Asymmetric Synthesis

DOI: 10.1002/anie.200704774

Toward a Computational Tool Predicting the Stereochemical Outcome of Asymmetric Reactions: Development and Application of a Rapid and Accurate Program Based on Organic Principles**

*Christopher R. Corbeil, Sabine Thielges, Jeremy A. Schwartzentruber, and Nicolas Moitessier**

Asymmetric catalyst discovery as currently practiced often relies on expensive, and sometimes serendipitous, stepwise optimization and/or library screening.[1] We believe that this paradigm is poised to change, as computational predictive methods have reached a level of accuracy that obviates many steps now done manually. We report herein the early version of a new program, ACE (asymmetric catalyst evaluation), its underlying concepts, and the assessment of its applicability and accuracy in distinguishing efficient asymmetric catalysts or chiral auxiliaries from inferior ones.

Although much effort has been directed toward the development of computer-aided drug design tools, there has been little investigation into computational tools for asymmetric catalyst design. Nowadays, the fields of quantum mechanics (QM) and quantum mechanics/molecular mechanics (QM/MM)[2] are highly developed and have yielded accurate predictions of asymmetric reaction stereoselectivities.[3–6] However, QM methods would require months of computation to screen a library of potential catalysts in the search for new ones. To address this issue, other methods were developed, which include reverse docking[7,8] and quantitative structure–selectivity relationships[9–11] and more specifically the use of quantum mechanics interaction fields.[12,13] As another alternative to QM techniques, molecular mechanics applied to ground-state structures have been used.[14] Advanced MM-based transition-state (TS) techniques, which accurately predict TS structures and their relative potential energies, have also been reported.[15] Although these methods (e.g., Q2MM (QM-guided MM),[16] using TSFF (transition-state force fields),[17] SEAM (seam of two potential-energy functions),[18,19] EVB (empirical valence bond),[20,21] and MCMM (multiconfiguration MM)[22]) have shown great potential in locating and investigating TSs, only a very few studies were reported that attempted to predict the stereochemical outcome of reactions,[7,8,14,23–28] with even fewer applications to the design of new asymmetric catalysts.[13,29,30] In fact, one major shortcoming of force fields is the lack of accurate parameters for metal complexes, which are necessary to model metal-catalyzed reactions and need to be specifically developed.[31]

ACE is a molecular-mechanics-based independent program that has been developed from simple organic chemistry principles. For example, the Hammond–Leffler postulate states that the TS is most similar to the species (reactants or products) which it is closest to in energy. Following this principle, ACE constructs TSs from a linear combination of reactants and products, including a factor λ describing the position of the TS on the potential-energy surface [Eq. (1),

$$\mathrm{TS} = (1-\lambda)\,\mathrm{reactant} + \lambda\,\mathrm{product} \quad (1)$$

$0 < \lambda < 1$]. A similar approach is used to locate transition states by the EVB method mentioned above, in which λ is changed stepwise from 0 to 1 to find the maximum energy corresponding to the TS. EVB has indeed been used successfully in the study of several enzymatic mechanisms.[21] Within ACE, interactions between two atoms forming a bond are described as both covalent-bond and nonbonded interactions with weights $(1-\lambda)$ and λ for each of these two types of interactions. Angles, torsion angles, and nonbonded interactions between atoms of the reacting center are also scaled by either $(1-\lambda)$ if found in the reactants or λ if found in the products. As a comparison, λ can be related to the Brønsted coefficient, which measures the role of the reacting partners in a TS.

As stated by Curtin and Hammett, stereomeric excesses can be derived from the difference in the diastereomeric TS energies, in this case the MM3* force field potential energies. This force field has already been used with the SEAM and TSFF approaches to predict TS energy differences.

