J|A|C|S ARTICLES

Published on Web 02/25/2009

Quantitative DFT Modeling of the Enantiomeric Excess for Dioxirane-Catalyzed Epoxidations

Severin T. Schneebeli, Michelle Lynn Hall, Ronald Breslow,* and Richard Friesner*

Department of Chemistry, Columbia University, New York, New York 10027

Received September 2, 2008; E-mail: rb33@columbia.edu; rich@chem.columbia.edu

Abstract: Herein we report the first fully quantum mechanical study of enantioselectivity for a large data set. We show that transition state modeling at the UB3LYP-DFT/6-31G* level of theory can accurately model enantioselectivity for various dioxirane-catalyzed asymmetric epoxidations. All the synthetically useful high selectivities are successfully "predicted" by this method. Our results hint at the utility of this method to further model other asymmetric reactions and facilitate the discovery process for the experimental organic chemist. Our work suggests the possibility of using computational methods not simply to explain organic phenomena, but also to predict them quantitatively.

Introduction

The development of tools to quantitatively model the stereochemical outcome of reactions is of great interest to the chemical community, as experimental screening of new compounds for enantioselective reactions is often labor-intensive and expensive. Furthermore, modeling stereochemistry presents a formidable challenge to the computational chemist, where a difference of only 1.8 kcal/mol in the stereochemically distinct transition states leads to product ratios of 96:4.

Moitessier et al. have recently proposed a computational tool to model diastereoselectivity employing a force-field-based approach.[1] Moitessier approximates transition state geometries as linear combinations of reactant and product geometries. Transition state energies are approximated by multiplying the MM3* force field terms for the forming and breaking bonds with a linear scaling factor. Of course, such an approach has the advantage that the computations are orders of magnitude faster than quantum mechanical calculations and therefore allow more efficient sampling of conformational space. However, the accuracy of such a crude transition state model is questionable, especially for flexible transition states. A quantum mechanical approach, however, promises higher accuracy for the transition state geometries and energies and does not require the parametrization of a transition state force field. The major disadvantage of a quantum mechanical approach is, of course, the much longer times required to run the calculations. The Jaguar suite of quantum chemistry programs[2] allows efficient computation of molecular integrals by exploiting pseudospectral methods[16] and therefore opens the possibility to model enantioselectivity of reactions of medium-sized organic molecules within affordable time. As a rough estimate, the computations described in this work took on the order of one week on a standard PC for a medium-sized transition state of the data set.

Herein we report the first fully quantum mechanical study of enantioselectivity for a large data set. Our study differs significantly from previous quantum mechanical studies of asymmetric reactions, which have been largely mechanistic or limited to only a handful of substrates. Bach et al. first explored computational methods to help understand asymmetric epoxidations,[3] paving the way for Houk and co-workers,[4] who would later contribute much to the understanding of the reaction's stereoselectivity. Houk showed that stereoselectivity in asymmetric epoxidations is qualitatively controlled by torsional steering, a differential stability of transition states mediated by varying degrees of torsional strain.[4] It has also been shown that the dioxirane-mediated epoxidation, among other epoxidations, is well described by B3LYP.[4,5] Sarzi-Amadè et al. have published several studies exploring the qualitative nature of enantioselectivity for peroxy acids and dioxiranes, utilizing B3LYP to obtain good qualitative agreement with experimental data.[6] Singleton et al. have also demonstrated that B3LYP can be used to predict kinetic isotope effects in very good agreement

(1) (a) Corbeil, C. R.; Thielges, S.; Schwartzentruber, J. A.; Moitessier, N. *Angew. Chem., Int. Ed.* **2008**, *47*, 2635–2638. (b) Moitessier, N.; Henry, C.; Len, C.; Chapleur, Y. *J. Org. Chem.* **2002**, *67*, 7275–7282.

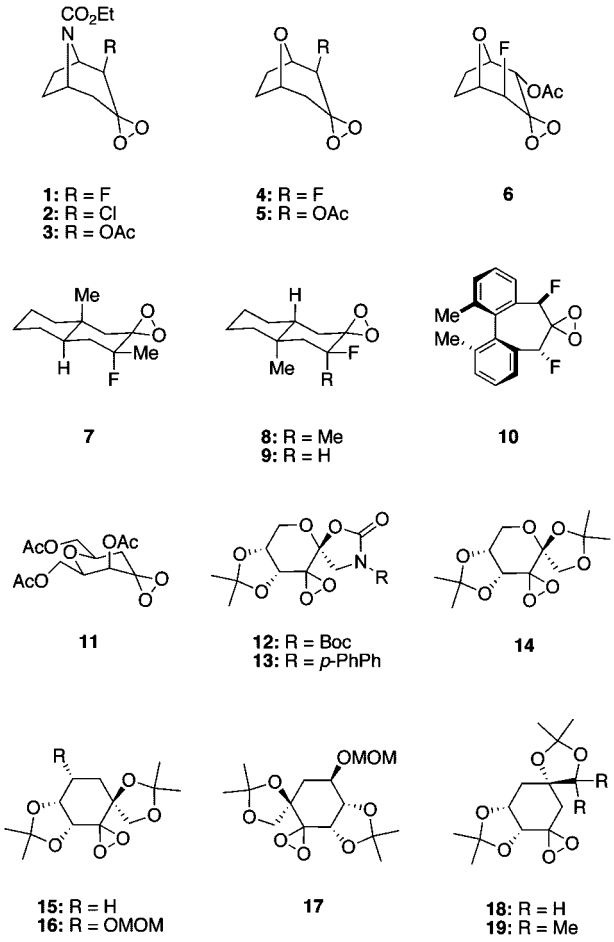
(2) *Jaguar 7.0*; Schrödinger, Inc.: Portland, OR, 2007.

(3) (a) Lang, T. J.; Wolber, G. J.; Bach, R. D. *J. Am. Chem. Soc.* **1981**, *103*, 3275. (b) Bach, R. D.; Wolber, G. J. *J. Am. Chem. Soc.* **1984**, *106*, 1410–1415. (c) Bach, R. D.; Coddens, B. A.; McDouall, J. J. W.; Schlegel, H. B.; Davis, F. A. *J. Org. Chem.* **1990**, *55*, 3325–3330. (d) Bach, R. D.; Owensby, A. L.; Gonzalez, C.; Schlegel, H. B.; McDouall, J. J. W. *J. Am. Chem. Soc.* **1991**, *113*, 2338–2339. (e) Bach, R. D.; Andres, J. L.; Davis, F. A. *J. Org. Chem.* **1992**, *57*, 613–618. (f) Bach, R. D.; Andres, J. L.; Owensby, A. L.; Schlegel, H. B.; McDouall, J. J. W. *J. Am. Chem. Soc.* **1992**, *114*, 7207–7217. (g) Bach, R. D.; Winter, J. E.; McDouall, J. J. W. *J. Am. Chem. Soc.* **1995**, *117*, 8586–8593. (h) Bach, R. D.; Glukhovtsev, M. N.; Gonzalez, C.; Marquez, M.; Estevez, C. M.; Baboul, A. G.; Schlegel, H. B. *J. Phys. Chem. A* **1997**, *101*, 6092–6100. (i) Bach, R. D.; Dmitrenko, O.; Adam, W.; Schambony, S. *J. Am. Chem. Soc.* **2003**, *125*, 924–934. (j) Dmitrenko, O.; Bach, R. D. *J. Phys. Chem. A.* **2004**, *108*, 6886–6892.

(4) (a) Houk, K. N.; Liu, J.; DeMello, N. C.; Condroski, K. R. *J. Am. Chem. Soc.* **1997**, *119*, 10147–10152. (b) Jenson, C.; Liu, J.; Houk, K. N.; Jorgenson, W. L. *J. Am. Chem. Soc.* **1997**, *119*, 12982–12983. (c) Washington, I.; Houk, K. N. *Angew. Chem., Int. Ed.* **2001**, *40*, 4485–4488. (d) Armstrong, A.; Washington, I.; Houk, K. N. *J. Am. Chem. Soc.* **2000**, *122*, 6297–6298. (e) Lucero, M. J.; Houk, K. N. *J. Org. Chem.* **1998**, *63*, 6973–6977.

