Rationalizing the Stereoselectivity of Osmium Tetroxide Asymmetric Dihydroxylations with Transition State Modeling Using Quantum Mechanics-Guided Molecular Mechanics

Per-Ola Norrby,*,† Torben Rasmussen,† Jan Haller,‡ Thomas Strassner,‡ and K. N. Houk‡

Contribution from the Department of Medicinal Chemistry, Royal Danish School of Pharmacy, Universitetsparken 2, DK-2100 Copenhagen, Denmark, and Department of Chemistry and Biochemistry, University of California, Los Angeles, California 90095

Received June 15, 1999

Abstract: A new transition state force field has been developed for the AD reaction, purely from quantum mechanical reference data. A new methodology was used for converting quantum mechanical normal modes into a form suitable for parametrization. The force field has been thoroughly validated by comparison to structural and energetic data, and by prediction of experimental enantioselectivities. Excellent agreement was observed, frequently within a few percent of the experimental enantioselectivity. The interactions responsible for enantioselectivity have been identified and compared to the Sharpless and Corey models.

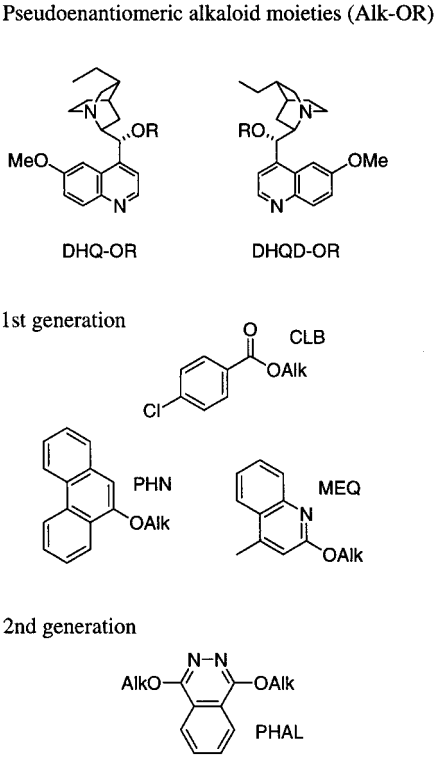
Introduction

The osmium tetroxide asymmetric dihydroxylation of alkenes (AD) is one of the most powerful tools in the field of asymmetric synthesis.[1,2] The scope is wide, the conditions are mild, and both yield and enantioselectivity are frequently very high. The AD reaction is run in the presence of suitably substituted cinchona ligands (Chart 1).

In the first generation ligands, the hydroxy group in dihydroquinine (DHQ) or dihydroquinidine (DHQD) is derivatized with any of several aromatic moieties.[3] In the second generation ligands, a symmetric linker couples two alkaloid units.[4] It has been demonstrated that the role of the second alkaloid moiety is to extend the binding surface for the substrate, not to interact with the osmium moiety. Very similar results can be obtained using large aromatic units without heteroatoms in lieu of the second alkaloid moiety.[5] However, the dimeric formulation facilitates synthesis of the ligand and increases the effective concentration of active quinuclidine units in solution.[1]

The overall catalytic cycle is shown in Scheme 1. Representatives of the two types of osmium complexes have been characterized by X-ray crystallography.[6] The hydrolysis or reoxidation step may be rate-limiting under standard conditions. The exact timing here is of crucial importance; direct oxidation of osmium before hydrolysis will lead to "second cycle" dihydroxylation, with a resulting marked decrease in stereoselectivity.[7] As long as this step can be controlled efficiently, the

Chart 1. Selected Ligands for the AD Reaction



* Email: peo@compchem.dfh.dk.
† Royal Danish School of Pharmacy.
‡ University of California.

(1) Kolb, H. C.; VanNieuwenzhe, M. S.; Sharpless, K. B. *Chem. Rev.* **1994**, *94*, 2483−2547.

(2) (a) Johnson, R. A.; Sharpless, K. B. In *Catalytic Asymmetric Synthesis;* Ojima, I., Ed.; VCH: New York, 1993; pp 227−272. (b) Lohray, B. B. *Tetrahedron: Asymmetry* **1992**, *3*, 1317−1349. (c) Kolb, H. C.; Sharpless, K. B. In *Transition Metals for Organic Synthesis;* Beller, M., Bolm, C., Eds.; VCH: Weinheim, 1998; Vol. 2, pp 219−242. (d) Svendsen, J. S.; Markó, I. In *Comprehensive Asymmetric Catalysis;* Jacobsen, E. N., Yamamoto, H., Eds.; Springer-Verlag: New York, in press. See also: (e) Jørgensen, K. A. In *Transition Metals for Organic Synthesis;* Beller, M., Bolm, C., Eds.; VCH: Weinheim, 1998; Vol. 2, pp 157−172.

(3) Sharpless, K. B.; Amberg, W.; Beller, M.; Chen, H.; Hartung, J.; Kawanami, Y.; Lübben, D.; Manoury, E.; Ogino, Y.; Shibata, T.; Ukita, T. *J. Org. Chem.* **1991**, *56*, 4585−4588.

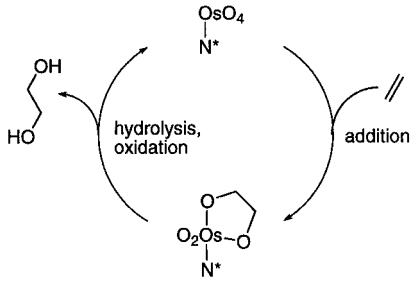
(4) Sharpless, K. B.; Amberg, W.; Bennani, Y. L.; Crispino, G. A.; Hartung, J.; Jeong, K.-S.; Kwong, H.-L.; Morikawa, K.; Wang, Z.-M.; Xu, D.; Zhang, X.-L. *J. Org. Chem.* **1992**, *57*, 2768−2771.

(5) Kolb, H. C.; Andersson, P. G.; Bennani, Y. L.; Crispino, G. A.; Jeong, K.-Y.; Kwong, H.-L.; Sharpless, K. B. *J. Am. Chem. Soc.* **1993**, *115*, 12226−12227.