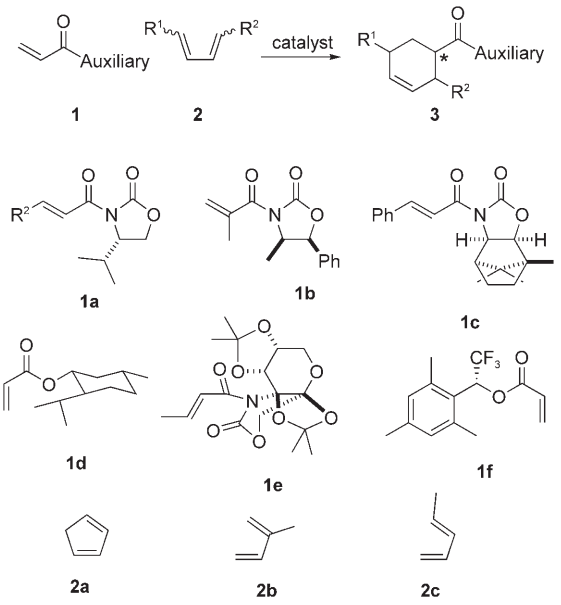
To assess the accuracy of our method, we investigated the asymmetric Diels–Alder cycloaddition using chiral auxiliaries (Scheme 1). For this purpose, 44 systems were selected from the literature involving varying dienes and dienophiles.

For each of the diene–dienophile pairs, reactants and products were built considering only an *endo* attack, known to be favored in this type of reaction. Prior to running the computation, λ must be set. It is well known that Diels–Alder reactions in the presence of strong Lewis acids have low energies of activation and early TSs, a situation which corresponds to low values of λ. In order to evaluate the impact of the selected λ value, values were used ranging from 0.10 to 0.60 in steps of 0.10.

[*] C. R. Corbeil, S. Thielges, J. A. Schwartzentruber, Prof. Dr. N. Moitessier
Department of Chemistry, McGill University
801 Sherbrooke Street W., Montréal, Québec H3A 2K6 (Canada)
Fax: (+1) 514-398-3797
E-mail: nicolas.moitessier@mcgill.ca
Homepage: moitessier-group.mcgill.ca

[**] We thank the Canadian Foundation for Innovation for financial support through the New Opportunities Fund program, NSERC, ViroChem Pharma, FQRNT, and CIHR for financial support and RQCHP for generous allocation of computer resources.

Supporting information for this article is available on the WWW under http://www.angewandte.org or from the author.



Scheme 1. General synthetic scheme and representative dienophiles **1a–f** and dienes **2a–c** used in the validation study.

ACE creates the TSs from reactants and products prepared using graphical interfaces and ESFF (extensible and systematic force field) charges[32] and carries out a conformational analysis using a genetic algorithm similar to the one previously implemented in our docking program FITTED 1.0.[33] This algorithm samples the conformational space of the transition structures. The potential energy was computed for each of the TSs, and diastereomeric excesses were derived and compared to the experimental data. Initially, the difference in potential energy between the diastereomeric TSs consistently overestimated the experimentally observed difference in free energy. A correction factor (0.5) was applied to the potential-energy difference to better align predictions with observations. Although this factor has no true physical meaning, it may reflect the difference between force-field potential energy in vacuo and experimental free energy in solvent as well as a steeper modeled potential-energy surface at the TS. Plots of $\Delta\Delta G$(predicted) versus $\Delta\Delta G$(experimental) are given in the Supporting Information. Overall, the rank-ordered list was not strongly affected by the value of λ when in the range 0.1–0.3. As the ranking is more important for virtual screening than the predicted absolute values, the selection of λ would not have much impact on the success of a screening campaign. However, increasing λ led to slightly ($\lambda = 0.4$) or significantly ($\lambda \geq 0.5$) reduced accuracies. These data demonstrate that λ does not have to be fully optimized but should be selected with care on the basis of the type of reaction analyzed. For this class of reactions, λ has to be lower than 0.5, suggesting an early TS, which is seen in DFT studies.[34] In fact, when using $\lambda = 0.1$ or $\lambda = 0.2$, the lengths of the forming/breaking bond (predicted to be in the range of 2.05–2.15 Å) match well to corresponding lengths computed using higher-level calculations (ranging from 2.05 to 2.55 Å).[34] However, the two forming bonds show the same distances with ACE, while the attack is usually asynchronous. Further development of the method is needed to account for this effect.