(5) (a) Singleton, D. A.; Merrigan, S. R.; Liu, J.; Houk, K. N. *J. Am. Chem. Soc.* **1997**, *119*, 3385–3386. (b) Singleton, D. A.; Wang, Z. *J. Am. Chem. Soc.* **2005**, *127*, 6679–6685. (c) Christian, C. F.; Takeya, T.; Szymanski, M. J.; Singleton, D. A. *J. Org. Chem.* **2007**, *72*, 6183–6189.

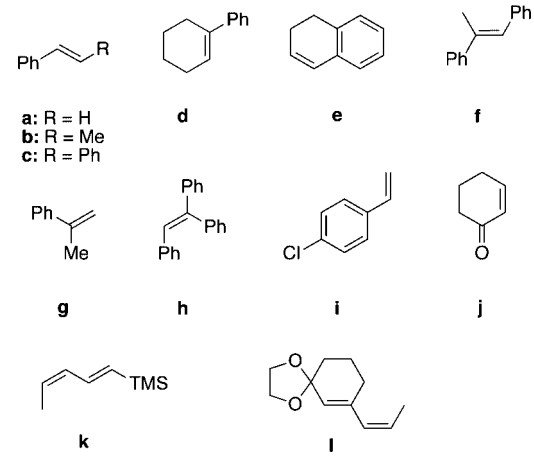
Chart 1. Dioxirane Catalysts Employed in This Study



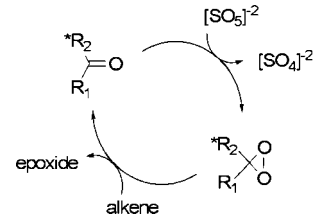
with experiment for the Shi epoxidation,[5b] further suggesting B3LYP's ability to accurately model dioxirane-mediated epoxidations even for larger complex systems. A recent study by Markad et al. discusses the facial selectivity in the dimethyldioxirane (DMDO) epoxidation of carbohydrate-based oxepines based on B3LYP calculations and compares the computational results to experimental findings.[6i]

Intrigued by whether computational chemistry could be used to explain organic phenomena of practical interest quantitatively, we employed transition state theory to model enantioselectivity on a large data set. Our test set comprised 46 asymmetric dioxirane-catalyzed epoxidation reactions, with the dioxirane catalysts and olefins shown in Charts 1 and 2, respectively. The data set is notable for the structural diversity of the dioxirane catalysts and of the olefin substrates as well as the broad range of experimental enantiomeric excess values (ranging from 0 to 98% ee). Furthermore, the data set is also marked by various

Chart 2. Olefins Employed in This Study



Scheme 1. Catalytic Cycle for the Asymmetric Epoxidation Reaction of Olefins with Dioxiranes



solvent systems and reaction temperatures. By utilizing a data set of such diversity, we aimed to thoroughly explore the boundaries and limitations of quantum mechanical methods to accurately reproduce experimental data with a keen eye on the future goal of prediction.

Our current test set includes many practically useful dioxirane-catalyzed epoxidations. There are, however, still a few interesting dioxirane catalysts absent from the data set, such as the binaphthyl-based catalysts reported by Yang.[7] Extending our data set to the aforementioned additional asymmetric epoxidation reactions as well as to other reaction types would, of course, be of interest but is beyond the scope of the current study. Furthermore, we propose that the diversity and size of our current data set is sufficient to excuse the aforementioned absences, being well beyond the scope of all previous studies.

The mechanism of the catalytic dioxirane-mediated epoxidation reaction of olefins is well understood and proceeds through the catalytic cycle shown in Scheme 1.[8] Computations by Bach et al.[3] and Houk et al.[4a] have shown the epoxidation of olefins by dioxiranes to proceed through a "spiro" transition state, as shown in Figure 1a. Although planar transition state structures (Figure 1b) are invoked by Shi and Bartlett in some cases,[9,10] computational evidence continues to support a spiro transition state[3a] or a hybrid planar-spiro transition state in more complex systems, as shown by Singleton.[5b]

(6) (a) Freccero, M.; Gandolfi, R.; Sarzi-Amadè, M.; Rastelli, A. *J. Org. Chem.* **2005**, *70*, 9573–9583. (b) Freccero, M; Gandolfi, R.; Sarzi-Amadè, M.; Rastelli, A. *J. Org. Chem.* **2004**, *69*, 7479–7485. (c) Freccero, M.; Gandolfi, R.; Sarzi-Amadè, M.; Rastelli, A. *J. Org. Chem.* **2002**, *67*, 8519–8527. (d) Freccero, M.; Gandolfi, R.; Sarzi-Amadè, M.; Rastelli, A. *J. Org. Chem.* **2000**, *65*, 8948–8959. (e) Freccero, M.; Gandolif, R.; Sarzi-Amadè, M.; Rastelli, A. *J. Org. Chem.* **2000**, *65*, 2030–2042. (f) Freccero, M.; Gandolfi, R.; Sarzi-Amadè, M.; Rastelli, A. *J. Org. Chem.* **1999**, *64*, 3853–3860. (g) Freccero, M.; Gandolfi, R.; Sarzi-Amadè, M.; Rastelli, A. *Tetrahedron* **1998**, *54*, 12323–12336. (h) Freccero, M.; Gandolfi, R.; Sarzi-Amadè, M.; Rastelli, A. *Tetrahedron* **1998**, *54*, 6123–6134. (i) Markad, S. D.; Xia, X.; Snyder, N. L.; Surana, B.; Morton, M. D.; Hadad, C. M.; Peczuh, M. W. *J. Org. Chem.* **2008**, *73*, 6341–6354.

(7) Yang, D. *Acc. Chem. Res.* **2004**, *37*, 497–505.

(8) Wang, Z.-X.; Tu, Y.; Frohn, M.; Zhang, J.-R.; Shi, Y. *J. Am. Chem. Soc.* **1997**, *119*, 11224.

(9) Bartlett, P. D. *Rec. Chem. Prog.* **1957**, *18*, 111.

(10) (a) Tian, H.; She, Z.; Shu, L.; Yu, H.; Shi, Y. *J. Am. Chem. Soc.* **2000**, *122*, 11551–11552. (b) Tian, H.; She, Z.; Xu, J.; Shi, Y. *Org. Lett.* **2001**, *3*, 1929–1931. (c) Shi, Y. *Acc. Chem. Res.* **2004**, *37*, 488–496.

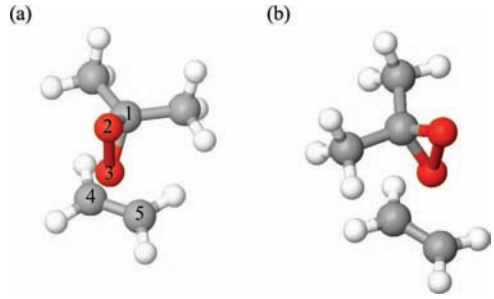


Figure 1. Spiro (a) and planar (b) transition states of dimethyldioxirane (DMDO) and ethylene.