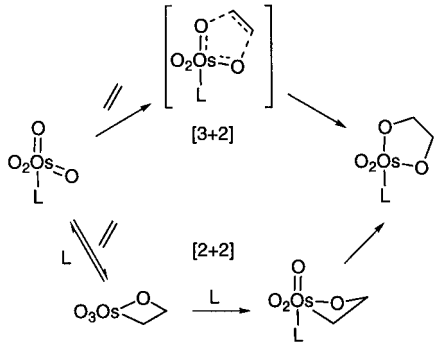
(6) (a) Svendsen, J. S.; Markó, I.; Jacobsen, E. N.; Rao, C. P.; Bott, S.; Sharpless, K. B. *J. Org. Chem.* **1989**, *54*, 2263−2264. (b) Pearlstein, R. M.; Blackburn, B. K.; Davis, W. M.; Sharpless, K. B. *Angew. Chem., Int. Ed. Engl.* **1990**, *29*, 639−641.

10.1021/ja992023n CCC: $18.00
Published on Web 10/14/1999

Scheme 1. The Catalytic Cycle in the AD Reaction; the Alkaloid Ligand Is Shown as N*



Scheme 2. Mechanistic Proposals for the Addition Step



stereoselectivity will be determined solely in the irreversible addition step.[8]

The exact mechanism of addition and the interactions determining the stereoselectivity have been subjects of intense debate. The ligand-accelerated version of the dihydroxylation was first studied by Criegee,[9] who proposed a concerted [3 + 2] addition (Scheme 2). Later, Sharpless argued that the electron-deficient osmium should be the electrophilic center, and proposed a mechanism where the alkene first coordinates to osmium, then slips into an osmaoxetane (formally a [2+2] addition, Scheme 2).[10] The two mechanisms are kinetically very similar,[11] making an experimental differentiation complex. Experimentally observed nonlinear behavior in modified Eyring-plots supports a mechanism with a reversibly formed intermediate.[12] The presence of an intermediate can also be inferred from recent observations of Michaelis−Menten effects in the AD reaction,[13] but the nature of the intermediate can still be debated. It has been suggested that a precomplex can be formed which is stabilized by strong nonbonded interactions between substituents in ligand and substrate.[13,14] This picture is in line with kinetic data for the combined effect of varying ligand and substrate sizes,[15] showing the accelerating effect of stabilizing interactions in the AD reaction (i.e., the fastest reaction in the study was between the bulkiest substrate and ligand). Electronic effects in both ligand and substrate can also be interpreted in favor of either mechanism.[16]

Great advances in electronic structure methods have recently made it possible to perform meaningful studies of model systems by high level methods. An early Extended Hückel investigation identified frontier orbitals that could rationalize the reaction in terms of the [3 + 2] mechanism.[17] On the other hand, the first DFT and GVB studies of the reaction class indicated that the metallaoxetane is a plausible intermediate.[18] However, more recent studies have located the proposed transition states in the reaction by DFT methods,[19,20] providing strong support for the concerted [3 + 2] mechanism. It has also been shown that the [3 + 2] TS, but not any of the [2 + 2] TS's, can rationalize observed isotope effects in the reaction,[20,21] and that calculations on model [3 + 2] TS's can be used to explain the diastereoselectivity in dihydroxylation of allylic ethers.[22]

In addition to the controversy about the mechanism, there has also been an ongoing debate about the source of the high stereoselectivity of the reaction. It was early recognized that qualitative stereoselectivity predictions for reactions employing AD ligands (Chart 1) could be obtained from a mnemonic device[3,4] featuring a simplistic model of the selectivity-determining TS, by assuming that two substituent positions *trans* to each other are hindered (one more than the other). Thus, substrates that have to interact with one of these sites (e.g., *cis*-disubstituted alkenes) react slowly and show low stereoselectivity.[23] The model has been refined to include an attractive interaction, to rationalize the observed influences of steric effects on reaction rates.[1,15] Sharpless et al. have attributed the stabilization to interactions between an alkene substituent and the aromatic −OR moiety in AD ligands.[15] These stabilizing interactions could also be identified by force field calculations on several substrates.[24] Despite the errors introduced by assuming a metallaoxetane center, the calculations could be used to rationalize structure/stereoselectivity correlations.[25] On the other hand, Corey et al. based a qualitative stereoselectivity model for the second generation ligands (Chart 1) on X-ray structures of derivatized ligands[26] and postulated that the attractive interactions are the result of sandwiching the alkene

(7) Wai, J. S. M.; Markó, I.; Svendsen, J. S.; Finn, M. G.; Jacobsen, E. N.; Sharpless, K. B. *J. Am. Chem. Soc.* **1989**, *111*, 1123−1125.

(8) (a) Kwong, H.-L.; Sorato, C.; Ogino, Y.; Chen, H.; Sharpless, K. B. *Tetrahedron Lett.* **1990**, *31*, 2999−3002. (b) Ogino, Y.; Chen, H.; Kwong, H.-L.; Sharpless, K. B. *Tetrahedron Lett.* **1991**, *32*, 3965−3968.

(9) Criegee, R. *Justus Liebigs Ann. Chem.* **1936**, *522*, 75−96.

(10) (a) Sharpless, K. B.; Teranishi, A. Y.; Bäckvall, J.-E. *J. Am. Chem. Soc.* **1977**, *99*, 3120−3128. (b) Hentges, S. G.; Sharpless, K. B. *J. Am. Chem. Soc.* **1980**, *102*, 4263−4265. (c) Jørgensen, K. A.; Schiøtt, B. *Chem. Rev.* **1990**, *90*, 0, 1483−1506.

(11) Norrby, P.-O.; Gable, K. P. *J. Chem. Soc., Perkin Trans. 2* **1996**, 171−178.

(12) Göbel, T.; Sharpless, K. B. *Angew. Chem., Int. Ed. Engl.* **1993**, *32*, 1329−1331.

