-S.; Chu, F.-Y.; Wang, K.-T. *Bioorg. Med. Chem. Lett.* **1991**, *1*, 339. (b) Ueji, S.; Fujio, R.; Okubo, N.; Miyazawa, T.; Kurita, S.; Kitadani, M.; Muromatsu, A. *Biotechnol. Lett.* **1992**, *14*, 163.

(6) In the present work, this mechanism is selected against through the use of crystalline chymotrypsin and crystalline subtilisin, whose conformations have been shown to be essentially unaffected by replacement of water by organic solvents as the medium. For example, it is reasoned that if the structure of chymotrypsin is nearly the same in water and in hexane, then it also should be the same in other organic solvents because the differences in physicochemical properties among them are less than those between water and hexane.

(7) This hypothesis has been proposed to explain solvent-induced variations in enzymatic selectivity: (a) Nakamura, K.; Takebe, Y.; Kitayama, T.; Ohno, A. *Tetrahedron Lett.* **1991**, *32*, 4941. (b) Secundo, F.; Riva, S.; Carrea, G. *Tetrahedron: Asymmetry* **1992**, *3*, 267.

(8) It is likely<sup>16,29</sup> that the active center of the enzyme is occupied by at least a few solvent molecules when a substrate molecule is not bound. These bound solvent molecules would be in a dynamic equilibrium between the enzyme active center and the bulk solvent, described by a binding constant (effectively an inhibition constant,  $K_I$ ). If  $1/(K_I V_M) \ll [S]/K_M$  (where  $V_M$  is the molar volume of the solvent), the bound solvent would exert little effect on the catalytic properties of the system. If, on the other hand, the opposite were true, the solvent would act as an effective inhibitor. If such a tightly bound solvent molecule affects each substrate binding mode equally, it would not influence the prochiral selectivity. If the solvent molecule is tightly bound in such a way that it only hinders one substrate binding mode, the prochiral selectivity would be affected in a manner that could not be predicted by the treatment used in the present work.

unsolvated in the transition state ( $S_U$ ). Similarly, the transition state is regarded as the sum of  $S_S$  and another portion which includes only the enzyme and the unsolvated substrate moiety in the transition state ( $ES^\ddagger_U$ ).  $\Delta G^\ddagger_A$  can be expressed as the sum of the energetic terms of the alternative path<sup>9</sup>

$$\Delta G^\ddagger_A = \Delta G^\ddagger_B + \Delta G^\text{tr}_E + \Delta G^\text{tr}_{S_S} + \Delta G^\text{tr}_{S_U} - \Delta G^\text{tr}_{ES^\ddagger_U} - \Delta G^\text{tr}_{S_S} \quad (1)$$

where  $\Delta G^\ddagger_B$  is the activation free energy for the reaction in solvent B, and  $\Delta G^\text{tr}$  is the free energy of transfer of the moiety indicated in the subscript from solvent A to B. Assuming that the solvated surfaces of E and  $ES^\ddagger_U$  for low-molecular-weight substrates are identical,<sup>10</sup>  $\Delta G^\text{tr}_E = \Delta G^\text{tr}_{ES^\ddagger_U}$ . Consequently, eq 1 can be simplified to:

$$\Delta G^\ddagger_A = \Delta G^\ddagger_B + \Delta G^\text{tr}_{S_U} \quad (2)$$

$\Delta G^\ddagger$  is related to  $k_{\text{cat}}/K_M$  as:<sup>11</sup>

$$\Delta G^\ddagger = -RT \ln \left[ \left( \frac{k_{\text{cat}}}{K_M} \right) \left( \frac{h}{\kappa T} \right) \right] \quad (3)$$

where  $h$ ,  $\kappa$ ,  $R$ ,  $k_{\text{cat}}$ ,  $K_M$ , and  $T$  are the Planck, Boltzmann, gas, catalytic, and Michaelis constants and temperature, respectively.  $\Delta G^\text{tr}$  can be expressed in terms of thermodynamic activity coefficients as  $RT \ln (x_B \gamma_B / x_A \gamma_A)$ , where  $\gamma$  and  $x$  are the activity coefficient and mole fraction, respectively, of the solute in the indicated solvent.<sup>12</sup> If  $\gamma'$  is defined as the activity coefficient of the unsolvated substrate moiety ( $S_U$ ), then

$$\Delta G^\text{tr}_{S_U} = RT \ln (x_B \gamma'_B / x_A \gamma'_A) \quad (4)$$

Substituting eqs 3 and 4 into 2 yields

$$(k_{\text{cat}}/K_M)_A = (\gamma'_A/\gamma'_B)(x_A/x_B)(k_{\text{cat}}/K_M)_B \quad (5)$$

Note that for dilute solutions  $x_A/x_B$  depends only on the molar volume of the solvents when the transfer is done at constant molar concentration.<sup>13</sup> The mole fraction ratio is thus the same for any substrate and cancels out when eq 5 is expressed for two substrates, I and II, and solved for the logarithm of the selectivity in solvent A:

(9) Each of the steps of the cycle in Scheme 1 is reversible. Single, rather than double, arrows are used in the scheme solely to illustrate the directionality in the definition of the energetic terms.

(10) Because the conformations of serine proteases (and most other enzymes) are not altered in the transition state, the solvent-accessible surface of the enzyme is assumed to be the same in ground and transition states. The only difference between E and  $ES^\ddagger_U$  is, apart from displacement of the solvent from the active site, the addition of  $S_U$ , which is unsolvated and thus does not contribute to the solvated surface of the complex.

(11) Fersht, A. *Enzyme Structure and Mechanism*, 2nd ed.; Freeman: New York, 1985.

(12) The free energy of a solute dissolved in a solvent is described by  $G = G^\circ + RT \ln (x\gamma)$ . Because the solvent is the primary variable in our work, the standard state is chosen as the pure liquid solute. Thus  $G^\circ$  is independent of the solvent, and the free energy of transfer of the solute from solvent A to B ( $\Delta G^\text{tr}$ ) is  $RT \ln (x_B \gamma_B / x_A \gamma_A)$ .

(13) For dilute solutions,  $x_{\text{solute}} \approx n_{\text{solute}}/n_{\text{solvent}} = [\text{solute}]V_M$ , where  $V_M$  is the molar volume of the solvent which is reciprocal to its molar concentration. For solvents A and B,  $x_A/x_B = [\text{solute}]V_{MA}/[\text{solute}]V_{MB}$ . If the solute concentration is kept the same in both solvents (i.e., its transfer from A to B is done at constant molar concentration),  $[\text{solute}]$  cancels out, yielding  $x_A/x_B = V_{MA}/V_{MB}$ , i.e., the ratio of the molar volumes of the solvents.

$$\log \left[ \frac{(k_{\text{cat}}/K_M)_I}{(k_{\text{cat}}/K_M)_{II}} \right]_A = \log \left( \frac{\gamma'_I}{\gamma'_{II}} \right)_A + \log \left[ \frac{\gamma'_{II}(k_{\text{cat}}/K_M)_I}{\gamma'_I(k_{\text{cat}}/K_M)_{II}} \right]_B \quad (6)$$

Equation 6 expresses the general relationship between enzymatic selectivity and solvent-transition-state interactions. In the present work, we specifically test eq 6 with respect to prochiral selectivity. While the same substrate leads to both the *R* and *S* products, the reaction proceeds through conformationally distinct transition states for the production of each enantiomer (a *pro-R* transition state results in the *R* product, while a *pro-S* transition state forms the *S* product). Thus, differences between  $\gamma'_I$  and  $\gamma'_{II}$  arise from variations in transition state solvation, not from chemical differences between two substrates. The parameters for substrates I and II in eq 6 can therefore be replaced with those for the *pro-R* and *pro-S* reaction pathways to describe prochiral selectivity. Also, if B is fixed as a reference solvent,<sup>14</sup> the final term in eq 6 will be a constant. Consequently, one arrives at the following equation describing the solvent dependence of the prochiral selectivity of an enzyme in terms of a  $\gamma'$  ratio:

$$\log \left[ \frac{(k_{\text{cat}}/K_M)_{\text{pro-S}}}{(k_{\text{cat}}/K_M)_{\text{pro-R}}} \right] = \log \left( \frac{\gamma'_{\text{pro-S}}}{\gamma'_{\text{pro-R}}} \right) + \text{const} \quad (7)$$

Unlike the situation for substrate specificity, where solvent-dependent variation in the activity coefficient ratio for the two substrates is primarily driven by chemical differences between the substrates,  $\gamma'$  for prochiral selectivity differs for each reaction pathway only due to differences in transition state solvation. In this work, we calculate  $\gamma'$  for both the *pro-R* and *pro-S* transition states using a three-step procedure. First, the desolvated portion of the substrate in the transition state is determined using molecular modeling based on the crystal structure of the enzyme. Second, this desolvated moiety is expressed in terms of distinct chemical groups to yield a model compound which approximates the portion of the substrate removed from the solvent in the transition state. Finally, the thermodynamic activity coefficient of this model compound is calculated using UNIFAC and then equated to  $\gamma'$ . According to eq 7, knowing only  $\gamma'_{\text{pro-R}}$  and  $\gamma'_{\text{pro-S}}$  for a series of solvents, it should be possible to predict the solvent dependence of prochiral selectivity.

