Rational Control of Enzymatic Enantioselectivity through Solvation Thermodynamics

Charles R. Wescott,[†] Hidetaka Noritomi,[§] and Alexander M. Klibanov*

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

Received April 29, 1996[⊗]

Abstract: The enantioselectivity of cross-linked crystals of γ-chymotrypsin in the transesterification of the medicinally important compound methyl 3-hydroxy-2-phenylpropionate (**1**) with propanol has been examined in a variety of organic solvents. The $(k_{cat}/K_M)_S/(k_{cat}/K_M)_R$ ratio in this enzymatic process can be forced to span a 20-fold range simply by switching from one solvent to another; in fact, while the enzyme strongly prefers the *S*-enantiomer of **1** in some solvents, the *R*-antipode is more reactive in others. These striking observations are quantitatively rationalized by accounting for the energetics of desolvation of *S*-**1** and *R*-**1** in the enzyme-bound transition states. In order to accomplish this, explicit rules have been established for the modeling and thermodynamic quantification of the partially desolvated substrate's transition state moieties.

Introduction

The exquisite stereoselectivity of enzymes is their most valuable attribute to the organic chemist.[1] Ironically, this same trait also limits the generality of enzymatic synthesis, because enzymes that catalyze the reaction of interest with the desired stereochemistry are not always available. Nonaqueous enzymology,[2] and especially the discovery that enzymatic selectivity can be markedly altered by the reaction medium,[3] thus greatly enhances the utility of enzyme-catalyzed syntheses. Our objective is to elucidate the mechanisms by which the solvent influences enzymatic stereoselectivity, and thereby enable the rational design of stereoselective systems on the basis of physicochemical properties of the substrate and solvent, as well as of enzyme structure.

Our pursuit of this goal first led to a thermodynamic model which explained the solvent dependence of the substrate specificity of subtilisin Carlsberg on the basis of the free energies of desolvation of the substrates.[4] The differential desolvation energy between two substrates was derived from their solvent-to-water partition coefficients, which were either measured experimentally[4a] or calculated[4b] using the UNIFAC[5] computer algorithm. However, since variations in the partition coefficients arose solely from chemical differences in the substrates, this initial model could not account for solvent-induced changes in selectivities involving chemically identical compounds (as in the case of prochiral selectivity or enantioselectivity).

Recently, we have further developed this methodology to explain the solvent effect on enzymatic prochiral selectivity by accounting for the role of partial desolvation of the enzyme−substrate transition state.[6] Transition states leading to the opposite enantiomers of a product may be desolvated to different extents, resulting in a nonzero differential desolvation energy even though the substrates are chemically identical. Implementation of this model to predict the solvent dependence of prochiral selectivity employed molecular modeling to determine the desolvated portions of the substrate in the *pro-R* and *pro-S* transition states, followed by the calculation of the thermodynamic activity coefficients of these moieties. However, due to conceptual difficulties in modeling partially desolvated surfaces of the transition states, application of this methodology was restricted to substrates specifically designed to interact with the enzyme in the desired manner.

In the present work, a set of rules is derived and validated for the modeling of the partially desolvated transition state

* To whom correspondence should be addressed.
† An NIH Biotechnology Predoctoral Trainee.
§ Permanent address: Department of Industrial Chemistry, Tokyo Metropolitan University, Minami-ohsawa, Hachioji, Tokyo 192-03, Japan.
⊗ Abstract published in *Advance ACS Abstracts,* October 15, 1996.

(1) (a) Simon, H.; Bader, J.; Gunther, H.; Neumann, S.; Thanos, J. *Angew. Chem., Int. Ed. Engl.* **1985**, *24*, 539. (b) Yamada, H.; Shimizu, S. *Angew. Chem., Int. Ed. Engl.* **1988**, *27*, 622. (c) Jones, J. B. *Tetrahedron* **1986**, *42*, 3351. (d) Faber, K. *Biotransformations in Organic Chemistry*; Springer-Verlag: Berlin, 1992. (e) Poppe, L.; Novak, L. *Selective Biocatalysis*; VCH Publishers: New York, 1992. (f) Sheldon, R. A. *Chirotechnology: Industrial Synthesis of Optically Active Compounds*; M. Dekker: New York, 1993. (g) Margolin, A. L. *Enzyme Microb. Technol.* **1993**, *15*, 266. (h) Wong, C.-H.; Whitesides, G. M. *Enzymes in Synthetic Organic Chemistry*; Pergamon: Oxford, 1994. (i) Roberts, S. M.; Turner, N. J.; Willetts, A. J.; Turner, M. K. *Introduction to Biocatalysis Using Enzymes and Microorganisms*; Cambridge University Press: New York, 1995. (j) Drauz, K.; Waldmann, H. *Enzyme Catalysis in Organic Synthesis*; VCH Publishers: New York, 1995.