(13) Corey, E. J.; Noe, M. C. *J. Am. Chem. Soc.* **1996**, *118*, 319−329.

(14) Ujaque, G.; Maseras, F.; Lledós, A. *J. Org. Chem.* **1997**, *62*, 7892−7894.

(15) Kolb, H. C.; Andersson, P. G.; Sharpless, K. B. *J. Am. Chem. Soc.* **1994**, *116*, 1278−1291.

(16) Nelson, D. W.; Sharpless, K. B.; Gypser, A.; Ho, P. T.; Kolb, H. C.; Kondo, T.; Kwong, H.-L.; McGrath, D. V.; Rubin, A. E.; Norrby, P.-O.; Gable, K. P. *J. Am. Chem. Soc.* **1997**, *119*, 1840−1858.

(17) (a) Jørgensen, K. A.; Hoffmann, R. *J. Am. Chem. Soc.* **1986**, *108*, 1867−1876. (b) Jørgensen, K. A. *Tetrahedron Lett.* **1990**, *31*, 6417−6420.

(18) (a) Rappé, A. K.; Goddard, W. A., III. *J. Am. Chem. Soc.* **1982**, *104*, 3287−3294. (b) Norrby, P.-O.; Kolb, H. C.; Sharpless, K. B. *Organometallics* **1994**, *13*, 344−347. (c) Veldkamp, A.; Frenking, G. *J. Am. Chem. Soc.* **1994**, *116*, 4937−4946. (d) Norrby, P.-O.; Becker, H.; Sharpless, K. B. *J. Am. Chem. Soc.*, **1996**, *118*, 35−42.

(19) (a) Pidun, U.; Boehme, C.; Frenking, G. *Angew. Chem., Int. Ed. Engl.* **1996**, *35*, 2817−2820. (b) Dapprich, S.; Ujaque, G.; Maseras, F.; Lledós, A.; Musaev, D. G.; Morokuma, K. *J. Am. Chem. Soc.* **1996**, *118*, 11660−11661. (c) Torrent, M.; Deng, L.; Duran, M.; Sola, M.; Ziegler, T. *Organometallics* **1997**, *16*, 13−19.

(20) DelMonte, A. J.; Haller, J.; Houk, K. N.; Sharpless, K. B.; Singleton, D. A.; Strassner, T.; Thomas, A. A. *J. Am. Chem. Soc.* **1997**, *119*, 9907−9908.

(21) Corey, E. J.; Noe, M. C.; Grogan, M. J. *Tetrahedron Lett.* **1996**, *37*, 4899−4902.

(22) Haller, J.; Strassner, T.; Houk, K. N. *J. Am. Chem. Soc.* **1997**, *119*, 8031−8034.

(23) Andersson, P. G.; Sharpless, K. B. *J. Am. Chem. Soc.* **1993**, *115*, 7047−7048.

(24) Norrby, P.-O.; Kolb, H. C.; Sharpless, K. B. *J. Am. Chem. Soc.* **1994**, *116*, 8470−8478.

(25) Becker, H.; Ho, P. T.; Kolb, H. C.; Loren, S.; Norrby, P.-O.; Sharpless, K. B. *Tetrahedron Lett.* **1994**, *35*, 7315−7318.

(26) (a) Corey, E. J.; Noe, M. C.; Sarshar, S. *Tetrahedron Lett.* **1994**, *35*, 2861−2864. (b) Corey, E. J.; Noe, M. C.; Grogan, M. J. *Tetrahedron Lett.* **1994**, *35*, 6427−6430.

substituent between the two quinoline units of the ligand. The model has been able to rationalize the observed stereoselectivity for many substrates.[27]

Both of the existing stereoselectivity models include a bias; the Sharpless model has been parametrized to fit the experimental selectivities,[24] whereas the Corey model is based on visual inspection of models, not on computed energy differences.[26] To achieve unbiased predictions, it is necessary to compute relative activation energies of competing paths without using the observations to be predicted in the model. Very recently, Maseras et al. published an IMOMM study of the dihydroxylation of styrene using a second generation ligand, $DHQD_2PYDZ$.[28] Transition states corresponding to the 12 possible approach vectors[15] were located and analyzed. Experimentally, this system yields an ee of 96%, compared to a calculated value of 99.4%, a remarkably good agreement. It was also concluded that the approach vectors corresponding to the Sharpless and Corey stereoselectivity models were indeed the lowest energy paths found, differing only slightly in energy ($\Delta\Delta E^* = 0.4$ kJ/mol, in favor of the Corey model).

It has been shown repeatedly that good predictions about relative activation barriers for diastereomeric transition states can be obtained from carefully designed force fields.[29] In line with our previous force field studies of the osmylation reaction,[24,30] we here want to demonstrate how an unbiased force field, created solely from data obtained by quantum mechanical (QM) calculations, can be used to rationalize the stereoselectivity of the AD reaction for widely different ligand-substrate combinations.

Methods

Detailed information about the force field development can be found in the Supporting Information. We have implemented our new parameter set in the MM3* force field within the MacroModel package, one of the most accurate force fields available today for the calculation of conformational energies.[31] Like most current molecular mechanics packages, MacroModel does not incorporate tools for reliable location of transition states, especially for bond formation. For this and other reasons, we have treated the transition state as an energy minimum in the force field.[29] This technique has proven to be remarkably successful in the rationalization and prediction of stereoselectivities in hydroborations,[32] Diels−Alder cycloadditions[33] and ene reactions,[34] radical cyclizations,[35] nucleophilic additions to carbonyls,[36] aldol reactions,[37] Horner-Wadsworth-Emmons reactions,[38,39] and osmium tetroxide dihydroxylations with diamine ligands.[30]

In the computational model, the apical and one equatorial oxygen in the OsO_4L complex are bound to one alkene carbon each, using the coordination bond type in MacroModel. Many of the existing parameters for substituents on the reacting alkene could therefore be used without modification. A few existing parameters were refined to fit the current context. The nitrogen ligand is bound to osmium using the same coordination bond type.[40] The apical and equatorial oxygens in the trigonal bipyramidal complex are identified by the angle to the coordinated nitrogen; an oxygen with an initial N−Os═O angle >140° is considered apical by the force field, and assigned parameters accordingly.