Physical Model and Computational Methodology

The enantiomeric excess is given by

$$\mathrm{ee} = \frac{k_{\mathrm{favored}} - k_{\mathrm{disfavored}}}{k_{\mathrm{favored}} + k_{\mathrm{disfavored}}} \tag{1}$$

where k is the reaction rate constant leading to either the favored or the disfavored enantiomer. The rate constant can be expressed via transition state theory as

$$k = A(T)\, \mathrm{e}^{-\Delta G^{\ddagger}/RT} \tag{2}$$

where $A(T)$ is a pre-exponential factor, assumed to be equivalent for both enantiomeric pathways, $\Delta G^{\ddagger}$ is the activation Gibbs free energy, R is the gas constant, and T is temperature in Kelvin. Substitution of eq 2 into eq 1 then allows us to write

$$\mathrm{ee} = \frac{\exp(-\delta\Delta G^{\ddagger}/RT) - 1}{\exp(-\delta\Delta G^{\ddagger}/RT) + 1} \tag{3}$$

where $\delta\Delta G^{\ddagger}$ is given by

$$\delta\Delta G^{\ddagger} = \Delta G^{\ddagger}_{\mathrm{favored}} - \Delta G^{\ddagger}_{\mathrm{disvavored}} \tag{4}$$

As the starting materials leading to the two different enantiomers are identical, and therefore also have identical free energy, we can write the difference in Gibbs free activation energy, $\delta\Delta G^{\ddagger}$, just as the difference in transition state free energies, ΔG^{TS}:

$$\delta\Delta G^{\ddagger} = \Delta G^{\mathrm{TS}} = G^{\mathrm{TS}}_{\mathrm{favored}} - G^{\mathrm{TS}}_{\mathrm{disfavored}} \tag{5}$$

Because the Gibbs free energy, G, is simply related to the Helmholtz free energy, A, as $G = A + pV$, where pV is assumed to be identical for both enantiomeric pathways, we set ΔG^{TS} equal to ΔA^{TS}. The Helmholtz free energy of a transition state ensemble i was then computed from an estimate of the partition function, Z_i:

$$A_i^{\mathrm{TS}} = -RT \ln Z_i \tag{6}$$

Neglecting entropic effects partially, the partition function Z_i was approximated as a Boltzmann-weighted average over unique transition states (shown schematically in Figure 2) leading to enantiomer i:

$$Z_i \approx \sum_j \exp\left(-\frac{E_j^{\mathrm{TS}}}{RT}\right) \tag{7}$$

E is the quantum mechanical transition state energy, and the index j loops over unique transition states leading to enantiomer i. Rearrangement of eq 3 gives an expression to compute ΔG^{TS} from the ee value and *vice versa*:

$$\Delta G^{\mathrm{TS}} = RT \ln\left(\frac{1 + \mathrm{ee}}{1 - \mathrm{ee}}\right) \tag{8}$$

For systems involving dioxirane catalysts and olefins with no rotational symmetry, eight geometrically distinct transition states were considered, as shown in Figure 2. For systems involving the C_{2h} symmetrical *trans*-stilbene (olefin **c**), the number of unique transition states is reduced by a factor of 2. The same is true for systems involving the C_2 symmetrical catalyst **10**.

To locate transition states in their lowest energy conformations, initial guesses for all transition states were first obtained using the spiro UB3LYP/6-31G* epoxidation transition state of DMDO and ethylene as a scaffold. (Figure 1a). While Shi invokes a planar transition state in some cases,[10] the planar transition state scaffold was not used, as Houk et al. have shown that the planar transition state for DMDO and ethylene is 7.4 kcal/mol higher in energy than the corresponding spiro transition state at the B3LYP/6-31G* level and represents only a secondary saddle point.[4a] We attempted to address the conformational flexibility of the systems by performing conformational searches of the transition state guesses using MacroModel[11] and the OPLS 2001 force field[12] while keeping the Cartesian coordinates of the reactive atoms of the transition state guesses frozen (atom numbers 1−5 in Figure 1a). For some systems involving catalysts **14**−**19** (Chart 1) with relatively flexible ring systems, conformational searching was also performed with the spiro carbon atoms of the dioxirane (C-1 in Figure 1a) unfrozen, therefore allowing rotational adjustment of the transition state geometry from strictly spiro to hybrid spiro-planar.[13] Finally, where we suspected that the above did not yield transition states in their lowest energy conformations, we further subjected the structures to manual conformational searching.

We further refined our initial transition state guesses by performing minimum energy structure optimizations at the B3LYP/6-31G* level while keeping three distances (C1−O3, C5−O3, and C4−O3 in Figure 1a) between the reactive atoms fixed. The quantum mechanically refined transition state guesses were then used as starting points for all subsequent transition state optimizations.

It has already been shown that dioxirane-mediated epoxidations are well described (at least qualitatively) by B3LYP.[4,5] Houk et al.[4a] and Bach et al.[3h] have suggested that the transition states for epoxidation by dioxiranes lie in shallow potential energy wells and that asymmetrical substitutions on the symmetrical DMDO and ethylene transition state scaffold lead to more asynchronous transition states. Furthermore, spin-unrestricted B3LYP (UB3LYP) leads to more asynchronous transition states than spin-restricted B3LYP.[4a] Therefore, we employed unrestricted density functional theory within Jaguar 7.0 for all quantum mechanical computations to capture the asynchronicity and flexibility of the transition states. Transition states were first located in a vacuum and then in the solution phase using the PBF solvation model as implemented in Jaguar 7.0. Solvation effects for the vacuum transition states were approximated from single-point calculations in the continuum solvent.[14] For CH_3CN/H_2O and dioxane/H_2O mixtures, dielectric constants were taken as those reported in experimental measurements.[15] Because no experimental dielectric constants were reported for the other solvent mixtures, DMM/CH_3CN/H_2O and DMM/DME/H_2O, the dielectric constant was approximated as the mole fraction

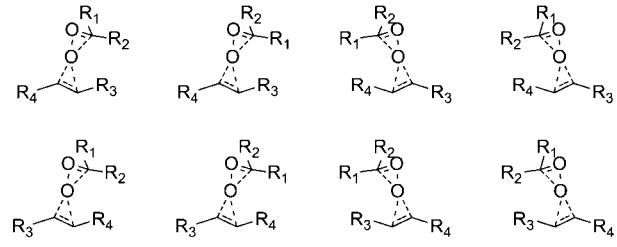


Figure 2. All unique transition states considered for systems where the catalysts and olefins show no rotational symmetry.

weighted average over all dielectric constants of the pure solvents. The dielectric constants ultimately employed were 55, 47, 20, 69, and 73 for 3:2 $CH_3CN:H_2O$ (v:v), 3:1 $CH_3CN:H_2O$, 2:1 dioxane/H_2O, 3:3:6.4 DMM:$CH_3CN:H_2O$, and 26:1:2.8 H_2O:DMM:DME solvent systems, respectively. The solvent probe radius was set to 1.40 Å for all calculations, as this probe radius was shown to reproduce solvation energies accurately for pure water.[16] The first shell correction factor term and the solute cavity energy term available in Jaguar 7.0 were included for all solvation calculations.[17] Single-point calculations at the B3LYP/6-311+G**(solution)//B3LYP/6-31G*(solution) level were also performed in order to check the basis set dependence of the results. Finally, all B3LYP transition states for a representative subset of the data set were fully reoptimized at the M06-2X[18]/6-31G*(solution) level employing Jaguar version 7.6, with the goal of obtaining a more accurate treatment of dispersion interactions, for which B3LYP is known to perform poorly.[19]

The nature of all vacuum B3LYP transition states was verified with frequency calculations, yielding only one large imaginary frequency. Harmonic zero-point energy corrections obtained from the frequency calculations of the vacuum B3LYP transition states were applied to all transition state energies. Because second derivatives within the solvation model can only be calculated via an impractically expensive numerical procedure, vibrational frequencies of the transition states in solution were not computed. Harmonic thermochemical and entropic corrections other than the zero-point energy, for which low-frequency vibrational modes contribute the most, were not included because the harmonic approximation is highly inaccurate for low-frequency vibrational modes.

Results and Discussion

Figure 3 shows the relationship between the computed and experimental enantioselectivity. Like others,[1] we found that ΔG^{TS} is systematically overestimated and scaled all ΔG^{TS} by a scaling factor of 2/3 accordingly before using ΔG^{TS} to calculate the enantiomeric excesses. Inspection of Figure 3 shows that our method is able to capture which enantiomer a system will prefer, except for a few cases where low enantioselectivity is predicted. Furthermore, it is clear that our method consistently and accurately captures whether a system will have low (0−50%), medium (50−85%), or high (>85%) enantioselectivity for the majority of the systems studied. Specifically, most systems with high experimental enantioselectivity also have high computed enantioselectivity. The same is true for both medium and low enantioselectivities, suggesting that our method can be used to model in a quantitative sense the enantioselectivity of other dioxirane-catalyzed epoxidations and perhaps even other enantioselective reactions as well.

