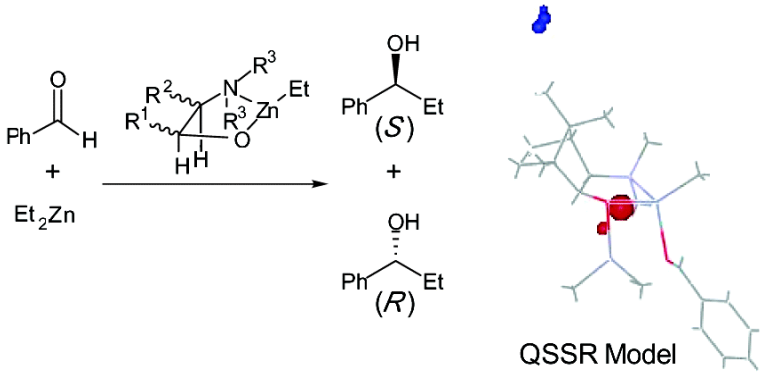
Communication

Quantum Mechanical Models Correlating Structure with Selectivity: Predicting the Enantioselectivity of β-Amino Alcohol Catalysts in Aldehyde Alkylation

Marisa C. Kozlowski, Steven L. Dixon, Manoranjan Panda, and Giorgio Lauri

J. Am. Chem. Soc., **2003**, 125 (22), 6614-6615• DOI: 10.1021/ja0293195 • Publication Date (Web): 08 May 2003



More About This Article

Additional resources and features associated with this article are available within the HTML version:

- Supporting Information
- Links to the 6 articles that cite this article, as of the time of this article download
- Access to high resolution figures
- Links to articles and content related to this article
- Copyright permission to reproduce figures and/or text from this article

View the Full Text HTML

J|A|C|S
COMMUNICATIONS

Published on Web 05/08/2003

Quantum Mechanical Models Correlating Structure with Selectivity: Predicting the Enantioselectivity of β-Amino Alcohol Catalysts in Aldehyde Alkylation

Marisa C. Kozlowski,*,† Steven L. Dixon,‡ Manoranjan Panda,† and Giorgio Lauri‡

Department of Chemistry, Roy and Diana Vagelos Laboratories, University of Pennsylvania, Philadelphia, Pennsylvania 19104-6323, and Pharmacopeia, Inc., Princeton New Jersey 08543-5350

Received November 12, 2002; E-mail: marisa@sas.upenn.edu

Over the last three decades, much effort has been focused on the development of chiral catalysts as means to synthesize molecules with defined handedness.[1] Computational chemistry methods[2] are increasingly practical for the selection of chiral ligands used in asymmetric transformations[3] due to improvements in their theoretical foundations coupled with increased computing speeds. In this report, we describe a quantitative structure selectivity relationship (QSSR) study of the asymmetric addition of Et_2Zn to benzaldehyde catalyzed by chiral β-amino alcohols (Figure 1). Related QSAR methods have received a great deal of attention with respect to the design of pharmaceutical agents,[4,5] but not with respect to small-molecule catalysts.[6,7]

To generate a valid QSSR model, we required experimental selectivities for a set of catalysts with similar characteristics. Since the report of DAIB (**9**) as a highly enantioselective catalyst,[8] a large number of related compounds have been examined.[9] We selected 18 β-amino alcohols (Figure 2) with similar reactivities and that encompass a wide enantioselectivity range. The mechanism of this reaction has also been studied extensively,[10] and μ-oxo transition structures have been proposed. Our goal was to correlate the structures of the aminoalkoxide zinc catalysts (Figure 1) with their selectivities in the addition of Et_2Zn to PhCHO. Unfortunately, calculations of the catalyst ground states afforded planar Zn geometries, which do not reflect the geometry in the catalysts as the reaction proceeds. Thus, several transition structures (two are shown in Figure 1) were located using PM3 methods[10c,11] to obtain an accurate picture of the chiral catalyst structures. While the PM3 transition-structure energies do not always correlate to product selectivity,[10f] the geometries are similar to those from much more time-consuming analyses (DFT, MP2).[10b]

For the subsequent analysis, catalyst geometries found in the *anti S* structures were used. The structures were aligned using the O, Zn, and O atoms (Figure 1, bold) and were divided into appropriate training and prediction sets (Table 1). To define a suitable dependent variable for the QSSR, the enantioselectivities obtained with the different catalysts for the reaction in Figure 1 were translated into ΔG values. The QSSR models were then constructed based on quantum mechanical models.[12] The aligned structures were placed on a grid and a set of PM3[13] *probe interaction energies* (PIEs) was computed separately for each structure using a carbon 2s electron probe at various grid points near the catalyst. Only the contributions from the catalyst portion to the PIEs were used; those from the Me_2Zn and PhCHO substrates were masked.