The parameters were refined to fit data from high quality QM calculations as described by Norrby and Liljefors.[45] A large set of TS structures and energies were already available, using ammonia as a simple ligand model.[20,22] A few additional TS structures were generated at the same level of theory, using trimethylamine as a more accurate model of the ligand.[41] Details of the fitting procedure are available in Supporting Information. However, we want to point out our novel use of QM normal mode data in the parametrization. Inclusion of such data in the parametrization leads to an accurate description of the PES around the stationary point,[42−45] but the normal mode corresponding to the reaction coordinate must be modified to a positive curvature to allow inclusion in the chosen force field paradigm.[38]

As opposed to original MM3,[46] the MM3* force field uses atomic point charges calculated from a charge flux parameter (denoted "bond dipole" in the force field). Previously, these "dipoles" have been modified by hand to yield a close correspondence with ChelpG charges,[47] and kept fixed during parameter refinement.[48] In the current implementation, the "dipoles" were refined, and the ChelpG charges for three structures were used as reference data. This procedure will keep the final atomic charges close to the ChelpG charges, but will allow small variations if a substantial improvement can be obtained for other data points by variation of the charges.

The final force field has been tested for ability to reproduce the reference data used in the parametrization (internal predictivity). This is shown in the results section, as overlays between selected reference and force field structures, and plots of other data points. The external predictivity (the ability to reproduce data that has not been included in the model) was then tested. We have selected a set of experimental selectivities from the literature, including both first and second generation ligands, and representatives of five of the six alkene classes.[23] The selectivities are calculated by an unbiased search for all low energy conformations, followed by Boltzmann averaging based on the calculated potential energies.[49] Several criteria must be fulfilled for this rather severe test to succeed. First of all, since only quantum mechanical data

(27) Corey, E. J.; Noe, M. C. *J. Am. Chem. Soc.* **1996**, *118*, 11038−11053, and references therein.
(28) Ujaque, G.; Maseras, F.; Lledós, A. *J. Am. Chem. Soc.* **1999**, *121*, 1317−1323.
(29) Eksterowicz, J. E.; Houk, K. N. *Chem. Rev.* **1993**, *93*, 2439−2461.
(30) Wu, Y.-D.; Wang, Y.; Houk, K. N. *J. Org. Chem.* **1992**, *57*, 1362−1369.
(31) Gundertofte, K.; Liljefors, T.; Norrby, P.-O.; Pettersson, I. *J. Comput. Chem.* **1996**, *17*, 429−449.
(32) (a) Houk, K. N.; Rondan, N. G.; Wu, Y.-D.; Metz, J. T.; Paddon-Row, M. N. *Tetrahedron* **1984**, *12*, 2257−2274. (b) Houk, K. N.; Paddon-Row, M. N.; Rondan, N. G.; Wu, Y.-D.; Brown, F. K.; Spellmeyer, D. C.; Metz, J. T.; Li, Y.; Loncharich, R. J. *Science* **1986**, *231*, 1108−1117.
(33) Brown, F. K.; Houk, K. N. *J. Am. Chem. Soc.* **1985**, *107*, 1971−1978.
(34) Thomas, B. E., IV.; Loncharich, R. J.; Houk, K. N. *J. Org. Chem.* **1992**, *57*, 1354−1362.
(35) (a) Spellmeyer, D. C.; Houk, K. N. *J. Org. Chem.* **1987**, *52*, 959−974. (b) Broeker, J. L.; Houk, K. N. *J. Org. Chem.* **1991**, *56*, 3651−3655.
(36) (a) Wu, Y.-D.; Houk, K. N. *J. Am. Chem. Soc.* **1987**, *109*, 908−910. (b) Wu, Y.-D.; Houk, K. N.; Trost, B. M. *J. Am. Chem. Soc.* **1987**, *109*, 5560−5561. (c) Mukherjee, D.; Wu, Y.-D.; Fronczek, F. R.; Houk, K. N. *J. Am. Chem. Soc.* **1988**, *110*, 3328−3330. (d) Wu, Y.-D.; Houk, K. N.; Florez, J.; Trost, B. M. *J. Org. Chem.* **1991**, *56*, 3656−3664.
(37) Bernardi, A.; Gennari, C.; Goodman, J. M.; Paterson, I. *Tetrahedron: Asymmetry* **1995**, *6*, 2613−2636.
(38) Norrby, P.-O. In *Transition State Modeling for Catalysis:* Truhlar, D. G., Morokuma, K., Eds.; ACS Symposium Series No. 721, 1999; Chapter 13, pp 163−172.
(39) (a) Tullis, J. S.; Vares, L.; Kann, N.; Norrby, P.-O.; Rein, T. *J. Org. Chem.* **1998**, *63*, 8284−8294. (b) Norrby, P.-O.; Brandt, P.; Rein, T. *J. Org. Chem.* **1999**, *64*, 5845−5852.
(40) Note that the ligand nitrogen will erroneously be assigned the N(sp^2) atom type by the graphical interface; the atom type must be modified to N(sp^3) (internally type 26) in the structure file before calculations can be run.
(41) In calculated ammine complexes, the N−H bonds are eclipsed with the Os═O bonds. Using a tertiary amine, the C−N bonds are now staggered with the Os═O bonds, in accordance with known structures (ref 6).
(42) Dasgupta, S.; Goddard, W. A., III *J. Chem. Phys.* **1989**, *90*, 7207−7215.
(43) Maple, J. R.; Hwang, M.-J.; Stockfisch, T. P.; Dinur, U.; Waldman, M.; Ewig, C. S.; Hagler, A. T. *J. Comput. Chem.* **1994**, *15*, 162−182.
(44) Halgren, T. A. *J. Comput. Chem.* **1996**, *17*, 490−519.
(45) Norrby, P.-O.; Liljefors, T. *J. Comput. Chem.* **1998**, *19*, 1146−1166.
(46) Allinger, N. L.; Yuh, Y. H.; Lii, J.-H. *J. Am. Chem. Soc.* **1989**, *111*, 8551−8566.
(47) Breneman, C. M.; Wiberg, K. B. *J. Comput. Chem.* **1990**, *11*, 361−373.
(48) Brandt, P.; Norrby, T.; Åkermark, B.; Norrby, P.-O. *Inorg. Chem.* **1998**, *37*, 4120−4127.
(49) We have ignored possible differences in solvation energy and vibrational contributions, as these can be expected to cancel in a comparison of diastereomeric structures with low polarity.