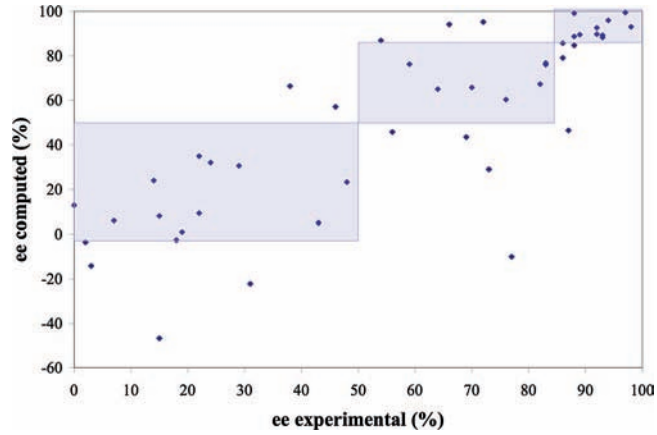


Figure 3. Experimental versus computed % ee for solution phase transition states at the B3LYP/6-31G*(solution) level. Shaded areas represent regions where computation and experiment fall within the same category of percent enantiomeric excess, i.e., low (0−50), medium (50−85), and high (85−100). Reactions that do not fall into the correct region are labeled in an identical type of graph in the Supporting Information. A linear least-squares fit through all the points gives a coefficient of determination $R^2 = 0.66$.

The experimental and computed free energy differences, ΔG^{TS}, between the transition state ensembles leading to favored and disfavored enantiomers are plotted in Figure 4. (Detailed results are listed in Table 1, while similar plots and tables for the results obtained at the B3LYP/6-311+G**(solution)//B3LYP/6-31G*(solution) and the M06-2X/6-31G*(solution) levels are available in the Supporting Information.) Five outliers with deviations from experiment larger than 1.5 kcal/mol are present at the B3LYP/6-31G*(solution) level. As discussed in detail later, all of these outliers share the unique structural motif of π-stacking between an oxygen or a chlorine atom and a benzene ring. Of course, those outliers significantly distort the linear least-squares fit shown in Figure 4b. Without the five outliers, the slope of the linear least-squares fit becomes significantly closer to unity, with a value of 1.33, while the correlation coefficient, R^2, increases to a respectable value of 0.78.

Interestingly, a comparison of parts a and b of Figure 4 shows a significant improvement in correlation with experimental results when the transition states are optimized in the solution phase, rather than simply estimating solvation energies with single-point calculations, with correlation coefficients of 0.46 and 0.64, respectively. This improvement is unsurprising, given that all systems treated were aqueous, suggesting that their transition state geometries might change significantly by reoptimizing the vacuum transition states within the solvation model. This strategy differs from those reported in a number of other publications, in which gas phase optimized transition states were employed exclusively under the assumption that the transition state geometries would not vary drastically from vacuum to solution phase.[20] Our results show the dubious nature of this common assumption. Figure 5 shows the lowest-energy favored and disfavored transition states in vacuum and solution phase for cases that show the most improvement following reoptimization in solution phase. Notably, the two forming bonds (O3−C4 and O3−C5 in Figure 1) are appreciably longer when the transition states are located in solution rather than vacuum

(11) *MacroModel 6.0*; Schrödinger, Inc.: Portland, OR, 1998.
(12) Jorgensen, W. L.; Maxwell, D. S.; Tirado-Rives, J. *J. Am. Chem. Soc.* **1996**, *118*, 11225–11236.
(13) This proved to be necessary as we observed many transition states that were neither completely planar nor completely spiro, but rather a hybrid. This phenomenon has also been reported by Singleton et al.[5b]
(14) Throughout text we will use B3LYP/6-31G*(solvation)//B3LYP/6-31G*(vacuum) to indicate single-point calculations within the solvation model on the transition states optimized in a vacuum.
(15) (a) Venables, D. S.; Schmuttenmaer, C. A. *J. Chem. Phys.* **1998**, *108*, 4935–4944. (b) Ahn-Ercan, G.; Krienke, H.; Schmeer, G. *J. Mol. Liq.* **2006**, *129*, 75–79.
(16) *Jaguar 7.0 User Manual*; Schrödinger, Inc.: Portland, OR, 2007.
(17) Keywords icavity=2 and isurf=1.
(18) Zhao, Y.; Truhlar, D. G. *Acc. Chem. Res.* **2008**, *41*, 157–167.
(19) (a) Johnson, E. R.; Wolkow, R. A.; DiLabio, G. A. *Chem. Phys. Lett.* **2004**, *394*, 334. (b) Tsuzuki, S.; Üthl, H. P. *J. Chem. Phys.* **2001**, *114*, 3949. (c) Ye, X.; Li, Z. H.; Wang, W.; Fan, K.; Xu, W.; Hua, Z. *Chem. Phys. Lett.* **2004**, *397*, 56. (d) Godfrey-Kittle, A.; Cafiero, M. *Int. J. Quantum Chem.* **2006**, *106*, 56.
(20) (a) Silva, M. A.; Bellenie, R.; Goodman, J. M. *Org. Lett.* **2004**, *6*, 2559. (b) Bach, R. D.; Glukhovtsev, M. N.; Gonzalez, C. *J. Am. Chem. Soc.* **1998**, *120*, 9902. (c) Cheong, P. H. Y.; Yun, H.; Danishefsky, S. J.; Houk, K. N. *Org. Lett.* **2006**, *8*, 1513.

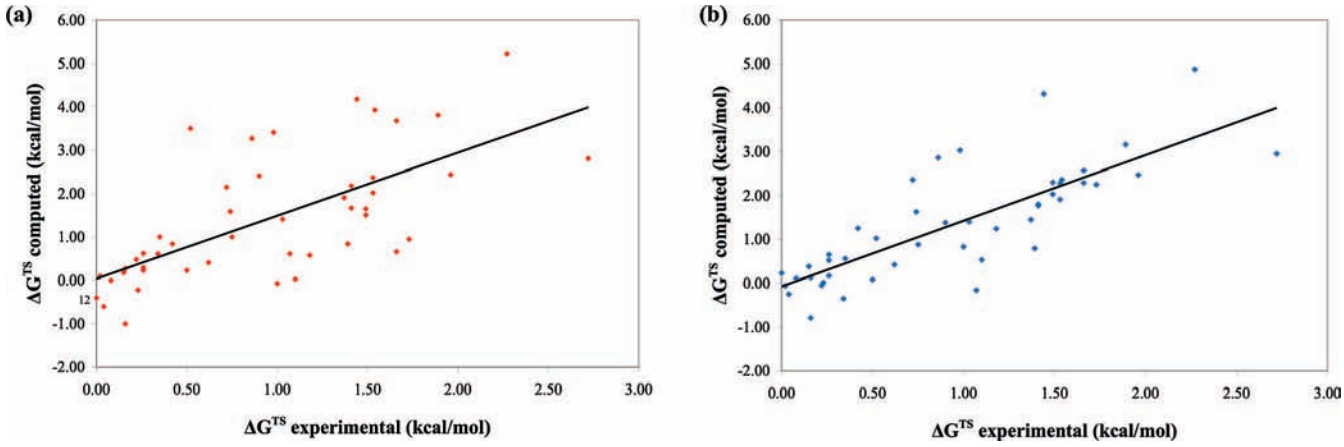


Figure 4. Experimental versus computed enantiomeric transition state energy differences, ΔG^{TS}, for all reactions in the test set. (a) B3LYP/6-31G*(solution)// B3LYP/6-31G*(vacuum) level. Solid line shows linear least-squares fit of $y = 1.45x + 0.04$, with coefficient of determination $R^2 = 0.46$. (b) B3LYP/6-31G*(solution) level ($y = 1.50x - 0.07$, $R^2 = 0.64$). Identical figures with all data points labeled with their corresponding reaction numbers are shown in the Supporting Information.

phase. In order to quantify this effect, we defined average forming bond lengths as a Boltzmann-weighted average of the corresponding bond lengths of all transition structures. The computed average bond lengths for each reaction are listed in the Supporting Information (Table S2). Each asynchronous transition state is characterized by two forming average bond lengths, one shorter then the other. In solution phase these bond lengths are 2.40 ± 0.02 Å (mean ± standard error; sample size $n = 46$) and 2.15 ± 0.01 Å ($n = 46$), respectively. These are longer than the corresponding averaged bonds lengths of the vacuum transition states, which are 2.30 ± 0.01 Å ($n = 46$) and 1.97 ± 0.01 Å ($n = 46$), respectively.