Applied to the entire set, ACE accurately predicted the correct isomer in 41 out of the 44 cases. The major failures (1–4 in Figure 1) were observed with polycyclic auxiliaries exemplified by **1e**. This finding suggests that the force-field description of complex molecules has to be refined.

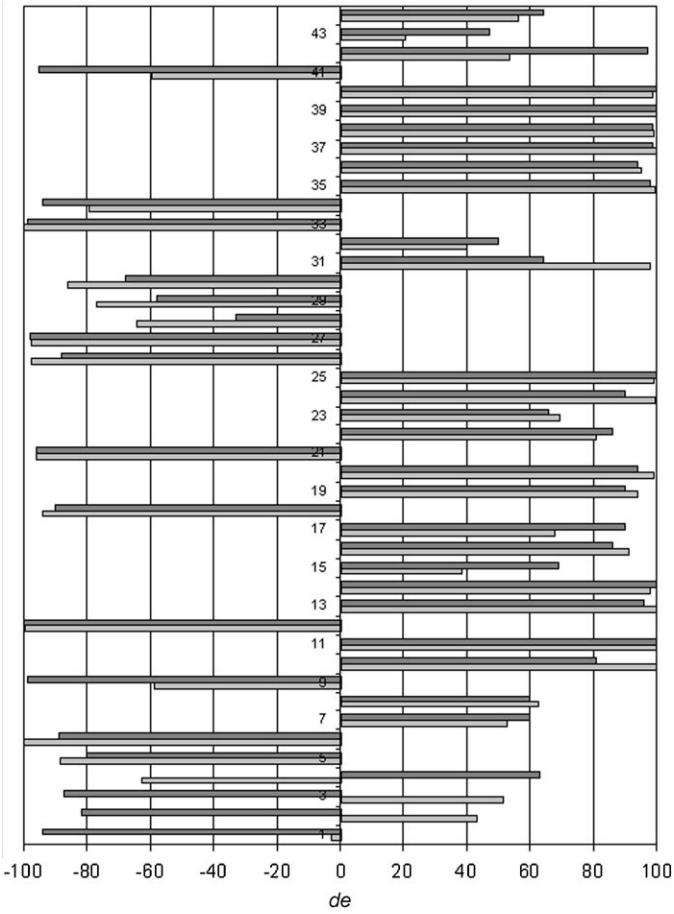
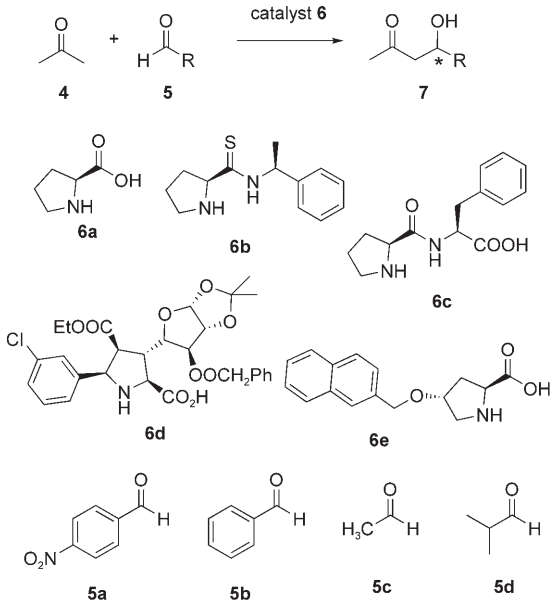


Figure 1. Predicted (light gray) versus observed (dark gray) diastereomeric excesses for 44 Diels–Alder reactions. Positive excess refers to the *R* isomer while negative excess refers to the *S* isomer ($\lambda = 0.20$).