Table 1. Calculated and Experimental Enantioselectivities

entry	alkene	DHQD ligand[a]	ee_{calc} (%)	ee_{exp} (%)	pocket[b]	ref
1	1-phenyl-cyclohexene	CLB	91	91	C (3)	3
2	styrene	CLB	70	74	S (1)	3
3	β,β-dimethyl styrene	CLB	72	74	S (7)	3
4	β-vinyl naphthalene	CLB	94	88	S (2)	3
5	*trans*-stilbene	CLB	98	99	both	3
6	*tert*-butyl ethene	CLB	70	44	S (3)	3
7	α-methyl styrene	CLB	65	62	equal	62
8	*cis*-β-methyl styrene	CLB	78	35	S (6)[c]	62
9	styrene	MEQ	94	87	C (7)	3
10	styrene	PHN	98	78	C (4)	3
11	*tert*-butyl ethene	PHN	89	79	S (1)	3
12	β-vinyl naphthalene	PHAL	100	98	C (9)	15
13	styrene	PHAL	97	97	C (3)	4
14	α-methyl styrene	PHAL	99	94	C (7)	4
15	*trans*-stilbene	PHAL	100	100	both	4

[a] Chart 1. [b] The pocket favored by the large substituent of the alkene; C = close to the quinoline moiety (as suggested by Corey et al.); S = over the linker unit (as suggested by Sharpless et al.). The number in parentheses is the difference between the two pockets (in kJ/mol). [c] For the *cis*-disubstituted alkene, the competing binding mode does not correspond to the pocket suggested by Corey et al., but rather to an approach where the methyl is in the Sharpless pocket and the phenyl lacks significant interactions with the ligand.

was used in the parametrization, the high level calculations must give a good description of the real transition state. Second, the force field must reproduce the PES faithfully, even for distorted structures and possible points outside the reference data set. Finally, the conformational search must find all relevant low energy conformations. Only when all these criteria are fulfilled can more than a spurious agreement be expected. The final results are shown in Table 1.

Computational Details. QM calculations were performed in Gaussian94.[50] All transition state structures were determined at the B3LYP[51] level using the LANL2DZ[52] basis set for osmium and 6-31G* for the remaining atoms. Normal modes were computed numerically and charges were calculated for the same structures using the ChelpG method.[47] Force field calculations were performed in MacroModel[53] on Silicon Graphics workstations, using a modified MM3* force field.[54] Conformational searches were performed using a combination of pseudo-systematic Monte Carlo[55] and Low Mode[56] searching. In our experience, the latter method is excellent for exhaustive searches of local regions of the potential energy surface (PES), but will not easily cross extensive regions of high energy, whereas the former explores the entire conformational space, but rather coarsely. Thus, the two methods complement each other perfectly. All searches were initialized by generation of low energy structures corresponding to the 12 possible approach vectors, whereupon 500 Monte Carlo steps were performed. The resulting output was subjected to 1000−2000 steps of Low Mode searching. The output was sorted into geometrically similar groups,[53] and each group resubmitted to a new Low Mode search. Finally, the relative rate through each distinct path was obtained from a Boltzmann distribution based on the calculated potential energy barriers and T = 298K. Entropy and solvation contributions were ignored.

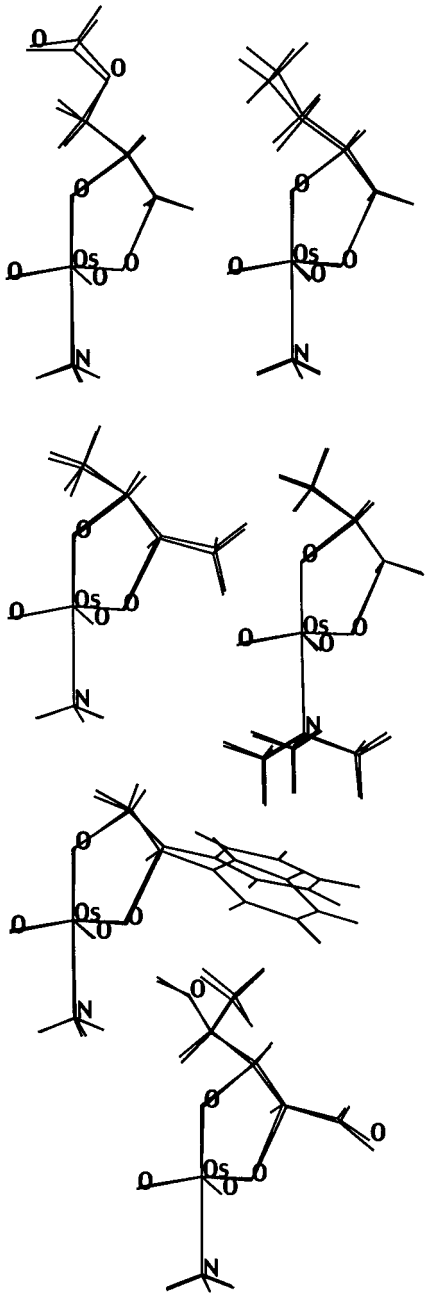


Figure 1. Selected overlays of force field and QM minima.