(2) (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) Koskinen, A. M. P.; Klibanov, A. M., Eds. *Enzymatic Reactions in Organic Media*; Blackie: London, 1996.

(3) For reviews, 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.

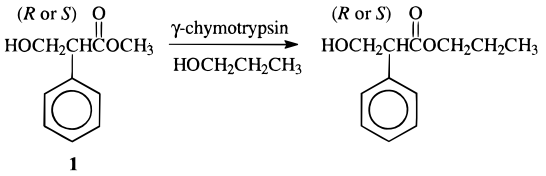
(4) (a) Wescott, C. R.; Klibanov, A. M. *J. Am. Chem. Soc.* **1993**, *115*, 1629. (b) Wescott, C. R.; Klibanov, A. M. *J. Am. Chem. Soc.* **1993**, *115*, 10362.

(5) UNIFAC is a computational method for the estimation of Rault's law activity coefficients. Being a group contribution method, it can calculate activity coefficients in systems for which there is no experimental data by assessing the individual contribution of each chemical 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. (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.

(6) Ke, T.; Wescott, C. R.; Klibanov, A. M. *J. Am. Chem. Soc.* **1996**, *118*, 3366.

S0002-7863(96)01394-7 CCC: $12.00

Scheme 1



1

moieties. Moreover, we expand our predictive repertoire to the solvent dependence of a new, key type of stereoselectivity, namely enantioselectivity. Consequently, we have been able to successfully apply our structure-based thermodynamic treatment to the quantitative analysis of the chiral resolution of a nondesigned, medicinally significant compound, methyl 3-hydroxy-2-phenylpropionate (**1**).

Results and Discussion

To test our methodology with enzymatic enantioselectivity, we have examined the transesterification of racemic **1** with propanol catalyzed by cross-linked crystals[7] (CLCs) of γ-chymotrypsin (Scheme 1). Various esters of the acid moiety of **1** (3-hydroxy-2-phenylpropionic acid, **2**) are potent anticholinergics, including atropine, hyoscyamine, and scopolamine.[8] While most synthetic methods produce racemates of these drugs, only the *S*-antipodes are pharmaceutically active.[8] γ-Chymotrypsin CLCs are employed as the catalyst herein because the crystalline form of the enzyme has been found to retain its native conformation in organic solvents,[9] thus allowing the use of structure-based molecular modeling.

Because enzyme CLCs (as well as nearly all other enzyme preparations) are insoluble in organic solvents, the transesterification in Scheme 1 is catalyzed in a heterogeneous system and thus is susceptible to rate limitation by diffusion of the substrate into the solid catalyst particle. To ensure that the initial velocities measured reflect the true kinetic constants of the enzyme, and not the mass transfer rates of the substrate through the crystals, enzymatic activity was examined as a function of the loading of the biocatalyst particles.[4a,7f,11] To this end, active γ-chymotrypsin was co-crystallized with varying amounts of this enzyme inactivated with diisopropyl fluorophosphate. In the absence of diffusional limitations, a plot of catalytic activity vs the fraction of active enzyme in the crystal should yield a straight line which passes through the origin.[11] If, however, the mass transfer contribution to the reaction rate is not negligible, a convex dependence should be observed,[11] because incremental increases in the enzyme loading do not produce similar amplifications in the catalytic efficiency of the CLC. Such a plot (Figure 1) for the transesterification of both enantiomers of **1** in cyclohexane reveals a linear dependence between catalytic activity and the fraction of active enzyme in the crystal, thus ruling out the possibility that the measured reactions are affected by mass transfer of the substrate.

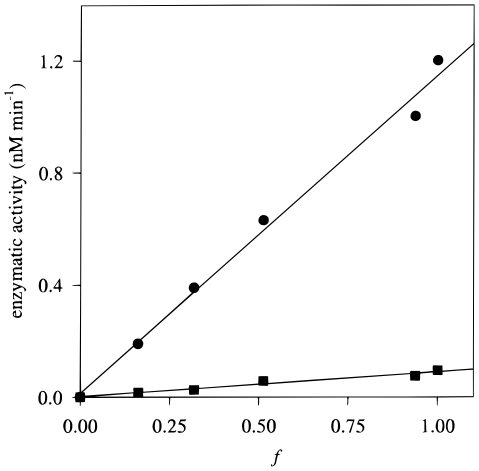


Figure 1. Dependence of the activity of γ-chymotrypsin cross-linked crystals on the fraction of active enzyme in the crystal (*f*). Activity is measured as the rate of enzymatic transesterification of the *S*- (●) or *R*- (■) enantiomers of **1** with propanol in cyclohexane. *f* is controlled by co-crystallizing native and inactivated (by diisopropyl fluorophosphate) γ-chymotrypsin in varying proportions (see Methods for details).