In practice, a tool like ACE would be of interest for its ability to discriminate very good auxiliaries from a list of potential auxiliaries. The predicted 20 best of the 44 systems were first considered. Experimentally, 19 of these 20 systems led to selectivities of over 80%, with 15 over 90% and 13 over 95%. On the other hand, the ten systems that were predicted to provide the lowest selectivities were considered. Six out of these ten systems had experimentally obtained diastereomeric excesses below 70% and only one obtained an excess over 95%. These data clearly show the potential of this method to discriminate between efficient and inefficient chiral auxiliaries.

The second reaction we investigated was the asymmetric organocatalyzed aldol reaction (Scheme 2). Reported reactions using various combinations of ketones, aldehydes, and



Scheme 2. General synthetic scheme and representative catalysts (**6a–e**) and aldehydes (**5a–d**) used in the validation study.

proline derivatives used as catalysts were selected, for a total of 40 combinations.

According to extensive experimental and DFT studies, this reaction involves the formation of a flexible macrocyclic TS[35,36] and so required sampling the conformational space of large rings. The corner-flapping approach[37] was implemented in ACE to carry out this conformational search. From DFT studies, the key TS is found to be closer in energy to the produced intermediate than to the starting reactants, implying a λ value greater than 0.5.[35] Figure 2 summarizes the results obtained with $\lambda = 0.60$, although λ in the range 0.60–0.75 led to similar results. As for the Diels–Alder reaction, ACE TSs can be compared to TSs developed using higher-level calculations. Figure 3 illustrates the superposition of the most energetically favored aldol TS structures as proposed by DFT and ACE. The lengths of the forming bond predicted by these two methods are within 0.1 Å.

In the asymmetric organocatalyzed aldol reaction, ACE was again accurate, with the correct isomers predicted in 38 cases out of 40. Most of the cases investigated herein are known to provide excesses below 80%, equivalent to a small difference in energies between diastereomeric TSs. This condition makes this second validation study more challenging. Extensive investigations did not reveal the cause of these two failures. Only one of the computed reactions experimentally showed an enantiomeric excess higher than 90% (4 in Figure 2) and was indeed predicted to lead to the highest selectivity of the set (prediction: 99%). This finding demonstrates that ACE could accurately guide the design of efficient catalysts.

As another validation, it is of interest to compare high-level calculation results, when available, with these results. Houk and co-workers have reported an exhaustive study on the proline-catalyzed aldol reaction of acetone with various aldehydes, in an attempt to assess the predictive power of