Results

The final force field parameters are listed in Supporting Information, and are also available for downloading via the Internet.[57] Selected structure overlays are shown in Figure 1. The overlays have been generated by superposition of the central OsO_4 moiety, to highlight positional deviations of the ligand

(50) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Gill, P. M. W.; Johnson, B. G.; Robb, M. A.; Cheesemen, J. R.; Keith, T. A.; Petersson, J. A.; Montgomery, J. A.; Raghavachari, K.; Al-Laham, M. A.; Zakrzewski, V. G.; Ortiz, J. V.; Foresman, J. B.; Cioslowski, J.; Stefanov, B. B.; Nanayakkara, A.; Challacombe, M.; Peng, C. Y.; Ayala, P. Y.; Chen, W.; Wong, M. W.; Andres, J. L.; Replogle, E. S.; Gomperts, R.; Martin, R. L.; Fox, D. J.; Binkley, J. S.; DeFrees, D. J.; Baker, J.; Stewart, J. J. P.; Head-Gordon, M.; Gonzales, C.; Pople, J. A. *Gaussian 94*; Gaussian Inc.: Pittsburgh, PA, 1995.

(51) (a) Becke, A. D. *J. Chem. Phys.* **1993**, *98*, 5648−5652. (b) Lee, C.; Yang, W.; Parr, R. G. *Phys. Rev.* **1988**, *37*, 785−789.

(52) Hay, P. J.; Wadt, W. R. *J. Chem. Phys.* **1985**, *82*, 299−310.

(53) *MacroModel*, V6.0; Mohamadi, F.; Richards, N. G. J.; Guida, W. C.; Liskamp, R.; Lipton, M.; Caulfield, C.; Chang, G.; Hendrickson, T.; Still, W. C. *J. Comput. Chem.* **1990**, *11*, 440−467. Note! A bug in MacroModel will give serious errors if conformational searches are performed using force fields with geometry-dependent parameters (like the current force field). It is necessary to divide the searches into groups where all structures have the same parameter values, and to ensure (by constraints) that no crossover takes place. Starting from V6.5, correct behavior can be obtained using the "DEBG 57" command.

(54) MM3* is based on the MM3(91) force field (ref 46). The major differences are: A substructure-matching scheme in lieu of the π-system calculation; use of atomic charges only, no dipoles; nondirectional hydrogen bonding adapted from the Amber* force field.

(55) Goodman, J. M.; Still, W. C. *J. Comput. Chem.* **1991**, *12*, 1110−1117.

(56) Kolossvary, I.; Guida, W. C. *J. Am. Chem. Soc.* **1996**, *118*, 5011−5019.

(57) A MacroModel force field file can be downloaded in a directly usuable format from http://compchem.dfh.dk/PeO/.

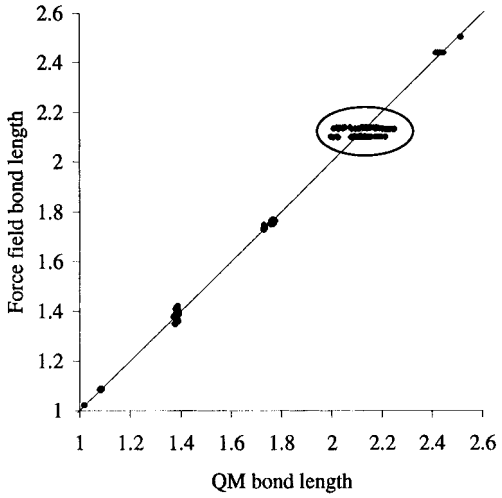


Figure 2. Force field vs QM bond lengths. The forming C··O bonds are encircled.

and alkene. Figure 2 shows the calculated lengths of all bonds for which a parameter has been determined. Additional comparisons of force field and reference data are available as Supporting Information.

The calculated and experimental enantioselectivities for 15 different substrate−ligand combinations are shown in Table 1. The major enantiomer is always formed through a TS where the large alkene substituent is stabilized by interactions with one of the two binding pockets offered by the ligand (as suggested by Sharpless et al.[15,24,25] and Corey et al.[26,27]). The geometries of the two possible approach vectors are illustrated in Figure 3. The particular binding mode preferred by each substrate is indicated in Table 1, together with the energetic preference over the other pocket (note that both pockets lead to the major enantiomer). Note that in the first generation ligands (entries 1−11), there can be no sandwich interaction. Thus, the binding does not correspond to the Corey suggestion. However, the alkene approach vector and the alkaloid geometry are similar, so this path has still been denoted "C" for consistency with the second generation ligand (entries 12−15).

Discussion

In transition state modeling by pure force field methods, it has in general been considered necessary to avoid true saddle point searches.[58] Traditionally, two methods have been used to achieve this: either the reaction center has been frozen at an appropriate geometry while the remainder of the system has been optimized, or a force field has been created that delivers the transition state structure as a minimum. The latter approach, while more flexible, has been notoriously hard to implement due to the problems of finding a unique and predictive set of parameters from the small amount of data points available. The problems associated with transition state parametrization have recently been alleviated by a new methodology, utilizing not only structures and energies but also modified normal modes from high level QM calculations.[38,39] The additional data allow a unique determination of many energy-related parameters, like force constants and torsional parameters. The effect of the modification of the normal modes is that all distortions along the reaction coordinate will result in steep energy increases, whereas distortions perpendicular to the reaction coordinate will correspond closely to the results from the high level QM calculations. The methodology has been implemented within a freely available framework capable of being adapted to parametrization of most current force fields.[45] From the point when all the QM data had been determined, the entire parametrization procedure was completed in about a month. Thus, it is now possible to go from QM data to a predictive model for experimentally interesting systems in a time frame that fits into normal project plans.[59]

Internal Validation. The first step in a validation should be a verification that the reference data are sufficiently well described by the force field. If possible, it should also be verified that the parameters are optimal within the chosen context. Both of these points have been amply demonstrated here. The structural agreement is demonstrated in Figures 1 and 2. Overall, the agreement is good. The mean absolute error over all interatomic distances (not just bonds) is only 0.05 Å. A few points are worth mentioning.