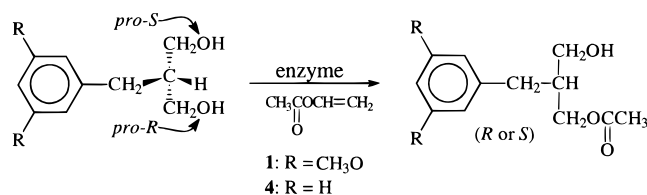
## Results and Discussion

As an initial test of the ability of the model described above to predict the solvent dependence of prochiral selectivity, we have examined the enzymatic acylation of a designed prochiral diol, 2-(3,5-dimethoxybenzyl)-1,3-propanediol (**1**), by vinyl acetate to produce the chiral monoester 3-hydroxy-2-(3,5-dimethoxybenzyl)propyl acetate (**2**) (see Scheme 3). The crystalline chymotrypsin, lightly cross-linked with glutaraldehyde, has been selected as a catalyst.<sup>15</sup> Its structure in an organic

(14) This reference solvent should not be confused with the standard state of the thermodynamic activity coefficients.<sup>12</sup> It should further be noted that consideration of this constant in terms of a "corrected substrate specificity" will depend on the standard state chosen for the activity coefficients (Janssen, A. E. M.; Halling, P. J. *J. Am. Chem. Soc.* **1994**, *116*, 9827).

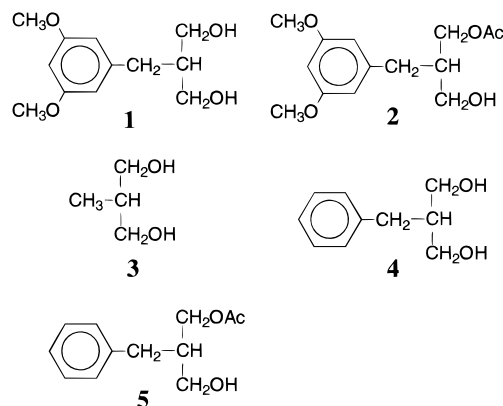
(15) Such cross-linked enzyme crystals (CLECs) have been employed as robust catalysts in synthetic transformations. (a) St. Clair, N. L.; Navia, M. A. *J. Am. Chem. Soc.* **1992**, *114*, 7314. (b) Sobolov, S. B.; Bartoszko-Malik, A.; Oeschger, T. R.; Montalbano, M. M. *Tetrahedron Lett.* **1994**, *35*, 7751. (c) Persichetti, R. A.; St. Clair, N. L.; Griffith, J. P.; Navia, M. A.; Margolin, A. L. *J. Am. Chem. Soc.* **1995**, *117*, 2732. (d) Lalonde, J. J.; Govardhan, C.; Khalaf, N.; Martinez, A. G.; Visuri, K.; Margolin, A. L. *J. Am. Chem. Soc.* **1995**, *117*, 6845.

## Scheme 3



solvent has been recently determined by X-ray crystallography<sup>16</sup> and found to be nearly the same as in water, thus enabling structure-based computer modeling.

The first step in the implementation of the model is the determination of the desolvated portions of **1** in the transition state for each product enantiomer. Since serine proteases suspended in organic solvents act via the ping-pong bi-bi mechanism,<sup>17</sup> at least two transition states are involved in each enzyme turnover, one for the acylation by vinyl acetate to form the acyl-enzyme intermediate and the other for the regeneration of the free enzyme by the nucleophile. The latter transition state leads to the production of the chiral product. Thus, two structural models were constructed (see Methods for details) of the transition state for the reaction of **1** with the acyl-

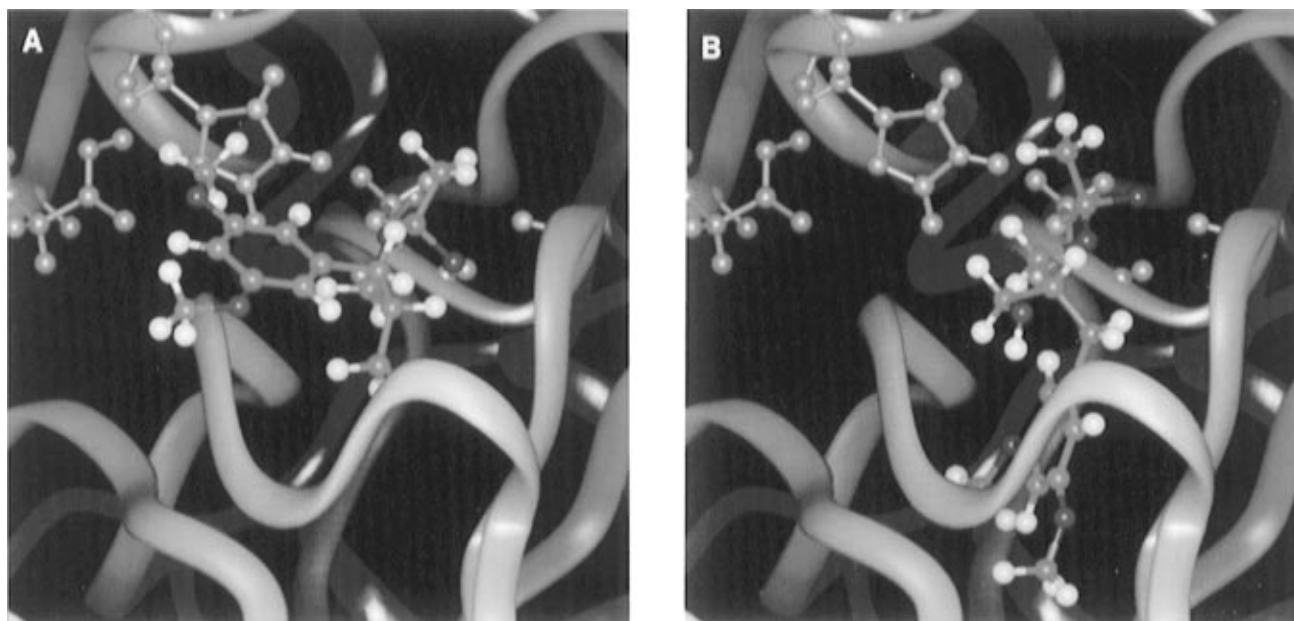


chymotrypsin. In the first (Figure 1A), **1** binds to the enzyme in such a way that the *S* product is formed. Due to steric constraints, the dimethoxyphenyl group in the *pro-S* transition state is unable to enter the S1 binding site of chymotrypsin and hence must extend away from the enzyme into the solvent. The second model (Figure 1B) depicts the transition state which leads to the *R* product. Unlike in the *pro-S* situation, the *pro-R* transition state adopts a conformation in which the dimethoxyphenyl group is buried in the S1 binding site of the enzyme. The determination of the solvent-accessible surface areas of **1** in each of the transition state models (Figure 2) reveals that the entire substrate is unsolvated in the *pro-R* transition state, while it is largely solvated in the *pro-S* transition state.