Table 1. Solvent Dependence of the Enantioselectivity of γ-Chymotrypsin Cross-Linked Crystals for the Transesterification of **1** with Propanol

solvent	$(k_{cat}/K_M)_S/(k_{cat}/K_M)_R$[a]	product ee, %[b] (preferred enantiomer)
cyclohexane	13	85 (*S*)
octane	8.8	79 (*S*)
hexane	8.0	77 (*S*)
toluene	5.6	69 (*S*)
isopropyl acetate	2.4	40 (*S*)
tetrahydrofuran	1.8	28 (*S*)
tert-butyl acetate	1.5	20 (*S*)
tert-butyl alcohol	0.91	4.6 (*R*)
tert-amyl alcohol	0.80	11 (*R*)
dioxane	0.74	15 (*R*)
propanol	0.73	15 (*R*)
acetone	0.64	21 (*R*)

[a] See Methods for details on the measurement of $(k_{cat}/K_M)_S/(k_{cat}/K_M)_R$. [b] Enantiomeric excesses (ee) were calculated from the enantioselectivities for a 5% conversion as described by Chen et al.[12]

To explore the effect of the solvent on the kinetic resolution of **1**, the enantioselectivity of γ-chymotrypsin CLCs for the reaction depicted in Scheme 1 was measured in a variety of organic solvents. The enantiomeric excess (ee) at 5% conversion was subsequently calculated[12] to quantify the efficiency of the resolution in each solvent (Table 1). Inspection of Table 1 reveals that the enantioselectivity, expressed as $(k_{cat}/K_M)_S/(k_{cat}/K_M)_R$, can be forced to span a 20-fold range simply by switching from one organic solvent to another under otherwise identical conditions. Perhaps even more striking is the fact that the enantioselectivity can actually be reversed through the choice

(7) Such cross-linked crystals of various enzymes 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.; Montelbano, 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. (e) Sobolov, S. B.; Leonida, M. D.; Bartoszko-Malik, A.; Veivodov, K. I.; McKinney, F.; Kim, J.; Fry, A. J. *J. Org. Chem.* **1996**, *61*, 2125. (f) Schmitke, J. L.; Wescott, C. R.; Klibanov, A. M. *J. Am. Chem. Soc.* **1996**, *118*, 3360.

(8) Reynolds, J. E. F., Ed. *Martindale, the Extra Pharmacopoeia*, 28th ed.; Pharmaceutical Press: London, 1982; p 304.

(9) It is reasoned that if the structure of uncross-linked chymotrypsin is nearly the same in water and in hexane,[10] 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. Furthermore, cross-linking followed by placement in acetonitrile does not affect the structure of two other serine proteases, subtilisin Carlsberg (Fitzpatrick, P. A.; Steinmetz, A. C. U.; Ringe, D.; Klibanov, A. M. *Proc. Natl. Acad. Sci. U.S.A.* **1993**, *90*, 8653. Fitzpatrick, P. A.; Ringe, D.; Klibanov, A. M. *Biochem. Biophys. Res. Commun.* **1994**, *198*, 675) and porcine pancreatic elastase (Allen, K. N.; Bellamacina, C. R.; Ding, X.; Jeffery, C. J.; Mattos, C.; Petsko, G. A.; Ringe, D. *J. Phys. Chem.* **1996**, *100*, 2605).

(10) (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.

(11) Boudart, M.; Burwell, R. L. In *Techniques in Chemistry*, 3rd ed.; Lewis, E. S., Ed.; Wiley: New York, 1973; Vol. 6, Chapter 12.

(12) Chen, C.-S.; Fujimoto, Y.; Girdaukas, G.; Sih, C. J. *J. Am. Chem. Soc.* **1982**, *104*, 7294.