The PIE data are the pool of independent variables from which linear regression equations of ΔG were generated. To prevent overfitting of the training set, regressions were constructed using only two grid points at a time. The two-variable fits were obtained using

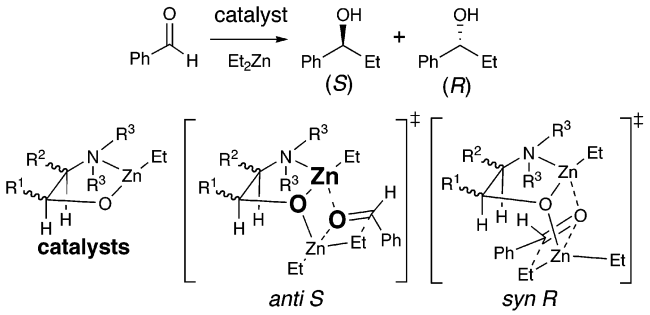


Figure 1. β-Aminoalkoxide-catalyzed Et_2Zn addition to PhCHO.

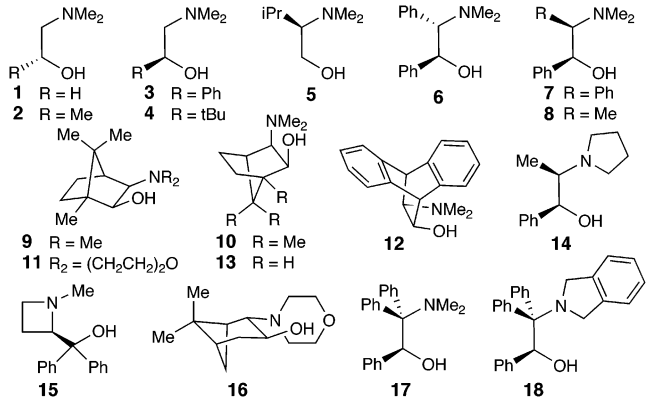


Figure 2. β-Amino alcohol catalysts.

a least-squares regression, with variables selected to optimize internal model fitness via simulated annealing. Results are reported for the *best individual* two-variable model with the lowest standard deviation of regression and for an *average* model comprising a weighted combination of all accepted two-variable regressions.

A model obtained from QSSR calculations using the β-aminoalkoxide catalysts is shown in Table 1. For the *best individual* model the calculated selectivity is expressed as $\Delta G_{fit} = a + c_1(PIE_1) + c_2(PIE_2)$. The two sets of PIE values appear to provide statistically independent information ($r = 0.18$). Most importantly, there is a strong relationship between the observed (ΔG) and calculated (ΔG_{fit}) selectivity for the training *and* prediction sets in Table 1. The model easily distinguishes among catalysts of low, moderate, and high selectivity.

The robustness of this method were established by a series of experiments (Table 2). As the grid spacing (2.0, 1.3, and 0.7 Å) was reduced, the calculated values (ΔG_{fit}) converged onto the experimental values. When the orientation of the grid was changed with respect to the aligned set of structures, the *average* models (grid1 vs grid2) were remarkably similar, also indicating that the calculations have converged.

† University of Pennsylvania.
‡ Pharmacopeia, Inc.

10.1021/ja0293195 CCC: $25.00

Table 1. QSSR Calculations Using Catalysts from **1**−**18**[a]

	expt.		*anti* S1 best				*anti* S1 avg	
cmpd	% ee[b]	ΔG[c]	% ee_{fit}	ΔG_{fit}[d]	PIE_1[e]	PIE_2[e]	% ee_{fit}	ΔG_{fit}
Training Set[f]								
1	0	0.00	28	0.32	25.92	4.76	26	0.30
3	59	0.76	70	0.97	21.79	3.54	68	0.92
4	93	1.85	89	1.61	20.62	4.44	90	1.63
5	49	0.60	24	0.27	26.25	4.87	34	0.39
6	66	0.88	72	1.00	21.92	3.72	70	0.96
7	73	1.04	72	1.00	24.28	5.44	76	1.10
8	81	1.26	79	1.18	23.15	5.10	81	1.26
9	98	2.56	98	2.44	18.39	5.11	98	2.47
10	95	2.04	97	2.35	18.53	4.95	97	2.38
11	98	2.56	97	2.35	19.14	5.39	97	2.37
12	96	2.17	97	2.26	21.55	6.94	96	2.18
13	94	1.94	91	1.69	20.92	4.88	93	1.83
17	94	1.94	95	2.02	21.35	6.12	93	1.89
18	97	2.33	97	2.43	20.97	6.97	96	2.23
Prediction Set[g]								
2	3	0.03	11	0.12	27.91	5.67	5	0.06
14	86	1.43	81	1.25	21.45	4.05	76	1.10
15	98	2.56	99	3.36	15.29	5.36	99	2.82
16	63	0.83	83	1.31	23.66	5.85	75	1.09