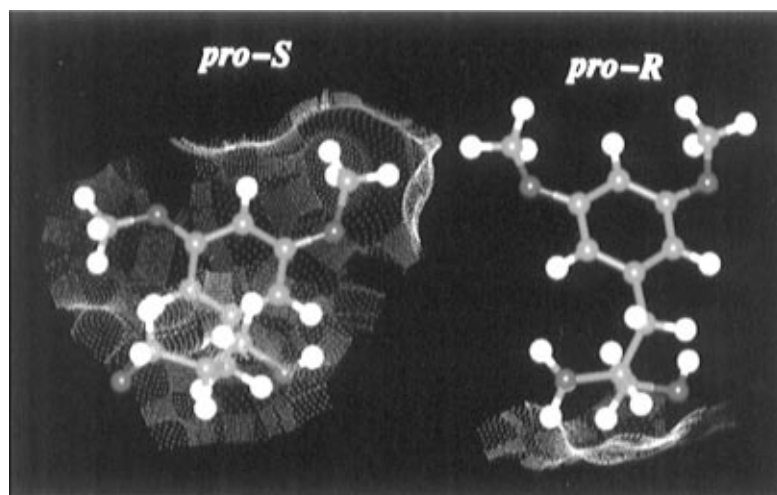
With the extent of solvation of the substrates elucidated, it is possible to proceed with the second phase of our strategy: the selection of model compounds which mimic the unsolvated portion of **1** in the *pro-S* and *pro-R* transition states. Table 1 quantitatively examines the extent to which each chemical group of the substrate is desolvated. The 2-methyl-1,3-propanediol (**3**) moiety is almost equally desolvated in the *pro-S* and *pro-R* transition states and consequently should be a constituent of both model compounds. The methoxy and phenyl groups are completely desolvated only in the *pro-R* transition state and are

(16) (a) Yennawar, N. H.; Yennawar, H. P.; Farber, G. K. *Biochemistry* **1994**, *33*, 7326. (b) Yennawar, H. P.; Yennawar, N. H.; Farber, G. K. *J. Am. Chem. Soc.* **1995**, *117*, 577.

(17) (a) Zaks, A.; Klivanov, A. M. *J. Biol. Chem.* **1988**, *263*, 3194. (b) Chatterjee, S.; Russell, A. J. *Enzyme Microb. Technol.* **1993**, *15*, 1022.



**Figure 1.** Conformation of substrate **1** in the *pro-S* (A) and *pro-R* (B) transition states with  $\gamma$ -chymotrypsin. See the Methods section for details on the construction of the molecular models.



**Figure 2.** Solvent-accessible surface areas of substrate **1** in the *pro-S* (left) and *pro-R* (right) transition states with  $\gamma$ -chymotrypsin. The solvent-accessible surfaces (represented by dot surfaces) were calculated using the Connolly method (see the Methods section for details).

**Table 1.** Extent of Desolvation of **1** in the Transition State with  $\gamma$ -Chymotrypsin

compound	group	percentage of desolvation <sup>a</sup>	
		<i>pro-S</i> (%)	<i>pro-R</i> (%)
	methoxy 1	35	100
	methoxy 2	49	100
	phenyl	14	100
	2-methyl-1,3-propanediol	78	71

<sup>a</sup> The desolvated surface area of each group of the substrate in the enzyme-bound transition state was determined using molecular models (see Figures 1 and 2) and expressed as a percentage of the solvent-accessible surface area of the group in the unbound substrate.

therefore only included in the *pro-R* model compound; their partial desolvation in the *pro-S* transition state is disregarded (see below). Thus analysis of the data in Table 1 leads to the choice of **3** and **1** as the *pro-S* and *pro-R* model compounds, respectively.

The final phase in the prediction of the solvent dependence of enzymatic prochiral selectivity is the employment of the model compounds to calculate  $\gamma'_{pro-R}$  and  $\gamma'_{pro-S}$  for use in eq

7. Because  $\gamma'$  is defined as the activity coefficient of the unsolvated substrate moiety in the transition state, it is approximated by the activity coefficient of the appropriate model compound. Consequently, UNIFAC-calculated activity coefficients of **3** and **1** are reported in Table 2 as  $\gamma'_{pro-S}$  and  $\gamma'_{pro-R}$ , respectively.

A profound solvent effect on the activity coefficient ratio is evident in Table 2:  $\gamma'_{pro-S}/\gamma'_{pro-R}$  varies over an order of magnitude when the solvent is changed from diisopropyl ether to acetonitrile. Equation 7 predicts that this variation in the activity coefficient ratio will be reflected in the same change in prochiral selectivity. As a test of this prediction, the prochiral selectivity of cross-linked crystalline  $\gamma$ -chymotrypsin toward **1** was measured experimentally in a variety of solvents (Table 3, first data column). In such solvents as diisopropyl ether and cyclohexane, the enzyme strongly favors production of the *R*-product. As predicted by eq 7, switching to solvents with higher  $\gamma'_{pro-S}/\gamma'_{pro-R}$  ratios, such as dioxane or tetrahydrofuran, brings about a concomitant change in prochiral selectivity, eventually resulting in the preferential formation of the *S*-product, e.g., in acetonitrile and methyl acetate (i.e., a solvent-

**Table 2.** Thermodynamic Activity Coefficients of Compounds Modeling the Desolvated Fractions of Substrates **1** and **4** in the Transition States with  $\gamma$ -Chymotrypsin Calculated Using UNIFAC

solvent	substrate <b>1</b>			substrate <b>4</b>		
	$\gamma'_{pro-S}$	$\gamma'_{pro-R}$	$\frac{\gamma'_{pro-S}}{\gamma'_{pro-R}}$	$\gamma'_{pro-S}$	$\gamma'_{pro-R}$	$\frac{\gamma'_{pro-S}}{\gamma'_{pro-R}}$
diisopropyl ether	73.5	99.6	0.738	73.8	36.2	2.04
dibutyl ether	51.4	69.4	0.741			
cyclohexane	480	266	1.80			
dioxane	5.09	2.37	2.15	5.09	3.36	1.51
<i>tert</i> -butyl acetate	13.5	4.59	2.94	13.5	5.16	2.62
tetrahydrofuran	8.13	2.65	3.07			
<i>p</i> -xylene	70.7	16.1	4.39			
toluene	66.9	11.9	5.62			
methyl acetate	7.25	0.986	7.35	7.25	2.14	3.39
propionitrile	3.10	0.401	7.73			
benzene	67.6	7.28	9.29	67.6	14.6	4.63
acetonitrile	2.71	0.278	9.75			

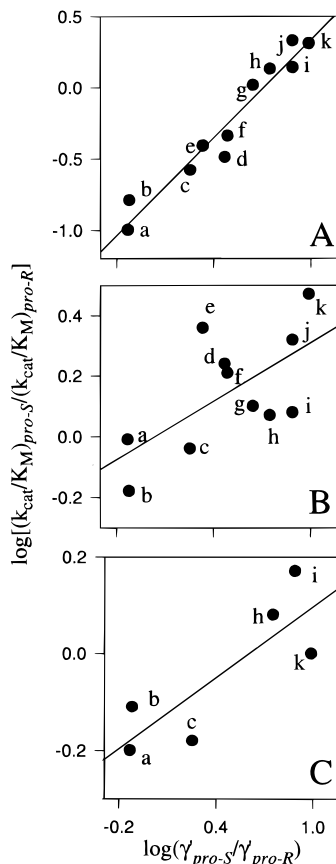
induced inversion of chymotrypsin's prochiral selectivity is observed). Indeed, the prochiral selectivity as a function of the solvent varies over the order of magnitude range predicted by eq 7.

Equation 7 further predicts that a double-logarithmic plot of  $(k_{cat}/K_M)_{pro-S}/(k_{cat}/K_M)_{pro-R}$  vs  $\gamma'_{pro-S}/\gamma'_{pro-R}$  will produce a linear correlation. Such a plot, presented in Figure 3A, does indeed yield a reasonable linear dependence ( $R^2 = 0.94$ ). The slope of the regression line of Figure 3A is 1.1, close to the predicted value of 1.0.