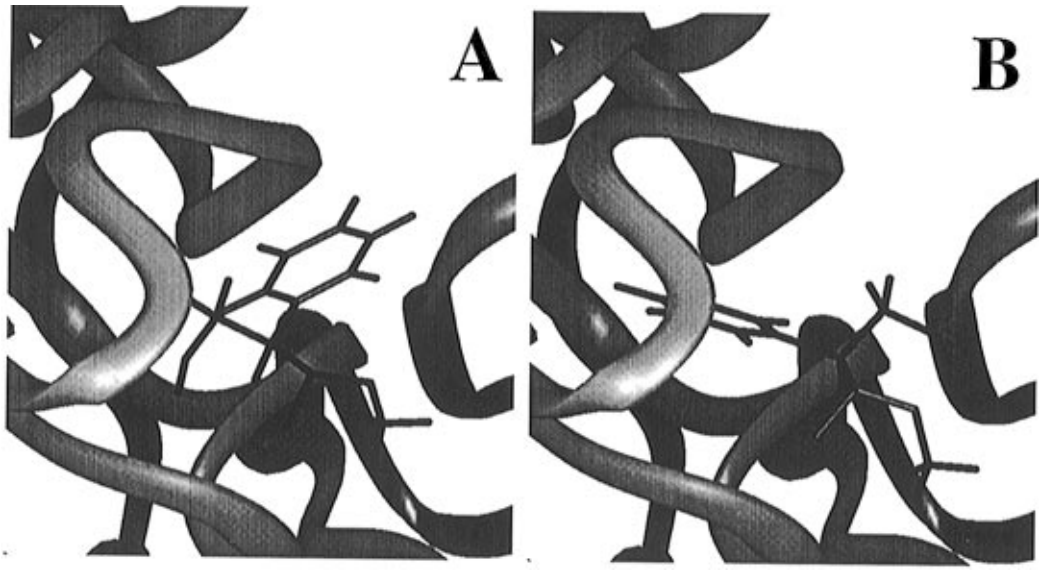


Figure 2. Molecular models of *S*- (A) or *R*- (B) **1** in the transition state for acylation of γ-chymotrypsin. See Methods for details on creation of the models.

of the solvent. For instance, in cyclohexane the enzyme preferentially transesterifies the *S*-enantiomer of **1**, while in acetone the *R*-antipode is preferred.

It has been shown that the solvent exerts its effect on the selectivity of enzyme-catalyzed processes through the differential Gibbs free energy of desolvation for the transition states of the reactions.[4,6] A quantitative analysis[6,13] of such a system, when applied to enantioselectivity herein, affords an equation relating the latter in a given solvent to the thermodynamic activity coefficients of the desolvated moieties of the substrate transition states (γ') in that solvent:

$$\log\left[\frac{(k_{cat}/K_M)_S}{(k_{cat}/K_M)_R}\right] = \log\left(\frac{\gamma'_S}{\gamma'_R}\right) + \text{constant} \qquad (1)$$

where subscripts *S* and *R* indicate parameters related to the respective enantiomers of the substrate, and "constant" contains kinetic and thermodynamic terms in an arbitrarily selected reaction medium.[6]

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 them,[4] γ' for enantioselectivity differs for each reaction pathway only due to differences in transition state solvation. In this work, we calculate γ' for both the *R* and *S* transition states using a three-step procedure. First, the desolvated portion of each enantiomer 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 approximated in terms of individual UNIFAC groups. Finally, the thermodynamic activity coefficient of this desolvated fragment is calculated using the UNIFAC group contribution method and then equated to γ'. According to eq 1, knowing only the ratio of γ'_R and γ'_S for a series of solvents, it should be possible to explain the observed solvent dependence of enantioselectivity.

As a first step in the calculation of γ', molecular models have been constructed for the *S* and *R* transition states for the acylation of γ-chymotrypsin by **1** (see Methods for details). Examination of the *S* transition state model (Figure 2A) reveals that the hydroxyl group of the substrate is buried in chymotrypsin's S1 binding pocket, while the phenyl group extends away from the enzyme toward the solvent. Figure 2B depicts the opposite situation for the *R* transition state: the aryl moiety is buried in the active center of the enzyme, while the hydroxyl group is oriented toward the solvent. The solvated surface areas for the transition states, calculated using the method of Connolly,[14] are displayed as dot surfaces in Figure 3. One can see that, for example, in the *S* transition state the hydroxyl group is desolvated, while the phenyl group is not. In contrast, the surface for the *R* transition state indicates the inverse desolvation pattern for these two groups.

With the desolvated portions of the transitions states ascertained, the next step in the calculation of γ' is the construction of molecular fragments, based on UNIFAC groups,[5] which approximate the desolvated portions of the substrates in the transition states. To this end, the enantiomers of **1** have been modeled in terms of the smallest possible UNIFAC groups, and the percent of desolvation of each such group is tabulated in Table 2. Groups are then included in the molecular fragment for a given substrate enantiomer if they are at least 50% desolvated. Groups desolvated to a lesser extent are considered solvated and thus not part of the desolvated substrate moiety. According to these rules, the desolvated portion of the *S* transition state is represented by one hydroxyl group, one aryl carbon, two aryl methine groups, one carbonyl group, and one aliphatic methine group. Similarly, the corresponding *R* molecular fragment consists of one aryl carbon, three aryl methine groups, one carbonyl group, and one aliphatic methine group.

Finally, the activity coefficients for the *S* and *R* model fragments are calculated using UNIFAC and equated with the activity coefficients of the desolvated portions of the corresponding transition states, γ'_S and γ'_R, respectively.