While locating transition states in the solution phase leads to significantly improved results, five outliers persist. Specifically, cases 41, 29, 33, 32, and 9 (referring to the entry numbers in Table 1, listed in decreasing order of error) are not treated well by our model in either vacuum or solution phase and show deviations from experimental ΔG^{TS} larger than 1.5 kcal/mol at the B3LYP/6-31G*(solution) level. All outliers share one common characteristic: the relevant disfavored transition states have either an oxygen or a chlorine atom stacked over the face of a benzene ring (as illustrated in Figure 7). This structural motif seems to be present only for cases with large discrepancies between computation and experiment. For each ensemble of favored and disfavored B3LYP/6-31G* solution phase transition states, the Boltzmann-weighted average of the number of oxygen, fluorine, or chlorine atoms present in a cone above the plane of each benzene ring was computed as shown in Figure 6a. The differences between these average values of disfavored and favored transition state ensembles were then plotted against the deviation of the computed ΔG^{TS} from experimental values (Figure 6b).

Based on the well-known deficiencies of B3LYP in treating dispersion and especially π-stacking interactions,[19] it is not surprising to see that π-stacking interactions with oxygen and chlorine pose an obstacle for our computational model. However, π-stacking interactions with the smaller and less polarizable fluorine seem to be treated much more accurately by our model. Therefore, because B3LYP generally does not take into account attractive π-stacking interactions, an inflated computed transition state free energy gap could be the result of an artificial energy increase of the disfavored transition state ensemble due to repulsive B3LYP errors for π-stacking interactions. Where π-stacking interactions with oxygen or chlorine are absent from a reaction, dispersion-related B3LYP errors are presumably minimized due to fortuitous error cancelation. However, since B3LYP errors for π-stacking interactions are often several kcal/mol higher than the corresponding errors for other types of dispersion interactions,[21,22] differential error likely accumulates where strong π-stacking interactions are present.

In order to test that incorrect treatment of dispersion interactions by B3LYP is a major source of error in our data set, all transition states of a representative subset of the data set were reoptimized at the M06-2X/6-31G*(solution) level.[23] Truhlar's M06-2X functional performs well for dispersion-dominated complexes as well as for activation energies.[18] However, the preliminary results obtained with M06-2X on our data set are inferior to the B3LYP results. (For a comparison of the statistical performances of all methods employed in this study, see Table 2.) The mean unsigned error, as well as the correlation between experimental and computed ΔG^{TS}, is clearly worse for M06-2X than for B3LYP. It is noteworthy, however, that three of the five B3LYP/6-31G*(solution) outliers improve significantly, namely systems 9, 32, and 33. For those systems, the errors with respect to experiment are reduced by 1.4, 1.3, and 0.8 kcal/mol, respectively, with the M06-2X optimizations. The B3LYP outlier 29 becomes significantly worse (with an error increase by 1.6 kcal/mol) when optimized with M06-2X, while for the fifth outlier, 41, M06-2X results are not available due to convergence problems for some of the transition states.

The M06-2X results are somewhat surprising in view of the strong performance demonstrated by this functional for data on a variety of training sets and for a wide range of properties. There are several possible explanations for the results, which we plan to pursue in future work:

(21) This is illustrated by comparing the B3LYP/6-31G* errors for the stacked (C_{2h}) and the T-shaped (C_{2v}) benzene dimers. Single-point B3LYP/6-31G* calculations on *ab initio* geometries reported by Jurečka[22] give binding energies of 2.58 and 0.18 kcal/mol for the stacked and the T-shaped benzene dimers, respectively, while best estimates for these interaction energies are −2.73 and −2.74 kcal/mol. Therefore, B3LYP seems to give the largest repulsive errors for systems showing strong π-stacking interactions.

(22) Jurečka, P.; Šponer, J.; Černyí, J.; Hobza, P. *Phys. Chem. Chem. Phys.* **2006**, *8*, 1985–1993.

(23) Zhao, Y.; Truhlar, D. *Theor. Chem. Acc.* **2008**, *120*, 215–241.

Table 1. B3LYP/6-31G*(solution)//B3LYP/6-31G*(vacuum) and B3LYP/6-31G*(solution) Results

			ΔG^{TS} (kcal/mol) *(% ee)*							
entry	catalyst	olefin	exp[c]		vac[d]		soln[e]		ref[a]	solv[b]
1	**1**	**e**	0.22	*(18)*	0.48	*(26)*	−0.05	*(−3)*	I	A
2	**1**	**b**	0.75	*(56)*	1.00	*(51)*	0.88	*(46)*	I	A
3	**1**	**c**	1.18	*(76)*	0.58	*(32)*	1.24	*(60)*	I	A
4	**1**	**f**	1.10	*(73)*	0.03	*(2)*	0.53	*(29)*	I	A
5	**1**	**d**	1.00	*(69)*	−0.07	*(−4)*	0.83	*(44)*	I	A
6	**1**	**g**	0.26	*(22)*	0.29	*(16)*	0.65	*(35)*	I	A
7	**1**	**a**	0.35	*(29)*	1.00	*(51)*	0.56	*(31)*	I	A
8	**1**	**h**	1.41	*(83)*	2.18	*(84)*	1.80	*(77)*	I	A
9	**2**	**c**	0.72	*(54)*	2.15	*(84)*	2.36	*(87)*	I	A
10	**3**	**c**	1.53	*(86)*	2.37	*(87)*	2.28	*(86)*	I	A
11	**8**	**b**	0.26	*(22)*	0.23	*(13)*	0.17	*(9)*	II	B
12	**9**	b	0.00	*(0)*	−0.40	*(−22)*	0.23	*(13)*	II	B
13	**7**	**b**	1.03	*(70)*	1.41	*(66)*	1.40	*(66)*	II	B
14	**7**	**c**	1.53	*(86)*	2.02	*(81)*	1.91	*(79)*	II	B
15	**10**	**b**	1.49	*(88)*	1.65	*(77)*	2.30	*(89)*	III	A
16	**10**	**c**	1.89	*(94)*	3.81	*(98)*	3.16	*(96.0)*	III	C
17	**10**	**d**	0.74	*(59)*	1.59	*(75)*	1.63	*(76)*	III	A
18	**10**	**i**	0.50	*(43)*	0.23	*(14)*	0.08	*(5)*	III	A
19	**4**	**c**	1.41	*(83)*	1.67	*(74)*	1.77	*(76)*	IV	A
20	**6**	**c**	0.90	*(64)*	2.41	*(88)*	1.38	*(65)*	IV	A
21	**6**	**a**	0.02	*(2)*	0.10	*(6)*	−0.06	*(−4)*	IV	A
22	**5**	**a**	0.62	*(48)*	0.41	*(23)*	0.42	*(23)*	V	A
23	**5**	**c**	1.96	*(93)*	2.44	*(88)*	2.47	*(88)*	V	A
24	**5**	**d**	1.37	*(82)*	1.91	*(79)*	1.45	*(67)*	V	A
25	**5**	**h**	2.72	*(98)*	2.81	*(92)*	2.95	*(93)*	V	A
26	**5**	**j**	0.04	*(3)*	−0.60	*(−32)*	−0.25	*(−14)*	V	A
27	**11**	**g**	0.08	*(7)*	0.00	*(0)*	0.11	*(6)*	VI	A
28	**11**	**a**	0.23	*(19)*	−0.22	*(−12)*	0.02	*(1)*	VI	A
29	**14**	**c**	2.27	*(97)*	5.22	*(99.7)*	4.88	*(99.5)*	VII	D
30	**15**	**c**	1.49	*(88)*	1.51	*(73)*	2.03	*(85)*	VII	D
31	**16**	**c**	1.54	*(89)*	3.93	*(98.4)*	2.36	*(90)*	VII	D
32	**18**	**c**	0.86	*(66)*	3.27	*(96.5)*	2.86	*(94)*	VII	D
33	**19**	**c**	0.98	*(72)*	3.41	*(97.0)*	3.03	*(95)*	VII	D
34	**14**	**a**	0.16	*(15)*	−1.00	*(−56)*	−0.79	*(−47)*	VII	D
35	**15**	**a**	0.16	*(15)*	0.27	*(17)*	0.13	*(8)*	VII	D
36	**16**	**a**	0.34	*(31)*	0.61	*(37)*	−0.35	*(−22)*	VII	D
37	**17**	**a**	0.26	*(24)*	0.62	*(38)*	0.52	*(32)*	VII	D
38	**18**	**a**	0.15	*(14)*	0.18	*(12)*	0.38	*(24)*	VII	D
39	**14**	**b**	1.66	*(92)*	3.68	*(98.2)*	2.57	*(93)*	VII	D
40	**15**	**b**	1.39	*(87)*	0.84	*(49)*	0.79	*(47)*	VII	D
41	**16**	**b**	1.44	*(88)*	4.18	*(99.0)*	4.32	*(99.2)*	VII	D
42	**17**	**b**	1.07	*(77)*	0.61	*(37)*	−0.16	*(−10)*	VII	D
43	**18**	**b**	0.52	*(46)*	3.50	*(97.7)*	1.02	*(57)*	VII	D
44	**19**	**b**	0.42	*(38)*	0.84	*(49)*	1.25	*(66)*	VII	D
45	**13**	**k**	1.66	*(92)*	0.66	*(40)*	2.29	*(90)*	VIII	E
46	**12**	**l**	1.73	*(93)*	0.95	*(54)*	2.25	*(89)*	VIII	E