Several sources of error are inherent to this treatment and may contribute to the 10% deviation of the slope from unity. First, the entire dimethoxyphenyl group of **1** in the *pro-S* transition state is treated as solvated. In reality, a fraction of its surface is partially desolvated (see Table 1), making it impossible to exactly represent the desolvated substrate moiety as a complete molecule. Second, there are unavoidable errors associated with UNIFAC, which produces a mere statistical estimate of the activity coefficient.<sup>4</sup>

Because most enzymes used in organic solvents to date have been prepared in the form of lyophilized powders,<sup>1</sup> it is important to ascertain whether the thermodynamic analytical methodology developed herein is applicable to such and other enzyme preparations besides cross-linked crystals. The determination of  $\gamma'$  is contingent on the ability to predict the specific interactions between the substrate and the enzyme. Hence, even a minor conformational perturbation of the enzyme active site could change the solvation of the transition state and thus invalidate the calculation of  $\gamma'$ . Recent hydrogen isotope exchange/NMR<sup>18</sup> and FTIR<sup>19</sup> experiments indicate that reversible conformational changes occur in proteins upon dehydration. Thus, we reasoned that lyophilized or precipitated (from water with cold acetone) chymotrypsin may not adhere to the solvent dependence of prochiral selectivity predicted on the basis of the enzyme crystal structure.

When the prochiral selectivity of lyophilized or acetone-precipitated chymotrypsin is measured for substrate **1** in several solvents (Table 3), the selectivity varies less than 3-fold, while the solvent effect predicted by eq 7 assuming an intact enzyme structure spans a factor of 13. Moreover, double-logarithmic



**Figure 3.** Dependence of the prochiral selectivity of different preparations of  $\gamma$ -chymotrypsin toward substrate **1** in various organic solvents on the ratio of the activity coefficients of the desolvated portions of the substrate in the *pro-S* and *pro-R* transition states (see eq 7). A—cross-linked enzyme crystals; B—lyophilized enzyme powder; C—acetone-precipitated enzyme powder. Solvents: a—diisopropyl ether; b—dibutyl ether; c—cyclohexane; d—*tert*-butyl acetate; e—dioxane; f—tetrahydrofuran; g—*p*-xylene; h—toluene; i—propionitrile; j—methyl acetate; k—acetonitrile. The straight lines are drawn using linear regression; see text for discussion. For experimental details, see the Methods section.

plots of the prochiral selectivity of lyophilized (Figure 3B) and acetone-precipitated (Figure 3C) chymotrypsin vs  $\gamma'_{pro-S}/\gamma'_{pro-R}$  do not yield the correlations obtained for structurally intact chymotrypsin: while theory predicts that the slopes of the aforementioned plots should be unity, those determined by linear regression are 0.32 and 0.24, respectively. Additionally, only a poor correlation between the prochiral selectivity of these enzyme powders and  $\gamma'_{pro-S}/\gamma'_{pro-R}$  is reflected by  $R^2$  values of 0.42 for lyophilized and 0.68 for acetone-precipitated chymotrypsin (as compared to 0.94 for the cross-linked crystals). These findings underscore the importance of dealing with structurally defined enzyme catalysts if a quantitative rationale involving stereoselectivity is sought after.

While our model requires the enzyme to be in a native (or at least a known) conformation in organic solvents, it places no further constraints on the catalyst. Therefore, the proposed analysis should be applicable to enzymes other than chymotrypsin. To test this prediction, the acylation of **1** by vinyl acetate, catalyzed by cross-linked crystalline subtilisin Carlsberg<sup>20</sup> (henceforth referred to as subtilisin), has been investigated using the same methodology as with the chymotrypsin-catalyzed reaction.

(20) Whose X-ray crystal structure in an anhydrous solvent was found to be virtually indistinguishable from that in water.

(18) (a) Desai, U. R.; Osterhout, J. J.; Klibanov, A. M. *J. Am. Chem. Soc.* **1994**, *116*, 9420. (b) Desai, U. R.; Klibanov, A. M. *J. Am. Chem. Soc.* **1995**, *117*, 3940.

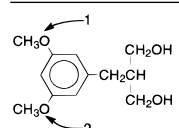
(19) (a) Prestrelski, S. J.; Tedeschi, N.; Arakawa, T.; Carpenter, J. F. *Biophys. J.* **1993**, *65*, 661. (b) Prestrelski, S. J.; Tedeschi, N.; Arakawa, T.; Carpenter, J. F. *Arch. Biochem. Biophys.* **1993**, *303*, 465. (c) Costantino, H. R.; Griebenow, K.; Mishra, P.; Langer, R.; Klibanov, A. M. *Biochim. Biophys. Acta* **1995**, *1253*, 69. (d) Griebenow, K.; Klibanov, A. M. *Proc. Natl. Acad. Sci. U.S.A.* **1995**, *92*, 10969.

**Table 3.** Solvent Dependence of the Prochiral Selectivity of  $\gamma$ -Chymotrypsin, Prepared by Different Methods, toward Substrate **1**

solvent	$(k_{\text{cat}}/K_M)_{\text{pro-S}}/(k_{\text{cat}}/K_M)_{\text{pro-R}}^a$		
	cross-linked crystalline <sup>b</sup>	lyophilized <sup>c</sup>	acetone-precipitated <sup>d</sup>
diisopropyl ether	0.10	0.98	0.63
dibutyl ether	0.17	0.67	0.77
cyclohexane	0.28	0.91	0.67
<i>tert</i> -butyl acetate	0.32	1.7	
dioxane	0.38	2.3	
tetrahydrofuran	0.45	1.6	
<i>p</i> -xylene	1.0	1.3	
toluene	1.3	1.2	1.2
propionitrile	1.4	2.1	
acetonitrile	1.9	2.9	1.0
methyl acetate	2.2	1.2	1.5

<sup>a</sup>  $(k_{\text{cat}}/K_M)_{\text{pro-S}}/(k_{\text{cat}}/K_M)_{\text{pro-R}}$  values were determined from the ratio of initial velocities for the production of each enantiomer as described in the Methods section. <sup>b</sup> The initial velocities,  $v_S$  and  $v_R$ , respectively, observed for 5 mg/mL crystalline  $\gamma$ -chymotrypsin were (in  $\mu\text{M}/\text{h}$ ) as follows: diisopropyl ether—0.52 and 5.2; dibutyl ether—0.34 and 2.0; cyclohexane—0.74 and 2.6; *tert*-butyl acetate—0.068 and 0.21; dioxane—0.63 and 1.6; tetrahydrofuran—1.4 and 3.1; *p*-xylene—5.6 and 5.6; toluene—31 and 24; propionitrile—4.7 and 3.3; acetonitrile—2.5 and 1.3; methyl acetate—1.8 and 0.82. <sup>c</sup> The initial velocities,  $v_S$  and  $v_R$ , respectively, observed for 15 mg/mL lyophilized  $\gamma$ -chymotrypsin were (in  $\mu\text{M}/\text{h}$ ) as follows: diisopropyl ether—44 and 45; dibutyl ether—15 and 23; cyclohexane—4.7 and 5.4; *tert*-butyl acetate—14 and 24; dioxane—3.8 and 8.6; tetrahydrofuran—11 and 18; *p*-xylene—120 and 96; toluene—53 and 45; propionitrile—6.7 and 3.2; acetonitrile—2.4 and 8.1; methyl acetate—11 and 9.2. <sup>d</sup> The initial velocities,  $v_S$  and  $v_R$ , respectively, observed for 15 mg/mL acetone-precipitated  $\gamma$ -chymotrypsin were (in  $\mu\text{M}/\text{h}$ ) as follows: diisopropyl ether—32 and 53; dibutyl ether—19 and 24; cyclohexane—7.9 and 12; toluene—73 and 61; acetonitrile—28 and 27; methyl acetate—18 and 12.

**Table 4.** Extent of Desolvation of **1** in the Transition State with Subtilisin Carlsberg

compound	group	percentage of desolvation <sup>a</sup>	
		<i>pro-S</i> (%)	<i>pro-R</i> (%)
	methoxy 1	9	14
	methoxy 2	54	60
	phenyl	55	52
	2-methyl-1,3-propanediol	56	62

<sup>a</sup> The desolvated surface area of each group of the substrate in the enzyme-bound transition state was determined using molecular models (similar to those in Figures 1 and 2) and expressed as a percentage of the solvent-accessible surface area of the group in the unbound substrate.

Because the S1 binding site of subtilisin is much more shallow than chymotrypsin's, no portion of substrate **1** in the either *pro-S* or *pro-R* transition state is fully shielded from the solvent. Furthermore, calculation of the unsolvated surfaces of the groups which make up the substrate reveals that there is little difference in the extent of desolvation of the two transition states (Table 4). Due to the analogous solvation patterns, the model compounds for the calculation of  $\gamma'_{\text{pro-R}}$  and  $\gamma'_{\text{pro-S}}$ , whatever they may be, should be identical, resulting in the prediction by eq 7 that the prochiral selectivity of subtilisin in the acetylation of **1** will be independent of the solvent. In support of this prediction, when  $(k_{\text{cat}}/K_M)_{\text{pro-S}}/(k_{\text{cat}}/K_M)_{\text{pro-R}}$  for subtilisin is measured in a range of solvents which span a 19-fold change in the prochiral selectivity of chymotrypsin for the same substrate, less than a 2-fold variation is observed (Table 5).