Equation 1 predicts that a double logarithmic plot of enantioselectivity vs the γ' ratio should be linear, with a slope of unity. Such a plot, presented in Figure 4A, does indeed follow the expected dependence: a least-squares fit to a linear model with a slope of 1 yields a correlation coefficient of 0.87. Furthermore, when the data are plotted in linear coordinates (Figure 4B), linear regression yields a correlation coefficient of 0.93. Thus eq 1 correctly predicts the solvent dependence of enzymatic enantioselectivity in a near quantitative fashion.

(13) This analysis assumes the absence of specific enzyme−solvent interactions; such interactions have been deemed to be negligible under similar circumstances.[4,6,7f]

(14) Connolly, M. L. *Science* **1983**, *221*, 709.

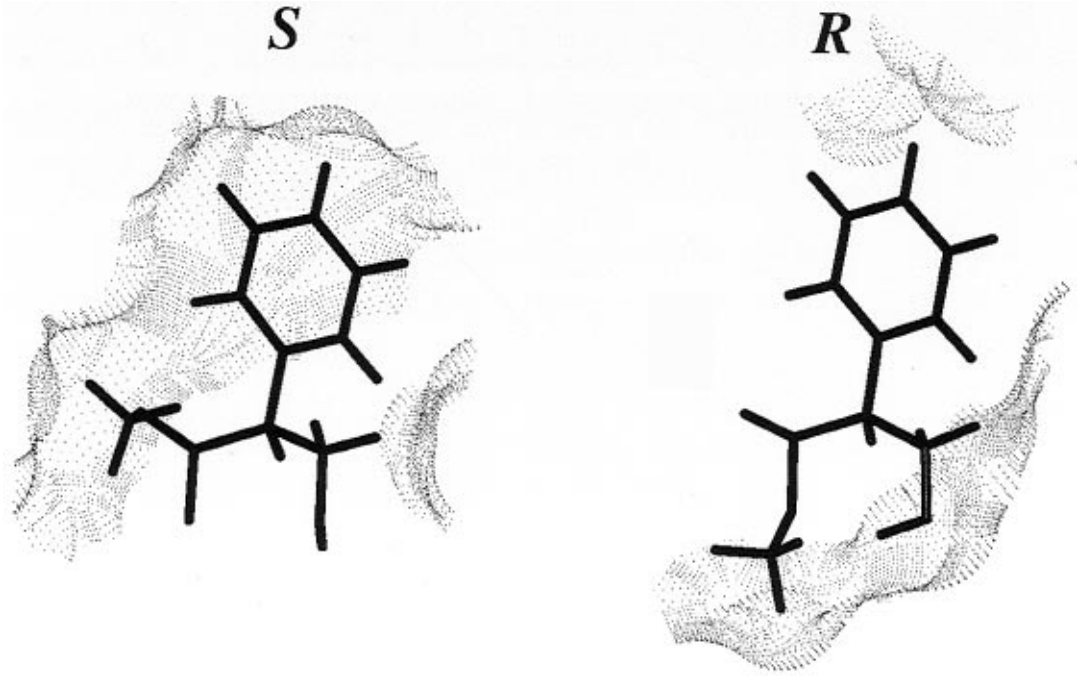


Figure 3. Solvent-accessible surface areas of **1** in the *S* (left) and *R* (right) transition states with γ-chymotrypsin. See Methods for details.

Table 2. Percent of Desolvation of Component Groups for *S*- or *R*-Enantiomers of the Transition States for the Acylation of γ-Chymotrypsin by **1**

	desolvation (%)[b]	
group[a]	*S*	*R*
hydroxyl	100	11
aryl C 1	100	100
aryl CH 2	86	35
aryl CH 3	0	100
aryl CH 4	11	54
aryl CH 5	42	33
aryl CH 6	67	100
carbonyl	100	100
methylene	49	36
methine	100	100
methyl	13	35

[a] The aryl units make up the phenyl group of **1**. [b] The percent of desolvation of a group is calculated as $[1 - (A_B/A_F)] \times 100\%$, where A_B is the solvent-accessible surface area of the substrate group in the enzyme-bound transition state (Figure 3), and A_F is the solvent-accessible surface area of the same substrate group in the free transition state.

The methodology employed above is surprisingly effective despite the fact that, while most substrate groups are in reality only partially desolvated in the transition state (Table 2), for the construction of the model fragments all groups are approximated as either 100% or 0% desolvated. Why does this approach work even though a continuous range of desolvated surface areas are approximated by a simple "all-or-nothing" model? To answer this question, we have examined how the method of treatment of partially desolvated groups (herein for simplicity defined as groups which are 20% to 80% desolvated) affects the predictive performance of our model. One extreme alternative to the partially desolvated groups is to include all of them in the model fragments. When these model fragments are used to calculate the activity coefficients of the desolvated portions of the *S* and *R* transition states, and the data are plotted as in Figure 4A, the resultant correlation coefficient (0.89) has been found to be virtually unaffected. A second extreme recourse to the treatment of the partially desolvated groups is to exclude all of them from the model fragments. Once again, fitting these data to a linear dependence with a slope of unity yields an unchanged correlation coefficient of 0.87. These results lead to the conclusion that, for the reaction in Scheme 1, the solvent dependence of the γ' ratio is quite insensitive to those moieties that happen to be partially desolvated and is determined only by those that happen to be essentially desolvated in the transition state (Table 2).