[a] Experimental references: (I) Armstrong, A.; Ahmed, G.; Dominguez-Fernandez, B.; Hayter, B. R.; Wailes, J. S. *J. Org. Chem.* **2002**, *67*, 8610. (II) Solladié-Cavallo, A.; Jierry, L.; Klein, A.; Schmitt, M.; Welter, R. *Tetrahedron: Asymmetry* **2004**, *15*, 3891. (III) Denmark, S. E.; Matsuhashi, H. *J. Org. Chem.* **2002**, *67*, 3479. (IV) Armstrong, A.; Dominguez-Fernandez, B.; Tsuchiya, T. *Tetrahedron* **2006**, *62*, 6614. (V) Armstrong, A.; Moss, W. O.; Reeves, J. R. *Tetrahedron: Asymmetry* **2001**, *12*, 2779. (VI) Armstrong, A.; Tsuchiya, T. *Tetrahedron* **2005**, *62*, 257. (VII) Wang, Z.; Shi, Y. *J. Org. Chem.* **2001**, *66*, 521. (VIII) Burke, C. P.; Shi, Y. *Angew. Chem., Int. Ed.* **2006**, *45*, 4475. [b] Solvent systems: (A) 3:2 CH_3CN:H_2O (v:v), dielectric constant employed in the continuum solvation model, $\varepsilon = 55$; (B) 2:1 dioxane:H_2O (v:v), $\varepsilon = 20$; (C) 3:1 CH_3CN:H_2O (v:v), $\varepsilon = 47$; (D) 3:3:6.4 DMM:CH_3CN:H_2O (v:v:v), $\varepsilon = 69$; (E) 26:1:2.8 H_2O:DMM:DME (v:v:v), $\varepsilon = 73$. [c] Experimental ΔG^{TS}. [d] Results at the B3LYP/6-31G*(solution)//B3LYP/6-31G*(vacuum) level. The ee values were obtained from scaled ΔG^{TS} (scaling factor = 2/3). [e] Results at the B3LYP/6-31G*(solution) level. The ee values were obtained from scaled ΔG^{TS} (scaling factor = 2/3).

(1) Our solvation model has not been reoptimized for M06-2X; we simply used the model previously optimized for B3LYP. Hence, if the charge distributions obtained with M06-02X are sufficiently different (and that might not be an enormous quantitative shift, in view of the small energy differences we are talking about), this could lead to degradation of the relative energetics.

(2) We spent significantly less time creating the optimized transition states for M06-2X than we did when working on the data set with B3LYP. There may be, for example, alternative minima with the former functional that we did not explore. Again, this could lead problems for specific compounds where one or both of the transition state structures are not in the lowest free energy state.

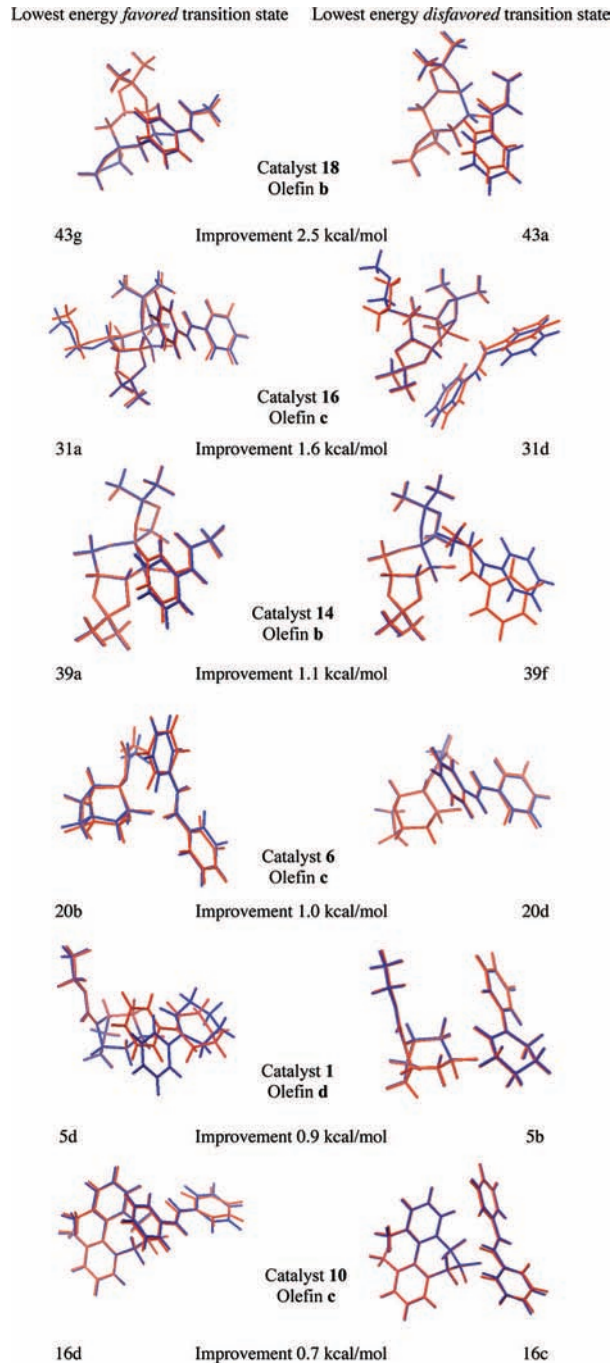


Figure 5. Lowest-energy B3LYP/6-31G*(vacuum) transition state structures (red) superimposed on the B3LYP/6-31G*(solution) transition state structures (blue) for systems that show the most significant improvements following reoptimization of the transition states in solution phase. Reductions in absolute deviations from experimental ΔG^{TS} are also shown.