In addition to being applicable to various enzymes, our treatment contains no restraints on the substrate. To verify its applicability to other substrates, we have examined the acylation of 2-benzyl-1,3-propanediol (**4**) (i.e., **1** devoid of both methoxy groups) by vinyl acetate catalyzed by cross-linked crystalline

**Table 5.** Solvent Dependence of Prochiral Selectivity of Cross-Linked Crystalline Subtilisin Carlsberg for Substrate **1**

solvent	$(k_{\text{cat}}/K_M)_{\text{pro-S}}^a$
	$(k_{\text{cat}}/K_M)_{\text{pro-R}}$
diisopropyl ether	1.2
cyclohexane	1.2
<i>tert</i> -butyl acetate	1.6
acetonitrile	2.2
toluene	2.3

<sup>a</sup>  $(k_{\text{cat}}/K_M)_{\text{pro-S}}/(k_{\text{cat}}/K_M)_{\text{pro-R}}$  values were determined from the ratio of initial velocities for the production of each enantiomer as described in the Methods section. The initial velocities,  $v_S$  and  $v_R$ , respectively, observed for 2.5 mg/mL crystalline subtilisin Carlsberg were (in  $\mu\text{M}/\text{h}$ ) as follows: diisopropyl ether—28 and 23; cyclohexane—20 and 17; *tert*-butyl acetate—4.3 and 2.7; acetonitrile—3.1 and 1.4; toluene—18 and 7.8.

**Table 6.** Solvent Dependence of Prochiral Selectivity of Cross-Linked Crystalline  $\gamma$ -Chymotrypsin for Substrate **4**

solvent	$(k_{\text{cat}}/K_M)_{\text{pro-S}}^a$
	$(k_{\text{cat}}/K_M)_{\text{pro-R}}$
dioxane	0.58
diisopropyl ether	0.64
<i>tert</i> -butyl acetate	0.88
methyl acetate	1.2
benzene	1.3

<sup>a</sup>  $(k_{\text{cat}}/K_M)_{\text{pro-S}}/(k_{\text{cat}}/K_M)_{\text{pro-R}}$  values were determined from the ratio of initial velocities for the production of each enantiomer as described in the Methods section. The initial velocities,  $v_S$  and  $v_R$ , respectively, observed for 5 mg/mL crystalline  $\gamma$ -chymotrypsin were (in  $\mu\text{M}/\text{h}$ ) as follows: dioxane—1.0 and 1.8; diisopropyl ether—2.3 and 3.6; *tert*-butyl acetate—3.0 and 3.4; methyl acetate—11 and 9.5; benzene—16 and 12.

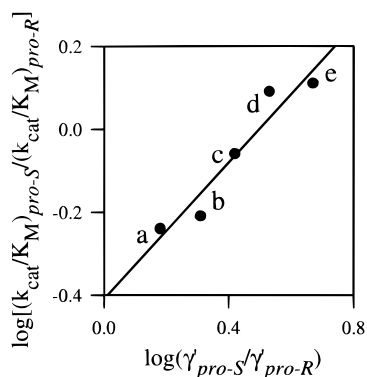
chymotrypsin. Modeling of the *pro-R* and *pro-S* transition states and subsequent calculation of the unsolvated areas of the substrates reveals that, as was the case with **1**, while the 2-methyl-1,3-propanediol fragment is unsolvated in both transition states, the phenyl group is only protected from the solvent in the *pro-R*. Consequently, the model compounds representing the unsolvated portions of **4** in the *pro-S* and *pro-R* transition states are **3** and **4**, respectively.

Table 2 (last column) lists the  $\gamma'$  ratios for substrate **4** in some of the same solvents used for **1**. Using these  $\gamma'$  ratios, our model predicts that **4** will behave differently than substrate **1** in two respects. First, the  $\gamma'$  ratio varies only 3-fold for **4**, whereas it spans a 12-fold range for **1**. Thus the prochiral selectivity for substrate **4** should be similarly less solvent-sensitive. Second, the  $\gamma'$  ratio for **4** is lower in dioxane than in diisopropyl ether, while the opposite is true for **1**. An analogous inversion in the ranking of these solvents with respect to prochiral selectivity is predicted by eq 7 upon switching from substrate **1** to **4**.

The measured prochiral selectivities of chymotrypsin in the acetylation of **4** are presented in Table 6. As predicted, the solvent effect on selectivity is much smaller for substrate **4** than for **1**, spanning less than 3 fold vs 22 fold. Also, the predicted inversion of the ranking of dioxane and diisopropyl ether with respect to prochiral selectivity is indeed observed. In further corroboration of eq 7, the double-logarithmic plot of prochiral selectivity vs  $\gamma'$  ratio (Figure 4) is linear (with a satisfactory correlation coefficient of 0.92) and has a slope of 0.82.

## Concluding Remarks

When crystalline enzymes are used as asymmetric catalysts in anhydrous organic solvents, the solvent dependence of enzymatic prochiral selectivity can be attributed primarily to changes in the relative solvation energies for the *pro-R* and *pro-S*



**Figure 4.** Dependence of the prochiral selectivity of  $\gamma$ -chymotrypsin toward substrate **4** in various organic solvents on the ratio of the activity coefficients of the desolvated portions of the substrate in the *pro-S* and *pro-R* transition states. Solvents: a—dioxane; b—diisopropyl ether; c—*tert*-butyl acetate; d—methyl acetate; e—benzene. See the legend to Figure 3 for other details.

binding modes of the substrate in the transition state. This work presents a quantitative model which satisfactorily predicts the solvent effect on prochiral selectivity solely on the basis of these solvation energies. Thus other factors not considered by the model, e.g., the effect of the solvent on the enzyme or displacement of bound solvent molecules from the active site by the substrate, are deemed relatively unimportant. While this model performs reasonably well with crystalline enzymes, its implementation with respect to amorphous (lyophilized or acetone-precipitated) enzyme powders is not possible due to the ill-defined, reversibly denatured structure of proteins in such dehydrated states. Because no specific assumptions are made regarding the enzyme or the substrate in the derivation of the equation, the model should be generally applicable. Moreover, our treatment should be similarly applicable to other types of enzymatic selectivity, such as enantioselectivity, chemoselectivity, regioselectivity, and substrate selectivity.

## Materials and Methods

**Enzymes.** *Rhizomucor miehei* lipase (EC 3.1.1.3) was obtained from Fluka. Subtilisin Carlsberg (alkaline protease from *Bacillus licheniformis*, EC 3.4.21.14) and bovine pancreatic  $\gamma$ - and  $\alpha$ -chymotrypsins (EC 3.4.21.1) were purchased from Sigma Chemical Co. Note that the  $\alpha$ - and  $\gamma$ -forms of chymotrypsin are covalently identical and interconvertible in a pH-dependent manner, differing primarily in their crystallization properties.<sup>21</sup>  $\gamma$ -Chymotrypsin crystals were created from the  $\alpha$ - form of the enzyme, following the method of Stoddard et al. (see the enzyme crystallization section for further details).

**Chemicals and Solvents.** All chemicals and solvents were purchased from Aldrich Chemical Co. The organic solvents were of the highest purity available from that vendor (analytical grade or better) and were dried prior to use to a water content below 0.01% (as determined by the Karl Fischer titration<sup>22</sup>) by shaking with Linde's 3-Å molecular sieves.

**Dimethyl 2-(3,5-dimethoxybenzyl)malonate** was prepared by dissolving 0.12 g (5 mmol) of metallic sodium in 20 mL of dry methanol, followed by addition of 0.6 g (5 mmol) of dimethyl malonate. The solution was cooled to 0 °C, and 0.76 g (4 mmol) of 3,5-dimethoxybenzyl chloride in 20 mL of dry THF was added dropwise over 30 min. The mixture was refluxed overnight under argon and cooled to room temperature, and then 20 mL of cold water was added to quench the reaction. The product was extracted with ethyl acetate and purified by distillation. The yield of the product was 1.0 g (70% of theor.). <sup>1</sup>H

NMR (CDCl<sub>3</sub>)  $\delta$  6.2–6.3 (3 H, m),  $\delta$  3.71 (6 H, s),  $\delta$  3.67 (6 H, s),  $\delta$  3.64 (1 H, t,  $J = 7.4$  Hz),  $\delta$  3.11 (2 H, d,  $J = 7.4$  Hz).