In order to rationalize why the groups partially desolvated in our case do not appreciably affect the solvent dependence of the γ' ratio, we have calculated the activity coefficients of all the individual component groups. As seen in Table 3, the activity coefficients of the partially desolvated groups vary relatively little throughout the series of solvents tested (0.26, 0.67, and 0.44 for the aryl CH, methyl, and methylene groups, respectively). In contrast, two of the groups for which the extent of desolvation is clearly defined, namely the hydroxyl and carbonyl groups, exhibit much larger respective activity coefficient changes of 8.0 and 2.9. Although the calculations of activity coefficients for individual groups are inherently qualitative,[15] they suggest that the solvent dependence of the γ' ratio is dominated by the interaction of the hydroxyl and carbonyl groups with the organic solvents. Consistent with this conclusion is the finding that γ' ratios calculated by including the hydroxyl in, or excluding it from, both transition states produce a correlation coefficient below 0.1 when the data are plotted as in Figure 4A. To generalize, different types of groups impact the free energy of desolvation of the transition states to varying extents. Therefore, in choosing a biocatalyst for the resolution of a chiral compound, one should seek an enzyme with an active center that maximizes the difference in desolvation of the "impactful" groups (such as OH) between the two enantiomers. This optimization of differential desolvation of the impactful groups can be performed at the expense of the nonimpactful groups without consequence.

An insight into the nature of the solvent−solute interactions can also be gained from Table 3. The activity coefficients for most of the groups in the table are below unity, indicating

(15) It should be cautioned that these comparisons are qualitative and indicative of general trends only. This is because the standard states of the individual group activity coefficients are different, and the contribution of each group to the overall activity coefficient of the substrate model is affected by inter-group interaction parameters.[5]

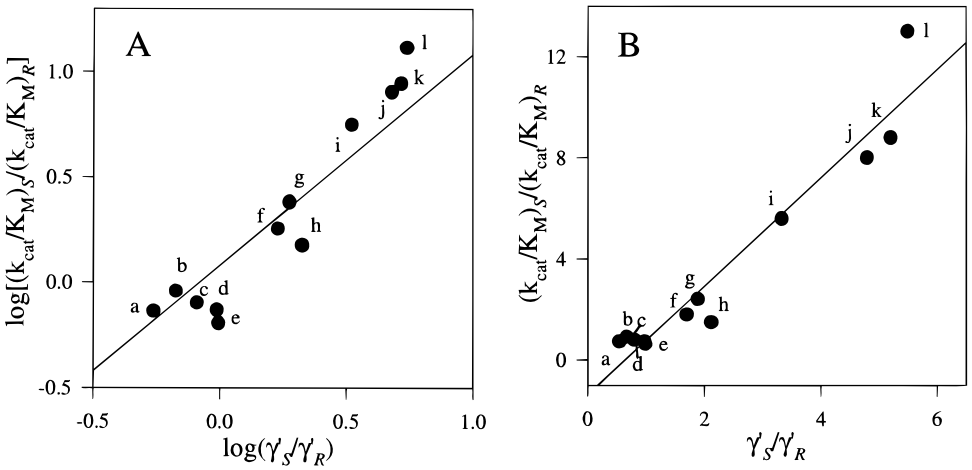


Figure 4. Dependence of the enantioselectivity of the transesterification in Scheme 1 catalyzed by cross-linked crystals of γ-chymotrypsin on the activity coefficient ratio for the desolvated portions of the substrates in the enzyme-bound transition states. (A) Double logarithmic plot with a least-squares fit to a linear dependence with the slope of unity, used to assess predictive ability of eq 1 (correlation coefficient 0.87). (B) Linear plot with linear regression (correlation coefficient 0.93). Solvents: (a) propanol, (b) *tert*-butyl alcohol, (c) *tert*-amyl alcohol, (d) dioxane, (e) acetone, (f) tetrahydrofuran, (g) isopropyl acetate, (h) *tert*-butyl acetate, (i) toluene, (j) hexane, (k) octane, (l) cyclohexane. See Methods for experimental details.