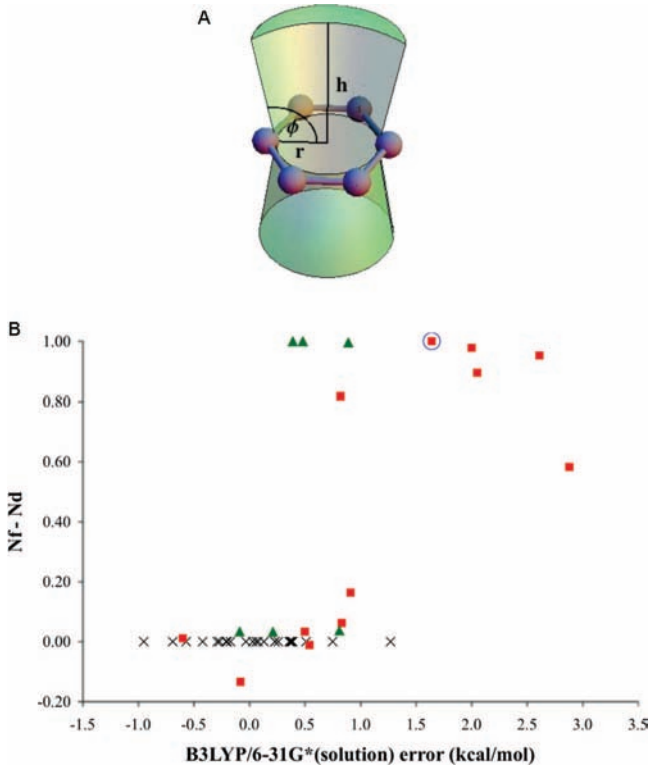


Figure 6. (a) Schematic illustration of the cone volume above benzene rings searched for oxygen, fluorine, and chlorine atoms. Benzene carbon atoms are shown in gray. $h = 4.0$ Å, $r = 1.5$ Å, and $\phi = 100°$. (b). Difference between disfavored and favored transition state ensembles in the total number of oxygen, fluorine, and chlorine atoms (Boltzmann-weighted average) present within the cone shown in panel a above each benzene ring, plotted against deviation from experiment. Black crosses, no atom in cone; red squares, oxygen in cone; green triangles, fluorine in cone; and blue circle, chlorine in cone.

(3) While M06-2X has displayed encouraging performance for a number of training sets, these data sets are relatively small and do not cover a broad range of chemical problems. M06-2X also contains a large number of fitting parameters (on the order of 35) as compared to the three parameters in B3LYP. Hence, there has to be some concern about overfitting (which would yield superior results for the training set as compared to what one might expect in an application not covered by those data sets), and what we have observed may be one example of this. Much more work has to be done to rigorously confirm (or deny) this interpretation; however, we feel that it is important to present the data and conclusions, not least so that other groups can access the data and check our results with their own calculations. To this end, we have supplied all of the data needed for such investigations in our Supporting Information.

Single-point calculations of the solution phase transition states at the B3LYP/6-311+G**(solution)//B3LYP/6-31G*(solution) level were also performed to probe basis set dependence. Table 2 shows that the correlation coefficient becomes notably worse compared to the results obtained with the smaller basis set. One possible reason might simply be random errors arising from the fact that the geometries are not actually optimized within the larger basis set. Singleton has shown that relative energies at the B3LYP/6-311+G**(vacuum)//B3LYP/6-31G*(vacuum) level vary by only about 0.2 kcal/mol from the energies at the B3LYP/6-311+G**(vacuum)//B3LYP/6-31+G**(vacuum) level for two Shi epoxidations.[5b] However, changes in geometry due to changes in basis set might be larger for transition states

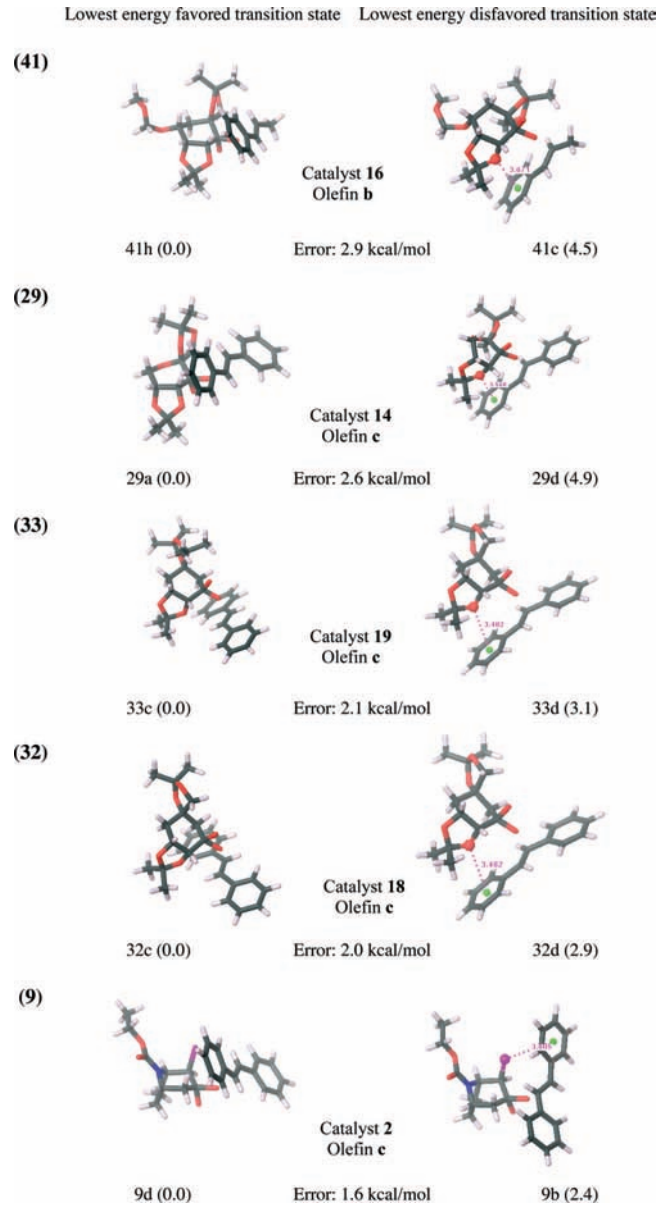


Figure 7. Lowest-energy favored and disfavored transition states at the B3LYP/6-31G* (solution) level for all cases where the error at this level is larger than 1.5 kcal/mol. Deviations (in kcal/mol) from the experimental transition state free energy differences, ΔG^{TS}, are given for each reaction. Relative energies of the transition states (in kcal/mol) are shown in parentheses. π-Stacking interaction distances between benzene middle points (green) and oxygen or chlorine atoms (in Å) are shown. Atom colors: black, carbon; red, oxygen; white, hydrogen; blue, nitrogen; and pink, chlorine.

optimized in solution phase since the solvation energies also depend on the basis sets used. In addition, since we seek to discriminate free energy differences within a range of only about 2.5 kcal/mol, even small errors in relative energies, which are likely to be negligible for other problems, lead to a dramatic degradation of the correlation with experiment.

Notably, four out of five outliers of the data set are Shi epoxidations, and it is therefore not surprising that the mean unsigned error (MUE) for the Shi epoxidations (Table 1, entries 29−46) is significantly higher than for the non-Shi epoxidations (entries 1−28), with values of 1.02 and 0.41 kcal/mol at the B3LYP/6-31G*(solvent) level, respectively. Inaccurate treatment of dispersion interactions by B3LYP is one likely reason for this discrepancy. However, additional factors could further

explain the computational model's difficulty in modeling the Shi epoxidation in particular.

First, dioxirane-catalyzed epoxidations have been shown to be particularly sensitive to solvent selection, presumably because the transition state exhibits enhanced hydrogen bonding.[4b] In cases where the olefin contains hydrophobic groups (such as phenyl substituents), hydrophobic effects likely also influence the enantioselectivity. However, continuum solvation models might not accurately capture such solvation effects. Notably, the sugar-derived dioxirane catalysts employed by Shi (catalysts **12**−**19**) contain many potential hydrogen bond acceptors, and the corresponding reactions are also performed in highly polar solvent systems. Therefore, errors due to approximating the solvent with a dielectric continuum are likely to be largest for the Shi epoxidation reactions. A plot of MUE against the number of hydrogen bond acceptors can be found in the Supporting Information (Figure S3).