**1** was prepared by dissolving 0.56 g (2 mmol) of dimethyl 2-(3,5-dimethoxybenzyl)malonate in 10 mL of dry THF, followed by a dropwise addition to 8 mL of a 1.0 M LiAlH<sub>4</sub> solution in ether at 0 °C. The mixture was stirred under argon overnight at room temperature, and then 0.5 mL of water was added to quench the reaction. The product was extracted using ether and, after rotary evaporation, purified by flash column chromatography (3:1 (v/v) ethyl acetate:hexane). The yield of **1** was 0.27 g (60% of theor.). <sup>1</sup>H NMR (CDCl<sub>3</sub>)  $\delta$  6.2–6.4 (3 H, m),  $\delta$  3.76 (2 H, dd,  $J = 10.2, 3.9$  Hz),  $\delta$  3.73 (6 H, s),  $\delta$  3.65 (2 H, dd,  $J = 10.2, 6.9$  Hz),  $\delta$  2.55 (2 H, d,  $J = 7.6$  Hz),  $\delta$  2.06 (1 H, m),  $\delta$  2.03 (2 H, s). Anal. Calcd for C<sub>12</sub>H<sub>18</sub>O<sub>4</sub>: C, 63.70; H, 8.02; O, 28.28. Found: C, 62.60; H, 8.10; O, 29.63.

**2** (a racemate used to calibrate the HPLC instrument) was prepared by dissolving 0.11 g (0.5 mmol) of **1** in 10 mL of dry ether, followed by addition of 0.05 mL of triethylamine at 0 °C and 0.05 mL of acetyl chloride dissolved in 5 mL of dry ether. The reaction was monitored by TLC and quenched by addition of 1 mL of water. The product was extracted with ethyl acetate and, after rotary evaporation, purified by flash column chromatography (1:1 (v/v) ethyl acetate:hexane). The yield was 54 mg (40% of theor.). <sup>1</sup>H NMR (CDCl<sub>3</sub>)  $\delta$  6.2–6.3 (3 H, m),  $\delta$  4.15 (1 H, dd,  $J = 12, 4.2$  Hz),  $\delta$  4.04 (1 H, dd,  $J = 12, 6.4$  Hz),  $\delta$  3.73 (6 H, s),  $\delta$  3.56 (1 H, dd,  $J = 12, 4.2$  Hz),  $\delta$  3.47 (1 H, dd,  $J = 12, 6.6$  Hz),  $\delta$  2.57 (1 H, dd,  $J = 13, 7.3$  Hz),  $\delta$  2.52 (1 H, dd,  $J = 13, 6.6$  Hz),  $\delta$  2.1 (1 H, m),  $\delta$  2.05 (3 H, s),  $\delta$  1.66 (1 H, m). Anal. Calcd for C<sub>14</sub>H<sub>20</sub>O<sub>5</sub>: C, 62.67; H, 7.51; O, 29.82. Found: C, 62.03; H, 7.62; O, 29.92.

**4** was prepared by dissolving 0.50 g (2 mmol) of diethyl benzylmalonate in 10 mL of dry THF, and adding this solution dropwise to 8 mL of 1.0 M LiAlH<sub>4</sub> dissolved in ether at 0 °C. The mixture was stirred under argon overnight at room temperature, and then 0.5 mL of water was added to quench the reaction. The product was extracted using ether and, after rotary evaporation, purified by flash column chromatography (3:1 (v/v) ethyl acetate:hexane). The yield was 0.20 g (60% of theor.). <sup>1</sup>H NMR (CDCl<sub>3</sub>)  $\delta$  7.1–7.4 (5 H, m),  $\delta$  3.85 (2 H, dd,  $J = 11, 3.6$  Hz),  $\delta$  3.72 (2 H, dd,  $J = 11, 6.7$  Hz),  $\delta$  2.67 (2 H, d,  $J = 7.6$  Hz),  $\delta$  2.1 (1 H, m),  $\delta$  2.08 (2 H, s). Anal. Calcd for C<sub>10</sub>H<sub>14</sub>O<sub>2</sub>: C, 72.26; H, 8.49; O, 19.25. Found: C, 72.50; H, 8.85; O, 17.77.

**5** (a racemate used to calibrate the HPLC instrument) was prepared by dissolving 83 mg (0.5 mmol) of **4** in 10 mL of dry ether, followed by addition of 0.05 mL of triethylamine at 0 °C and 0.05 mL of acetyl chloride dissolved in 5 mL of dry ether. The reaction was followed by TLC and quenched after 20 min with 1 mL of water. The product was extracted with ethyl acetate and purified by flash column chromatography (1:1 (v/v) ethyl acetate:hexane). The yield of **5** was 45 mg (43% of theor.). <sup>1</sup>H NMR (CDCl<sub>3</sub>)  $\delta$  7.35 (5 H, m),  $\delta$  4.28 (1 H, dd,  $J = 4.6, 11.3$  Hz),  $\delta$  4.18 (1 H, dd,  $J = 6.4, 11.2$  Hz),  $\delta$  3.70 (1 H, dd,  $J = 4.6, 11.2$  Hz),  $\delta$  3.60 Hz (1 H, dd,  $J = 6.2, 11.3$  Hz),  $\delta$  2.77 (1H, dd,  $J = 7.8, 11.5$  Hz),  $\delta$  2.73 (1 H, dd,  $J = 7.8, 11.5$  Hz),  $\delta$  2.2 (1 H, m),  $\delta$  2.18 (3H, s),  $\delta$  1.94 (1H, m). Anal. Calcd for C<sub>12</sub>H<sub>16</sub>O<sub>3</sub>: C, 69.21; H, 7.74; O, 23.05. Found: C, 68.21; H, 7.88; O, 24.01.

**Enzymatically Prepared (R)-2.** In a 20-mL screw-cap scintillation vial, 45 mg (0.2 mmol) of **1** was dissolved in 10 mL of dried diisopropyl ether. Then, 20 mg of cross-linked  $\gamma$ -chymotrypsin crystals (see below) was added to the vial, followed by addition of 20  $\mu$ L of deionized water and 0.4 mL (10 mmol) of vinyl acetate. The vial was shaken at 45 °C and 300 rpm for 12 h. The mixture was filtered, and the crystalline enzyme was washed 3 times with 10 mL of diisopropyl ether. The washings were combined with the filtrate, and the subsequent workup of the mixture was the same as described above for the chemically synthesized racemic **2**. The yield of **2** ( $[\alpha]_D^{28} +20.3^\circ$  (CHCl<sub>3</sub>), 82% ee by chiral HPLC) was 36 mg (71% of theor.); determination of its absolute configuration is described below. <sup>1</sup>H NMR (CDCl<sub>3</sub>)  $\delta$  6.2–6.3 (3 H, m),  $\delta$  4.15 (1 H, dd,  $J = 12, 4.2$  Hz),  $\delta$  4.04 (1 H, dd,  $J = 12, 6.4$  Hz),  $\delta$  3.73 (6 H, s),  $\delta$  3.56 (1 H, dd,  $J = 12, 4.2$  Hz),  $\delta$  3.47 (1 H, dd,  $J = 12, 6.6$  Hz),  $\delta$  2.57 (1 H, dd,  $J = 13, 7.3$  Hz),  $\delta$  2.52 (1 H, dd,  $J = 13, 6.6$  Hz),  $\delta$  2.1 (1 H, m),  $\delta$  2.05 (3 H, s),  $\delta$  1.66 (1 H, m). Anal. Calcd for C<sub>14</sub>H<sub>20</sub>O<sub>5</sub>: C, 62.67; H, 7.51; O, 29.82. Found: C, 62.03; H, 7.62; O, 29.92.

(21) Cohen, G. H.; Silverton, W. E.; Davies, D. R. *J. Mol. Biol.* **1981**, *148*, 449.