First of all, some systematic errors in the reaction coordinate are unavoidable using the current procedure. The response of true transition states to steric crowding is opposite to what is expected for a minimum: the bulky substrates actually have shortened forming C−O bonds compared to unsubstituted alkenes. The force field will instead elongate the bonds slightly in order to relieve the nonbonded repulsions. Our tentative solution has been to minimize the error by introducing a large artificial eigenvalue for the normal mode corresponding to the reaction coordinate.[38] We cannot distort the complex correctly; therefore, we try to minimize *all* distortions along the reaction coordinate in the hope that selectivity-determining interactions will only be weakly dependent on small deviations along the reaction coordinate. It can be seen in Figure 2 that the force field follows our intentions: there is substantial variation in the lengths of the forming C−O bonds in the QM structures but almost none in the corresponding force field results. A more complex correction would be to correct the parameters iteratively using the quotient of the original and imposed force constant for the normal mode corresponding to the reaction coordinate in response to distortions. However, this would require identification of the reaction coordinate with one or very few internal coordinates. More importantly, a dynamic parameter update is not easily introduced into current molecular mechanics packages. It was therefore decided not to implement the correction until it was shown to be needed. The comparison to experimental results (vide infra) fully validate this simplification.

An important structural parameter is the rotation around the N−Os bond. In the ammine complexes, the N−H bonds are eclipsed with the Os═O bonds,[19] whereas in tertiary amine complexes the bonds are staggered[6] (Figure 1). Maseras et al. have argued that the position of the hydrogens in the QM structures makes development of force field parameters problematic in this case.[28] The natural answer is to include data with the correct orientation in the reference data set. The good structural correspondence clearly shows that it is possible to design a force field that describes both types of systems correctly. The reaction center seems fairly independent of the ligand model used. The propensity of an N−H bond to eclipse with an Os═O bond can easily be understood on electrostatic grounds. By a correct description of the van der Waals interactions and the electrostatics, the conformational preferences of both ammine and amine complexes are well-described.

(58) For examples of exceptions, see: (a) Jensen, F. *J. Comput. Chem.* **1994**, *15*, 1199−1216. (b) Åqvist, J.; Warshel, A. *Chem. Rev.* **1993**, *93*, 2523−2544.

(59) As a personal experience, introduction of force field methods into selectivity predictions has previously been limited by the long development time; by the time the model was finished, the synthetic project could easily have moved into another phase or been abandoned altogether.

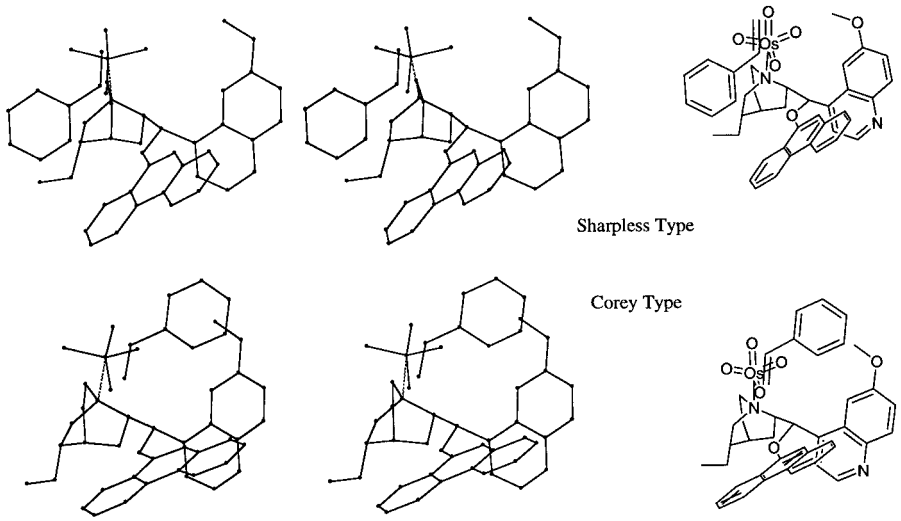


Figure 3. Stereoview of the two paths leading to the major enantiomer (entry 10, Table 1).

Comparison to Experiment. At this point, we want to reiterate the fact that no experimental information has gone into our development of the force field. The only bias is in the choice of mechanism to consider and in the weight factors for the various types of reference data. Calculated enantioselectivities can therefore be taken as predictions, not rationalizations.

The correlation shown in Table 1 is excellent. In all cases the correct enantiomer is predicted. In most cases, the predictions are also within a few percent of the experimental enantioselectivity.[60] Given that the experimental selectivities generally vary by a few percent depending on reaction conditions, most predictions are within the experimental deviation. For the few points that are slightly less well described, entries 6 and 8, the calculations overestimate the enantioselectivity (by 2.0 and 3.8 kJ/mol, respectively). These are two of the slowest substrates. In these cases, it is not impossible that a side reaction (e.g., the known second cycle[7,8]) competes with the regular AD reaction, lowering the selectivity. The same may in fact be true for the best substrate of them all, entry 15. Superficially, this seems to be an exact prediction, but from an energy comparison, it is actually the worst correspondence in the study. The experimental value has been determined to 99.8%, whereas the calculated value is 99.999% (the best conformation of the minor enantiomer is in fact outside the ligand pocket shown in Figure 4).