Table 3. Solvent Dependence of the Activity Coefficients of Component Groups of the Model Fragments for the Enantiomers of **1**

	activity coefficients[a]						
solvent	hydroxyl	aryl C	aryl CH	carbonyl	methyl	methylene	methine[b]
cyclohexane	9.3	0.35	0.37	4.1	0.56	0.42	0.34
octane	6.8	0.28	0.28	3.3	0.43	0.33	0.27
hexane	7.7	0.35	0.35	3.8	0.51	0.40	0.33
toluene	5.4	0.34	0.35	1.7	0.63	0.45	0.35
isopropyl acetate	3.0	0.37	0.35	1.6	0.65	0.48	0.38
tetrahydrofuran	3.6	0.51	0.46	1.7	0.76	0.60	0.51
tert-butyl acetate	3.0	0.33	0.31	1.6	0.56	0.42	0.34
tert-butyl alcohol	1.3	0.43	0.42	2.5	0.71	0.54	0.43
tert-amyl alcohol	1.4	0.37	0.37	2.3	0.61	0.46	0.37
dioxane	2.1	0.55	0.47	1.2	0.90	0.69	0.56
propanol	1.3	0.51	0.50	2.7	0.86	0.64	0.51
acetone	2.5	0.56	0.54	1.6	1.1	0.77	0.59

[a] Activity coefficients were calculated using UNIFAC[5] (see Methods for details). The aryl units make up the phenyl group of **1**. [b] Aliphatic.

thermodynamic stabilization of most groups by the solvent. The hydroxyl and carbonyl groups, however, consistently feature activity coefficients greater than unity, i.e., these groups are destabilized by the solvents. Solvent control of the γ' ratio (and thereby the enantioselectivity) for this system is thus exercised through thermodynamically unfavorable interactions between the solvents and the hydroxyl and carbonyl groups of the substrate.

Because a given substrate molecule can be represented by several different combinations of groups, it is important to assess the effect of approximating the substrate using different types of fragments. To this end, substrate **1** has been modeled using a series of incrementally larger groups, and the ability of each member of the series to predict the solvent dependence of enantioselectivity has been assessed via the ensuing correlation coefficient. One can see in Table 4 that as the size of the groups increases, the ability to approximate the desolvated portion of the substrate in the transition state deteriorates. This results in increasing error in the calculation of the γ' ratio, which erodes and ultimately destroys the predicted power of eq 1. Therefore, to optimize the performance of this methodology, one should model the substrate in terms of the smallest possible fragments for the determination of γ'.

In closing, the present study constitutes a further step toward the ultimate goal of predicting enzymatic selectivity on the basis of the enzyme structure and physicochemical properties of the substrates and solvents.

Materials and Methods

Enzymes. Native and diisopropyl−fluorophosphate-inactivated α-chymotrypsins (EC 3.4.21.1) were purchased from Sigma Chemical Co. γ-Chymotrypsin crystals were created from the α-form of the enzyme, following the method of Stoddard *et al.*[16] Partially inactivated γ-chymotrypsin crystals (used in the diffusional limitation experiments) were grown from mother liquors containing varying ratios of the diisopropyl−fluorophosphate-inactivated enzyme. Co-crystallization of the native and inhibited forms of the enzyme was confirmed by measuring the activity of single crystals dissolved in water.[7f] For use in organic solvents, crystals were cross-linked and prepared for catalysis as described previously.[6]

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[17]) by shaking with Linde's 3-Å molecular sieves.

1 was synthesized by refluxing 1 g of **2** in 25 mL of anhydrous methanol containing 5 drops of concentrated H_2SO_4 for 12 h. The reaction mixture was subsequently concentrated by rotary evaporation, dissolved in 50 mL of diethyl ether, and then washed with five 10-mL aliquots of 5% $NaHCO_3$ and with 10 mL of deionized water. The crude product was recovered from the organic phase by rotary evaporation

(16) Stoddard, B. L.; Bruhnke, J.; Porter, N.; Ringe, D.; Petsko, G. A. *Biochemistry* **1990**, *29*, 4871.
(17) Laitinen, H. A.; Harris, W. E. *Chemical Analysis*, 2nd ed.; McGraw-Hill: New York, 1975; pp 361−363.

Table 4. Correlation Coefficients for Representations of the Transition State of **1** Using Successively Larger Groups

group	desolvation (%)[a] *S*	desolvation (%)[a] *R*	corr coeff[b]
representation 1			0.87
see Table 2			
representation 2			0.80
hydroxyl	100	11	
phenyl	38	74	
methyl	49	36	
methyl acetate	54	59	
representation 3			0.78
methanol	70	26	
phenyl	38	74	
methyl acetate	54	59	
representation 4			0.68
ethanol	73	36	
phenyl	38	74	
methyl formate	50	56	
representation 5			0.00
2-phenylethanol	46	62	
methyl formate	50	56	