(22) Laitinen, H. A.; Harris, W. E. *Chemical Analysis*, 2nd ed.; McGraw-Hill: New York, 1975; pp 361–363.

**Enzymatically Prepared (R)-5.** In a 20-mL screw-cap scintillation vial, 62 mg (0.3 mol) of **4** was dissolved in 10 mL of dry diisopropyl ether. Then 30 mg of lyophilized *R. miehei* lipase powder was added, followed by 0.4 mL (4 mmol) of vinyl acetate. The vial was shaken at 45 °C and 300 rpm for 12 h. The mixture was filtered, and the enzyme powder was washed 3 times with 10 mL of diisopropyl ether. The washings were combined with the filtrate, and the subsequent workup of the mixture was the same as described above for the chemically synthesized racemic **5**. The yield of (R)-**5** was 34 mg (55% of theor.). The observed  $[\alpha]_{\text{D}}^{28} +24.4^{\circ}$  (CHCl<sub>3</sub>) (83% ee by chiral HPLC) is in agreement with the value of  $[\alpha]_{\text{D}}^{24} +24.2^{\circ}$  (86% ee) reported by Tsuji et al.<sup>23</sup> <sup>1</sup>H NMR (CDCl<sub>3</sub>)  $\delta$  7.35 (5 H, m),  $\delta$  4.28 (1 H, dd,  $J = 4.6, 11.3$  Hz),  $\delta$  4.18 (1 H, dd,  $J = 6.4, 11.2$  Hz),  $\delta$  3.70 (1 H, dd,  $J = 4.6, 11.2$  Hz),  $\delta$  3.60 Hz (1 H, dd,  $J = 6.2, 11.3$  Hz),  $\delta$  2.77 (1H, dd,  $J = 7.8, 11.5$  Hz),  $\delta$  2.73 (1 H, dd,  $J = 7.8, 11.5$  Hz),  $\delta$  2.2 (1 H, m),  $\delta$  2.18 (3H, s),  $\delta$  1.94 (1H, m). Anal. Calcd for C<sub>12</sub>H<sub>16</sub>O<sub>3</sub>: C, 69.21; H, 7.74; O, 23.05. Found: C, 68.71; H, 7.88; O, 23.51.

**Anthroate derivatives of 2 and (R)-5** were prepared as described by Wiesler and Nakanishi.<sup>24</sup> The products were purified by flash column chromatography (1:3 (v/v) ethyl acetate:hexane) and characterized by <sup>1</sup>H NMR (CDCl<sub>3</sub>). For the derivative of **2**:  $\delta$  8.55 (1 H, s),  $\delta$  8.3 (2 H, m),  $\delta$  8.1 (2 H, m),  $\delta$  7.5 (4 H, m),  $\delta$  6.3 (3 H, m),  $\delta$  4.36 (1 H, dd,  $J = 4.7, 11.2$  Hz),  $\delta$  4.29 (1 H, dd,  $J = 6.5, 11.4$  Hz),  $\delta$  4.14 (1 H, dd,  $J = 4.6, 11.3$  Hz),  $\delta$  4.06 (1 H, dd,  $J = 6.4, 11.2$  Hz),  $\delta$  3.74 (6 H, s),  $\delta$  2.57 (1 H, dd,  $J = 8.0, 11.6$  Hz),  $\delta$  2.54 (1 H, dd,  $J = 7.8, 11.4$  Hz),  $\delta$  2.1 (1 H, m),  $\delta$  2.05 (3 H, s). For the derivative of (R)-**5**:  $\delta$  8.57 (1 H, s),  $\delta$  8.3 (2 H, m),  $\delta$  8.0 (2 H, m),  $\delta$  7.5 (4 H, m),  $\delta$  7.3 (5 H, m),  $\delta$  4.51 (1 H, dd,  $J = 4.6, 11.3$  Hz),  $\delta$  4.40 (1 H, dd,  $J = 6.2, 11.3$  Hz),  $\delta$  4.28 (1 H, d,  $J = 4.6, 11.3$  Hz),  $\delta$  4.18 (1 H, dd,  $J = 6.5$  Hz, 11.2 Hz),  $\delta$  2.78 (1 H, dd,  $J = 7.8, 11.5$  Hz),  $\delta$  2.72 (1 H, dd,  $J = 7.7, 11.5$  Hz),  $\delta$  2.2 (1 H, m),  $\delta$  2.18 (3 H, s). Both of the derivatives were further purified by HPLC prior to CD spectrum measurement.

**Enzyme Crystallization.** Crystallization of  $\gamma$ -chymotrypsin followed the method described by Stoddard et al.<sup>25</sup> except that it was scaled up 5-fold. Crystals typically appeared within 1 week and were harvested after 2–3 weeks. The crystals were transferred from the mother liquor to 1.5-mL microcentrifuge tubes, with some 2 mg of crystals placed in each tube. One milliliter of cross-linking solution (1.5% (v/v) glutaraldehyde, 17% (w/v) Na<sub>2</sub>SO<sub>4</sub>, 30 mM sodium cacodylate adjusted to pH 7.5 by 1 M HCl) was added to each tube. After a brief shaking and a 20-min incubation at room temperature, each tube was centrifuged, and the supernate was discarded. The crystals were washed with deionized water (5 times), with 20 mM phosphate buffer, pH 7.8 (5 times), and stored in this buffer for 18 h at 4 °C before use. Subtilisin Carlsberg was crystallized following the procedure of Niedhart and Petsko.<sup>26</sup> Cross-linking of subtilisin was accomplished in the same manner as described above for chymotrypsin. The cross-linked enzyme crystals were washed with water, which was subsequently removed by vacuum filtration prior to placement in organic solvents.

**Enzyme Lyophilization.**  $\gamma$ -Chymotrypsin and *R. miehei* lipase were lyophilized from 5 mg/mL solutions in 20 mM sodium phosphate buffer (pH 7.8 for chymotrypsin, pH 7.0 for lipase). Both enzymes were freeze-dried for at least 24 h.

**Enzyme Precipitation.** One hundred milligrams of  $\gamma$ -chymotrypsin was dissolved in 1 mL of 20 mM Na<sub>2</sub>HPO<sub>4</sub> solution at 4 °C. The pH was adjusted to 7.8, followed by addition of 10 mL of acetone at 0 °C to precipitate the enzyme. The precipitated protein was incubated at 4 °C for 20 min, centrifuged, washed three times with cold acetone, and dried in a vacuum desiccator at 4 °C before use.

**Determination of the Absolute Configuration of 2.** The CD chirality<sup>27</sup> method was used to determine the absolute configuration at C-2 in the enzymatically prepared **2**. Because **2** has a single

chromophore, an additional one was added to facilitate the use of CD. To this end, the free hydroxyl group in **2** was acylated by 9-anthroyl chloride. The absolute configuration of the enzymatically prepared **5** was assigned to be *R* by comparison of optical rotation data with the literature value, and (R)-**5** was used as a reference compound for **2**. The CD spectrum of the anthroate derivative of (R)-**5** was measured in hexane to validate this method.

The molecular models of the anthroate derivatives<sup>28</sup> of (R)-**5** and **2** were constructed using the Insight II and Discover programs. The initial structures were energy-minimized using the steepest descent method for 100 iterations, followed by conjugate gradient minimization until the maximum derivative was less than 0.001 kcal/Å. In the low energy conformation of the derivative of (R)-**5**, the alignment of the electronic transitions in the two chromophores (<sup>1</sup>B<sub>u</sub> transition for the anthroate chromophore and CT transition for the benzyl group) shows positive chirality, which stands for positive first and negative second Cotton effects in the CD spectrum. The situation is the same for the derivative of (R)-**2** and the opposite for that of (S)-**2**.

The CD spectrum of the derivative of (R)-**5** features a positive first Cotton effect around 250 nm and a negative second Cotton effect around 210 nm, indicating positive chirality alignment of the two chromophores. This result is consistent with the known absolute configuration of (R)-**5**. In the CD spectrum of the derivative of **2**, there is a strong positive first Cotton effect at 252 nm and a negative second Cotton effect at 212 nm, indicating that the absolute configuration of enzymatically prepared **2** is *R*.