Looking at the source of the enantioselectivity, our conclusions for the reaction of styrene using the $DHQD_2PHAL$ ligand are very similar to those of Maseras et al. for the reaction of styrene with the very similar $DHQD_2PYDZ$ ligand.[28] Two ligand pockets can be utilized for binding to the substrate, leading to two favored approach vectors, both resulting in the same enantiomer (Figure 3). There is little energetic difference between the two pockets, ~3 kJ/mol (entry 13; this is a sum over several contributing conformers). It should be noted that *trans*-disubstituted alkenes utilize both pockets for binding. Entry 15 is illustrated in Figure 4. Both phenyl groups of the stilbene experience stabilizing interactions, one with the PHAL linker and one with the bystander quinoline. This result is in perfect agreement with the experimental observation that *trans*-substituted alkenes are the best substrates for the AD reaction,

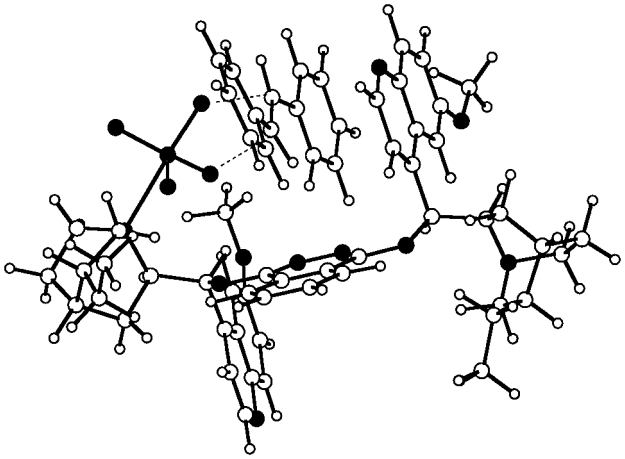


Figure 4. The optimum structure for AD-reaction of stilbene with the $DHQD_2PHAL$-ligand.

both in terms of reaction rate[23] and enantioselectivity.[1] It can further be seen in Figure 4 that also the alkene core, not only the substituents, is stabilized by the bystander quinoline moiety.

Comparing the binding depicted in Figure 4 to the pocket suggested by Corey et al., it can be seen that the quinolines do not really form a sandwich. However, the position of the bystander quinoline is quite flexible. With a longer distance between the alkene and the aromatic moiety (as in the allyl benzoates investigated by Corey et al.[26]), the bystander can rotate to form a proper sandwich structure.

Extending the analysis to the first generation ligands, it can be seen that both pockets are utilized for most substrates, with small and varying preferences for one over the other. The CLB ligand generally favors the binding mode suggested by Sharpless, the PHN ligand is intermediate, and the MEQ favors the same approach vector used in the Corey model for the second generation ligands. The preference can be easily understood from Figure 5. The middle model depicts interactions of the alkene substituents with the PHN moiety (see also the stereoview in Figure 3). The phenanthrene is large enough to stabilize either the R_1 or the R_3 substituent. The CLB unit corresponds closely to the lower part of the PHN moiety, interacting only with R_1. On the other hand, the MEQ unit simulates the upper half of PHN, providing stabilization only to the R_3 substituent.[61] It can

(60) Note that when comparing enantioselectivities, the accuracy is not the same in the entire range. For example, assuming that the calculations are accurate to within ~2 kJ/mol (see Supporting Information and ref 31), a predicted ee of 99.0% should fall within the range 97.7−99.5%, whereas the same accuracy in energy for a predicted value of 75% would correspond to the range 51−88%.

(61) The orientation shown in Figure 5 is strongly favored for all ligands in the current study.

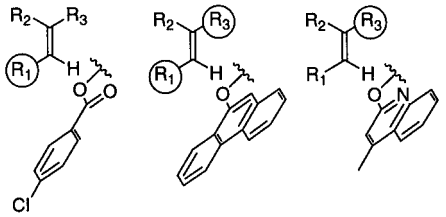


Figure 5. Interactions of first generation ligands with the substrate. Encircled groups experience stabilizing interactions with the ether/ester moiety of the ligand.

also be seen that one substituent is strongly hindered (shown as H in Figure 5), in perfect agreement with the mnemonic device for the reaction.[3,4] Thus, an aromatic alkene substituent cannot occupy the R_3 position if it is cis to another substituent (entry 8, Table 1), whereas an α−substituent disfavors binding in position R_1. The steric hindrance in this position is the main reason for the generally low rate and stereoselectivity for tetrasubstituted alkenes (no orientation can fit well into the binding pocket), and the excellent results for tri- and *trans*-disubstituted alkenes (only one enantiomer can result from binding with a hydrogen in the indicated position).[23]

Conclusions

The results presented here have important implications both for the AD reaction and for modeling of reaction transition states. From a mechanistic viewpoint, the currently favored [3 + 2] mechanism[20] has gained additional support. The interactions leading to enantioselectivity in several substrate-ligand combinations have been identified, allowing easier matching of the optimum ligand for each substrate. For the synthetic chemist, we have made available an easily used force field that can be run using a widely available program package. Synthetic proposals can therefore be tested before implementation. Knowledge of the transition state geometry also allows rational design of new ligands for specific substrates.

From a computational viewpoint, it has here been demonstrated that excellent predictivity can be obtained from pure force field calculations on transition state models. The method for producing the force field is unbiased and straightforward, and should be applicable to many reactions where a good QM description of the transition state(s) is available. The methodology is also very efficient from a resource usage point. Starting from the available QM data, the force field was developed in about a month on an SGI Octane workstation. The subsequent conformational search (including minimization of >100 000 structures) required less than a month. Both in terms of accuracy and resource usage, the methodology described herein is a strong contender with the increasingly popular QM/MM methodology.[28]

Acknowledgment. We are grateful to the Danish Technical Research council and the National Institute of General Medical Sciences, National Institutes of Health, for financial support of this research and to the NPACI-San Diego Supercomputing Center and UCLA Office of Academic Computing for generous allocations of computer time. J.H. thanks the Alexander von Humboldt foundation for a Feodor Lynen fellowship. T.S. thanks the DFG (Deutsche Forschungsgemeinschaft) for a research grant. K.N.H. acknowledges with thanks the early work at UCLA on force field transition state modeling of the asymmetric osmylation reaction by Professors Yun-Dong Wu (Hong Kong University of Science and Technology) and Satomi Niwayama (Oklahoma State University).

Supporting Information Available: Details of force field development (PDF). Force field parameters in MacroModel format (ASCII). This material is available free of charge via the Internet at http://pubs.acs.org.

JA992023N

(62) Sharpless, K. B.; Wang, L., unpublished data. Private communication from K. Barry Sharpless and Bruce Bender.