[a] The percent of desolvation of a group is calculated as described in footnote *b* to Table 2. [b] To assess the performance of each representation of the substrate molecule in the prediction of enantioselectivity, a double logarithmic plot of the latter vs the activity coefficient ratio of the desolvated substrate moiety in the enzyme-bound transition state (calculated for the indicated representation of the substrate) was fit to a linear dependence with a slope of unity (as predicted by eq 1), and the resultant correlation coefficient was calculated. By fixing the slope to 1, both random and systematic errors (due to deviation of the slope from unity) are reflected in the correlation coefficient.

and subsequently purified by vacuum distillation. ^{1}H NMR ($CDCl_3$) δ 7.3−7.4 (5 H, m), 4.1−4.2 (1 H, m), 3.8−3.9 (2 H, m), 3.7 (3 H, s), 2.2 (1 H, s).

Propyl 3-hydroxy-2-phenylpropionate, a racemic mixture used to calibrate the HPLC instrument, was synthesized from **2** following the procedure used for **1**, except that methanol was substituted with 50 mL of propanol. ^{1}H NMR ($CDCl_3$) δ 7.3−7.4 (5 H, m), 4.1−4.2 (1 H, m), 4.1 (2 H, t, J = 6.0 Hz), 3.8−3.9 (2 H, m), 1.8 (1 H, s), 1.6−1.7 (2 H, m), 0.8−0.9 (3 H, t, J = 7.7 Hz).

Propyl (*S*)-3-hydroxy-2-phenylpropionate, used to assign the *S*-product HPLC peak, was synthesized by refluxing 0.5 g of scopolamine·HBr in 10 mL of anhydrous propanol containing 5 drops of concentrated H_2SO_4 for 3 days. The crude product was recovered as in the synthesis of **1** and purified by TLC. ^{1}H NMR ($CDCl_3$) δ 7.3−7.4 (5 H, m), 4.1−4.2 (1 H, m), 4.1 (2 H, t, J = 6.0 Hz), 3.8−3.9 (2 H, m), 2.1 (1 H, s), 1.6−1.7 (2 H, m), 0.8−0.9 (3 H, t, J = 7.7 Hz).

Kinetic Measurements. One milliliter of solvent containing 100 mM racemic **1** and 100 mM propanol was added to 10 mg of cross-linked γ-chymotrypsin crystals. Then 0.2% (v/v) water was added to the suspension to enhance the rate of enzymatic transesterification. To verify that the added water was soluble in each of the solvent systems, an even larger amount of water, 0.3% (v/v), was added to 1 mL of octane (the most hydrophobic solvent used herein) containing 100 mM ester substrate and 100 mM propanol. The mixture was shaken at 45 °C for 15 min and allowed to settle for 10 min. Then 0.5 mL was withdrawn from the top of the solution and the water content was determined to be 1.46 mg (i.e., approximately 0.3% (v/v)) by Karl Fischer titration. The hydrolysis product **2** was not detected during any of the reactions studied. The effect of the added water on γ' is accounted for by the explicit inclusion of 0.2% (v/v) water in the UNIFAC calculations (see Activity Coefficient Calculation below for details). Note that any competing hydrolysis would merely reduce the concentration of the acyl-enzyme available for reaction with propanol, equally reducing the rate of production of both enantiomers of the propyl ester product, and thus leaving the enantioselectivity 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 transesterifications which lead to the *R* and *S* products take place in the same reaction mixture and the substrate enantiomers compete for the same population of free enzyme, the ratio of initial velocities of the reactions is equal to $(k_{cat}/K_M)_S/(k_{cat}/K_M)_R$.[3a,4a]

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 0.5 mL/min separated the *R* and *S* enantiomers of **1** with retention times of 17 and 19 min, respectively. The products were quantified using a UV absorbance detector at 220 nm.

Activity Coefficient Calculation. All activity coefficients were calculated using the UNIFAC group contribution method.[5] Calculations explicitly included the effects of 100 mM propanol and 0.2% (v/v) water.

Structural Modeling. Molecular models were produced using the crystal structure of γ-chymotrypsin in hexane (Brookhaven data bank entry 1GMC).[10] Because the transition state for the acylation of a serine protease is structurally similar to the corresponding tetrahedral intermediate for the reaction,[18] transition states were modeled as the tetrahedral intermediates for the reactions. Such models were produced using the two-step procedure outlined below and described in detail previously.[6] First, potential binding modes of each enantiomer of the substrate were generated by performing molecular dynamics simulations, followed by energy minimization. The carbonyl oxygen of the substrate 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 substrate from diffusing too far from the enzyme. This substrate 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 the second step, each substrate 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 solvent-accessible surface areas were calculated by the method of Connolly.[14]

JA961394Q

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

(19) This work was financially supported by NIH grant GM39794. H.N. gratefully acknowledges a fellowship from the Tokyo Metropolitan Government.