**Structural Modeling.** The enzyme crystal structures used were those of  $\gamma$ -chymotrypsin in hexane<sup>16</sup> (Brookhaven entry 1GMD) and subtilisin Carlsberg in acetonitrile<sup>29</sup> (Brookhaven entry 1SCB). Because the transition state for the acylation or deacylation of a serine protease is structurally similar to the corresponding tetrahedral intermediate for the reaction,<sup>30</sup> transition states were modeled as the tetrahedral intermediates for the reactions. Such models were produced using a two-step procedure. First, potential binding modes of the chiral products were generated by performing molecular dynamics simulations, followed by energy minimization. The carbonyl oxygen of the product was tethered to the oxyanion binding site using a harmonic potential with a force constant selected to allow widely different conformations to be explored, while preventing the product from diffusing too far from the enzyme. Second, each product binding mode thus identified was used as a template for creating an initial model of the tetrahedral intermediate. The low-energy conformation of each of these starting models was found using molecular dynamics simulations and energy minimizations. The lowest-energy conformer of the tetrahedral intermediate was selected as the model of the transition state.

The first step (the product binding mode search) is necessary because the covalently bound tetrahedral intermediate is sufficiently sterically constrained that molecular dynamics simulations do not sample highly different conformations separated by large energetic barriers. For example, in the case of the *pro-R* transition state for the deacylation of acetyl-chymotrypsin by **1**, an initial tetrahedral intermediate model which starts with the dimethoxyphenyl group bound in the S1 pocket is unable to span the energetic barrier to sample conformations in the S1' binding pocket during molecular dynamics simulation. The product binding mode study identifies both these, as well as other potential starting structures, allowing each of these types of conformations to be examined in the modeling of the tetrahedral intermediate.

Molecular modeling and dynamics simulations were performed with the Insight II and Discover programs<sup>31</sup> as follows: The initial structures were energy-minimized using the steepest descent method for 50 iterations, followed by conjugate gradient minimization until the maximum derivative was less than 0.001 kcal/Å. The minimized structure was then subjected to 40 ps of molecular dynamics at 900 K with steps of 1 fs. After each simulated picosecond, the atomic

(28) Harada, N.; Ono, H.; Uda, H.; Parveen, M.; Khan, N. U.-D.; Achari, B.; Dutta, K. P. *J. Am. Chem. Soc.* **1992**, *114*, 7687.

(29) (a) Fitzpatrick, P. A.; Steinmetz, A. C. U.; Ringe, D.; Klivanov, A. M. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 8653. (b) Fitzpatrick, P. A.; Ringe, D.; Klivanov, A.M. *Biochem. Biophys. Res. Commun.* **1994**, *198*, 675.

(30) Warshel, A.; Naray-Szabo, G.; Sussman, F.; Hwang, J.-K. *Biochemistry* **1989**, *28*, 3629.

(31) Biosym Inc.; San Diego: CA.

(23) Tsuji, K.; Terao, Y.; Achiwa, K. *Tetrahedron Lett.* **1989**, *30*, 6189.

(24) Wiesler, W. T.; Nakanishi, K. *J. Am. Chem. Soc.* **1990**, *112*, 5574.

(25) Stoddard, B. L.; Bruhnke, J.; Porter, N.; Ringe, D.; Petsko, G. A. *Biochemistry* **1990**, *29*, 4871.

(26) Niedhart, D. J.; Petsko, G. A. *Protein Eng.* **1988**, *2*, 271.

(27) Harada, N.; Nakanishi, K. *Circular Dichroic Spectroscopy Excitation Coupling in Organic Stereochemistry*; University Science Books: Mill Valley, CA, 1983.



coordinates were saved, resulting in 40 independent structures with different conformations. The resulting structures were then minimized as outlined above, except the minimization proceeded until the maximum derivative was less than 0.0001 kcal/Å. During all minimizations and molecular dynamics simulations, only the atoms of the substrate and those of the catalytic triad's serine were allowed to move, and a cutoff distance of 11 Å was used with the CVFF force field provided with the Discover program. Because solvent molecules and counterions were not included in the simulations, all protein residues were modeled in their un-ionized forms. Of the 40 minimized structures, the lowest energy conformer was selected, and the solvent-accessible surface area was calculated using the Connolly algorithm, as implemented in the Insight software package.

In support of the validity of the structural modeling methods described above, we were able to use this procedure to correctly predict the conformation of *N*-acetyl-L-phenylalanine trifluoromethyl ketone in its hemiketal complex with chymotrypsin.<sup>32</sup>

**Activity Coefficient Calculation.** All activity coefficients were calculated using the UNIFAC method.<sup>4</sup> Because UNIFAC is a group contribution method, it allows the estimation of activity coefficients in systems for which there is no experimental data by assessing the individual contribution of each group which makes up the system. Use of this method requires three types of parameters for each group in the system: the group's surface area, the volume of the group, and empirically determined parameters which reflect the free energy of interaction between a given group and every other group in the system.

As a test of the accuracy of the UNIFAC calculations, we compared some activity coefficients derived from published vapor-liquid equilibrium (VLE) data to those calculated using UNIFAC. The types of systems for which such data are available are quite limited, but we were able to find VLE data for two compounds (3-methylphenol and 2-methyl-1-propanol) which represent some of the functional groups present in our model molecules in the most nonideal solvent observed in the present work, cyclohexane.<sup>33</sup> Interpreting the VLE data using the Wilson equation of state, for 298 K and a mole fraction of 0.001, the activity coefficients for 3-methylphenol and 2-methyl-1-propanol are 47 and 29, respectively. Under identical conditions, UNIFAC predicts an activity coefficient of 31 for 3-methylphenol, and 21 for 2-methyl-1-propanol. While the individual activity coefficients pre-

dicted by UNIFAC are underestimated by about 30%, the activity coefficient ratio (the quantity used in our work) is estimated to within 6%.

Activity coefficients reported in Table 2 include the effects of 100 mM vinyl acetate and 0.2% (v/v) water.

**Kinetic Measurements.** One milliliter of solvent containing 100 mM vinyl acetate and 10 mM prochiral diol was added to 5 mg of crystals or 15 mg of acetone-precipitated or lyophilized enzyme. Then 0.2% (v/v) water was added to the system to enhance the rate of enzymatic transesterification.<sup>17a</sup> In the presence of the dissolved substrates, the amount of added water was soluble in each of the solvents used. The vinyl acetate hydrolysis product, acetic acid, was not detected during any of the reactions studied. Note that any competing hydrolysis would merely reduce the concentration of acetyl-enzyme available for reaction with the prochiral diol, equally reducing the rate of production of both enantiomers of the chiral monoester product, and thus leaving the prochiral selectivity unaffected. The suspensions were shaken at 45 °C and 300 rpm. Periodically, a 10-μL sample was withdrawn and assayed by chiral HPLC. Because the reactions which lead to the *R*- and *S*-products take place in the same vial, and compete for the same population of free enzyme, the ratio of the initial velocities of the reactions is equal to  $(k_{\text{cat}}/K_{\text{M}})_{\text{pro-}S}/(k_{\text{cat}}/K_{\text{M}})_{\text{pro-}R}$ .<sup>2a,3a</sup> Mass transfer constraints cannot alter the measured initial velocity ratio because both products are generated from the same substrate. Initial velocity ratios were measured two to four times to assure reproducibility. Standard deviations for the ratios were within 13% of the mean values.

Chiral HPLC separations were performed using a Chiralcel OD-H column and a mobile phase of 95:5 (v/v) hexane:2-propanol. A flow rate of 0.8 mL/min separated the enantiomers of **2** with retention times of 29 and 32 min for the *R* and *S* enantiomers, respectively. Chiral resolution of **5** was achieved with a flow rate of 0.5 mL/min, resulting in 30 and 32 min retention times for the respective *R* and *S* enantiomers. The products were quantified using a UV absorbance detector tuned to 220 nm.

**Acknowledgment.** This work was financially supported by NIH Grant GM39794. We thank Ms. Jennifer L. Schmitke for providing the cross-linked crystals of subtilisin and Prof. Gregory Farber and Ms. Anke Steinmetz for helpful advice on the preparation of crystalline  $\gamma$ -chymotrypsin.

JA952674T

(32) Brady, K.; Wei, A.; Ringe, D.; Abeles, R. H. *Biochemistry*, **1990**, *29*, 7600.

(33) Gmehling, J.; Onken, U.; Weidlich, U. Vapor-Liquid Equilibrium Data Collection, Supplement 2d. DECHEMA: Frankfurt, 1982.