Second, as noted by Singleton and Houk,[4,5b] the Shi epoxidation can have many conformationally distinct low-energy transition states. To quantify the degree of conformational flexibility for each reaction, each catalyst and olefin was subjected to conformational searching within MacroModel using the OPLS 2001 force field employing 20 000 steps. The sum of unique conformations of catalyst and olefin conformations within 5 kcal/mol was then taken as a measure of the conformational flexibility of a system. Accordingly, we found that the Shi epoxidation catalysts (entries 29−46 in Table 1) show a distinctly higher degree of conformational flexibility compared to the rest of the data set, with values of 52 ± 11 (mean ± s.e.; $n = 18$) and 17 ± 11 ($n = 28$), respectively. Even though we tried searching conformational space as completely as possible, according to our protocol described in the methodology section, we cannot exclude the possibility that for some flexible systems there might exist low-energy conformations of transition states not considered in this study. Therefore, part of the observed errors for the more flexible Shi epoxidations could also result from incomplete conformational searching.

Much debate within the literature has centered around whether the dioxirane-mediated epoxidation proceeds through a spiro or planar transition state, as shown in Figure 1. Computational studies at various levels of theory have shown the preference of spiro over planar transition states for small model systems.[3,4,24] However, Shi has proposed that some reactions with sterically hindered transition states as well as reactions involving *cis*-olefins proceed through planar transition states.[10] In agreement with the results obtained by Singleton,[5b] the results of our work indicate that all relevant low-energy transition states are either spiro or hybrid spiro-planar; no relevant transition states with a completely planar geometry were found. In order to classify the degree of rotation in transition states for each reaction, Boltzmann-weighted averages of the dihedral angles, θ, as shown in Figure 8, were computed.

For each ensemble of favored and disfavored transition states, the averaged dihedral angles of the solution phase transition states are listed in the Supporting Information (Table S3). Interestingly, dihedral angles for the *disfavored* transition states deviate more from 90° (which corresponds to a perfect spiro transition state) than do the dihedral angles for the *favored* transition states, with average values for the whole data set of 117 ± 2° (mean ± s.e.; $n = 46$) and 100 ± 1° ($n = 46$), respectively. This suggests that the disfavored transition states are often distorted from spiro in order to reduce steric clashes. However, while Shi proposed that this distortion leads to planar transition states, we show that less distorted hybrid spiro-planar transition states result instead. In addition, the degree of rotation in the transition states seems to be at least partly responsible for the increased energy of the disfavored transition states as compared to the favored ones.

Other groups have illustrated the importance of asynchronous transition structures.[3−5] Unsurprisingly, we find that the transition state asynchronicity is generally larger for asymmetrically substituted than for symmetrically substituted olefins. In order to quantify these effects, we defined asynchronicity as the absolute difference between the two forming bond lengths (O3−C4 and O3−C5 distances in Figure 1). Representative bond lengths for each reaction were computed as the Boltzmann-weighted averages of the corresponding shorter and longer distances of all transition states. Asynchronicities for each reaction were then computed as absolute differences between the Boltzman-weighted distance averages. Table 3 shows mean asynchronicities for all olefins present in the data set (ordered according to increasing asynchronicity). The electron deficient 2-cyclohexenone (olefin **j**) forms transition states with lowest asynchronicity. The symmetrically substituted *trans*-stilbene (olefin **c**) forms transition states with clearly lower asynchronicity than the other asymmetrically substituted olefins. Con-

Table 2. Performance of the Methods Employed in the Current Study

method	MUE[a]	max error[b]	R^2 [c]
B3LYP/6-31G*(solv)//B3LYP/6-31G*(vac)	0.85	3.0	0.46
B3LYP/6-31G*(solv)	0.65	2.9	0.64
B3LYP/6-311+G**(solv)//B3LYP/6-31G*(solv)	0.73	3.1	0.36
M06-2X/6-31G*(solv)[d]	1.76	5.2	0.12

[a] Mean unsigned error (kcal/mol) from experimental ΔG^{TS}. [b] Maximum absolute deviation (kcal/mol) from experimental ΔG^{TS}. [c] Correlation coefficient for a linear least-squares fit to a plot of computed versus experimental ΔG^{TS}. [d] Due to problems in converging some of the transition state geometries, systems 27, 28, 31, 36, 37, 41, 45, and 46 are not included in the M06-2X data set.

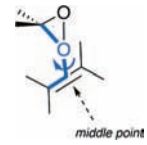


Figure 8. Schematic illustration of the dihedral angle θ (shown in blue).

Table 3. Asynchronicities for All Olefins of the Data Set

olefin	asynchronicity (Å) (mean ± s.e.[a])	n[b]
j	0.07	1
c	0.11 ± 0.02	13
f	0.19	1
b	0.25 ± 0.01	11
e	0.26	1
d	0.30 ± 0.02	3
i	0.33	1
l	0.34	1
h	0.37 ± 0.06	2
a	0.38 ± 0.01	9
g	0.41 ± 0.01	2
k	0.45	1

[a] Standard error. [b] Sample size.

(24) Crehuet, R.; Anglada, J. M.; Cremer, D.; Bofill, J. M. *J. Phys. Chem. A* **2002**, *106*, 3917–3929.

sistent with previous findings,[5b,25] a dioxirane-catalyzed epoxidation is an electrophilic process; i.e., we find a Boltzmann-weighted partial charge[26] on the olefin part of the transition states with an average value of 0.38 ± 0.01 (mean ± s.e.; n = 46). Considering the electrophilic nature of the epoxidations, it is not surprising that we find the longer of the two forming bonds always to be the one between the dioxirane oxygen and the carbon α to conjugated substituents such as phenyl or alkene.[27] Olefin **h** displays the longer forming bond, as expected, always between the dioxirane oxygen and the carbon α to the *two* phenyl substituents. Olefin **f**, containing two symmetrically arranged phenyl substituents, and olefin **j**, with very low asynchronicity, both show no clear trends in which of the two forming bond lengths is shorter in the transition states.

Conclusions

In conclusion, we present the first fully quantum mechanical study of quantitative computational modeling of enantioselectivity for a large data set. Our results indicate that optimization of transition states within a solvation model is critical to the accurate and robust prediction of experimental results. We find that the two forming bonds (O3−C4 and O3−C5 in Figure 1) are appreciably longer when the transition states are located in solution rather than in vacuum phase. We generally located distorted spiro transition states (Figure 1) and showed that the degree of rotation of the transition states (measured by the dihedral angle θ shown in Figure 7) seems to play an important role in determining the enantioselectivity of the reaction. Specifically, we find that the disfavored transition states show, on average, a stronger distortion from spiro than the favored ones. Finally, we show that B3LYP's problems in treating dispersion and especially π-stacking interactions are likely to be part of the reason for the few observed large discrepancies between experiment and computation. Our results illuminate the possibility of employing quantum mechanical predictions to facilitate the discovery of new reactions and catalysts in an efficient and promising way. Future work will continue to explore ways of increasing the accuracy of our first-generation model. Such approaches, for example, include applying semiempirical dispersion corrections to density functional theory. Because construction of the manual transition state guesses involves a significant amount of work, we are currently automating our protocol in order to simplify large-scale virtual screening of chiral catalysts.

Acknowledgment. R.B. thanks the National Science Foundation (CHE-06-40407) and R.A.F. thanks the Department of Energy (90ER14162) and the National Institutes of Health (GM 40526) for financial support of this research.

Supporting Information Available: Coordinates, energies, and imaginary frequencies of transition states. Results at the M06-2X/6-31G*(solvation) and B3LYP/6-311+G**(solvation)//B3LYP/6-31G*(solvation) level as well as additional plots and tables mentioned in the main text, including pictures of all relevant transition states not shown in the main text. This material is available free of charge via the Internet at http://pubs.acs.org.

JA806951R

(25) Düfert, A.; Werz, D. B. *J. Org. Chem.* **2008**, *73*, 5514–5519.
(26) Computed as the sum over all Mulliken charges of the olefin atoms at the B3LYP/6-31G*(solution) level.
(27) The only exceptions to this rule are a few high-energy transition states, i.e., 41a, 41e, and 42b (solution phase structures) as well as 39f, 40a, 41a, 41e, and 42b (vacuum structures).