[a] Catalyst geometries taken from *anti S* transition structures. Grid1 orientation, 0.7 Å grid spacing. [b] (*S*)-product. [c] The % ee is converted to ΔG (kcal/mol) using $\Delta G = RT \ln K$, K is ratio of the (*R*) and (*S*) enantiomers. [d] $\Delta G_{fit} = a + c_1(PIE_1) + c_2(PIE_2)$; $a = 5.48$ kcal/mol, $c_1 = -0.27$, $c_2 = 0.36$. [e] Probe interaction energies (kcal/mol) at the two grid points identified in the QSSR analysis. [f] best,avg: SD = 0.23, 0.17 kcal/mol; $R^2 = 0.93, 0.95$. [g] best,avg: RMSE = 0.49, 0.29 kcal/mol; $R^2 = 0.72, 0.90$; CC = 0.95, 0.96.

Table 2. Statistical Summary of the QSSR Models

TS[a]/grid spacing	model	RMSE[b]	R^2[c]	CC[d]	predicted R^2[e]	N[f]
grid2/2.0 Å	best	0.81	0.23	0.50	0.32	54
	avg	1.27	-0.87	-0.29	-0.66	54
grid2/ 1.3 Å	best	0.29	0.90	0.99	0.92	174
	avg	0.34	0.86	0.98	0.88	174
grid2/ 0.7 Å	best	0.34	0.86	0.93	0.88	1077
	avg	**0.30**	**0.90**	**0.95**	**0.91**	1077
grid1/0.7 Å	best	0.49	0.72	0.95	0.75	1036
	avg	**0.29**	**0.90**	**0.96**	**0.92**	1036

[a] Grid1 and grid2 refer to the two different grid orientations. [b] RMS error between prediction set ΔG and ΔG_{fit}. [c] Coefficient of multiple determination (percentage of variance accounted for by the model). [d] Correlation coefficient (how well the prediction set selectivity order is calculated) [e] Prediction R^2 from training set mean. [f] Number of sampled gridpoints.

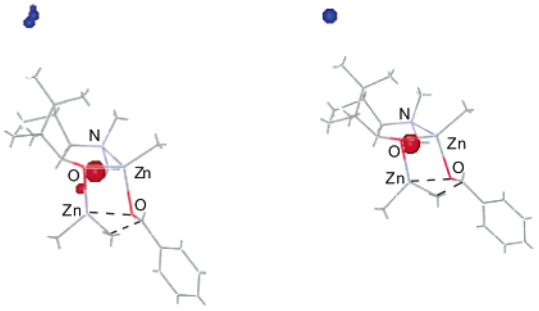


Figure 3. Models from QSSR calculations (grid1, 0.7 Å) superimposed on the *anti S* transition structure from the DAIB ligand (**9**). Left = *average* model; Right = *best individual* model. Blue = ee increases with increasing PIE values; Red = ee decreases.

Comparison of the *best individual* vs *average* models indicates that the latter provide consistently better results. The most reliable external predictions occur when these models yield similar results and grid point arrangements. For example, the grid points in the 0.7 Å *best individual* model provide the major components of the corresponding *average* model (Figure 3). Moreover, the tight spacing of the *average* model grid points indicates that PIE values from nearby catalyst components reproducibly correlate with selectivity. In a cross-validation study, several different training/prediction sets were generated, and similar results were obtained.

Analysis of the positions of the grid points identified in the QSSR model (Figure 3) provides insight into the catalyst features important to the stereochemical induction. One point was found near the sterically large bridging dimethylmethylene of DAIB (**9**) which is also where many other catalysts position large bulky groups. Examination of the transition structures leading to the enantiomeric products indicates this large group is more readily accommodated in the *anti S* rather than the *anti R*, *syn R*, or *syn S* approaches due to steric interactions. A second grid point was located on the bottom face of the five-membered β-aminoalkoxide zinc chelate. Nearby catalyst components destabilize the *syn R* structure via steric interactions and stabilize the *anti S* structure via electrostatic interactions with the carbonyl dipole. We conclude that the QSSR model can correlate chemically relevant portions of the catalyst structure with the observed selectivities.

In summary, we have determined that grid-based QSAR methods can be used to generate statistically valid models that predict the selectivities of catalysts directly from their structure. The strength of the present method is that empirical models obtained from a small set of experimentally determined selectivities and relatively simple theoretical calculations yield selectivity predictions that are as accurate as those derived from much higher-level calculations (DFT, MP2) of transition-structure isomers. Only minutes of computing time are required, and easily interpretable models are obtained which provide realistic and insightful predictions.

Acknowledgment. Financial support of this research was provided by the National Institutes of Health (GM59945). We thank Professor Kenneth Merz for making available to us the DivCon and QM-QSAR programs prior to their release.

Supporting Information Available: Details of the *best individual* model of each calculation described above (PDF). This material is available free of charge via the Internet at http://pubs.acs.org.

References

(1) *Comprehensive Asymmetric Catalysis*; Jacobsen, E. N., Pfaltz, A., Yamamoto, H., Eds.; Springer: New York; 1999.
(2) (a) For reviews see: *Transition State Modeling for Catalysis*; Truhlar, D. G., Morokuma, K., Eds.; ACS Symposium Series 721; American Chemical Society: Washington, DC, 1999; Chapters 1, 3, 13, 14. (b) Kozlowski, M. C.; Panda, M. *J. Mol. Graphics Modell.* **2002**, *20*, 399. (c) Kozlowski, M. C.; Panda, M. *J. Org. Chem.* **2003**, *68*, 2061.
(3) Use of PM3 to design a catalyst for the reaction in Figure 1: Vidal-Ferran, A.; Moyano, A.; Pericás, M. A.; Riera, A. *Tetrahedron Lett.* **1997**, *38*, 8773.
(4) (a) Goodford, P. *J. Med. Chem.* **1985**, *28*, 849. (b) Cramer, R. D.; Patterson, D. E.; Bunce, J. D. *J. Am. Chem. Soc.* **1988**, *110*, 5959.
(5) *3D-QSAR in Drug Design*; Kubinyi, H., Folkers, G., Martin, Y. C., Eds.; Kluwer: Leiden, 1993−1996; Vols. 1−3. Particularly, Vol. 3: Martin, Y. C. Recent Progress in CoMFA Methodology and Related Techniques.
(6) Oslob, J. D.; Åkermark, B.; Helquist, P.; Norrby, P.-O. *Organometallics* **1997**, *16*, 3015.
(7) For chiral recognition in chromatography and catalysis: (a) Lipkowitz, K. B. *Acc. Chem. Res.* **2000**, *33*, 555. (b) Lipkowitz, K. B.; D'Hue, C. A.; Sakamoto, T.; Stack, J. N. *J. Am. Chem. Soc.* **2002**, 124, 14255.
(8) Kitamura, M.; Suga, S.; Kawai, K.; Noyori, R. *J. Am. Chem. Soc.* **1986**, *108*, 6071.
(9) Pu, L.; Yu, H.-B. *Chem. Rev.* **2001**, *101*, 757 and references therein.
(10) (a) Yamakawa, M.; Noyori, R. *J. Am. Chem. Soc.* **1995**, *117*, 6327. (b) Yamakawa, M.; Noyori, R. *Organometallics* **1999**, *18*, 128. (c) Goldfuss, B.; Houk, K. N. *J. Org. Chem.* **1998**, *63*, 8998. (d) Goldfuss, B.; Steigelmann, M.; Khan, S. I.; Houk, K. N. *J. Org. Chem.* **2000**, *65*, 77. (e) Vazquez, J.; Pericas, M. A.; Maseras, F.; Lledos, A. *J. Org. Chem.* **2000**, *65*, 7303. (f) Panda, M.; Phuan, P.-W.; Kozlowski, M. C. *J. Org. Chem.* **2003**, *68*, 564.
(11) Four transition structures (*anti S*, *anti R*, *syn S*, and *syn R*) are possible for each catalyst. PM3 transition-structure energies correctly predict the sign of the enantioselection in most cases, but quantitative predictions fail in some instances (ref 10f).
(12) For a complete description of this methodology, see: Dixon, S. L.; Lauri, G.; Merz, K. M., Jr. Manuscript in preparation.
(13) Dixon, S. L.; Merz, K. M., Jr. *J. Chem. Phys.* **1997**, *107*, 879.

JA0293195