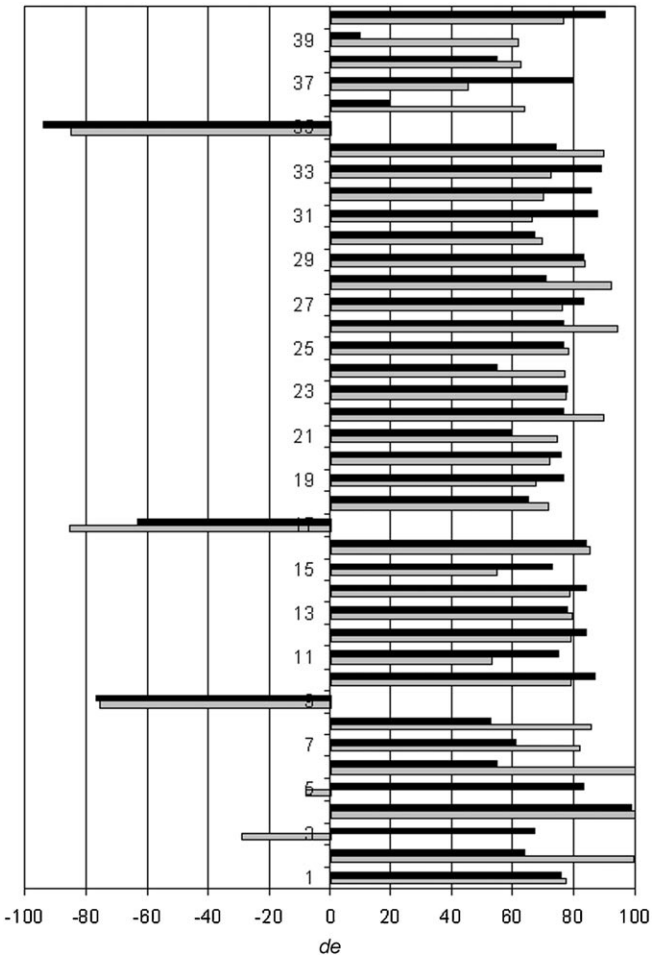


Figure 2. Predicted (gray) versus observed (black) diastereomeric excesses for 17 selected cases. Positive excess refers to the *R* isomer while negative excess refers to the *S* isomer ($\lambda = 0.6$). The complete data (40 cases) is given as Supporting Information.

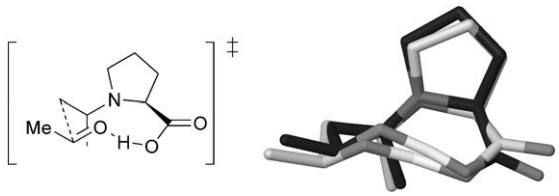


Figure 3. Predicted TS structure for the reaction involving **4**, **5c**, and **6a**. Gray: DFT prediction, black: ACE predictions.

DFT.[4] As shown in Figures 3 and 4, ACE shows accuracy close to DFT but within a much shorter period of time. This unexpectedly high accuracy might be attributable to the exhaustive conformational search of the macrocyclic TSs carried out by ACE but not by DFT techniques. In fact, ACE could be used as a conformational search engine providing high-quality starting structures for further DFT studies. Moreover, this software provides good-quality transition structures that can be used for rationalizations of data in place of CPK (Corey–Pauling–Koltun) models, as additional pairwise interaction energies can be generated as output.

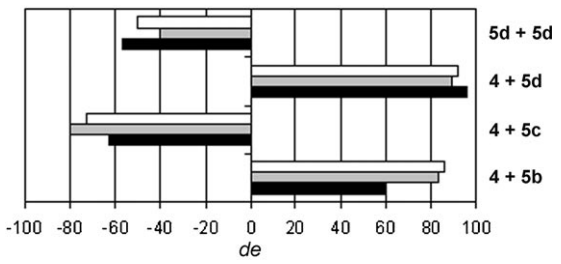


Figure 4. ACE predictions (gray) and DFT predictions (white) versus observed (black) diastereomeric excesses for four selected cases.

The trade-off between computing speed and accuracy of predictions is well known. Herein, we have presented a unique computational tool, ACE, which performs conformational sampling, TS potential-energy optimization, and TS relative-energy evaluation within less than an hour on a standard PC. Application of this tool to two well-established reactions has revealed its good accuracy in predicting enantiomeric and diastereomeric excesses. Future enhancements, applications, and validations are ongoing to improve and assess the predictive power and versatility of the software as well as its transferability to other reactions. Metal-catalyzed reactions are being investigated. However, the early version of ACE shows considerable promise and we believe should be transferable to any other reactions with well-known mechanisms.

Received: October 15, 2007
Revised: November 27, 2007
Published online: February 25, 2008

Keywords: asymmetric synthesis · computer chemistry · prediction · transition states

[1] M. B. Francis, E. N. Jacobsen, *Angew. Chem.* **1999**, *111*, 987; *Angew. Chem. Int. Ed.* **1999**, *38*, 937.
[2] H. Lin, D. G. Truhlar, *Theor. Chem. Acc.* **2007**, *117*, 185.
[3] M. Panda, P. W. Phuan, M. C. Kozlowski, *J. Org. Chem.* **2003**, *68*, 564.
[4] S. Bahmanyar, K. N. Houk, H. J. Martin, B. List, *J. Am. Chem. Soc.* **2003**, *125*, 2475.
[5] J. I. Garcia, G. Jimenez-Oses, V. Martinez-Merino, J. A. Mayoral, E. Pires, I. Villalba, *Chem. Eur. J.* **2007**, *13*, 4064.
[6] T. P. M. Goumans, A. W. Ehlers, K. Lammertsma, *Organometallics* **2005**, *24*, 3200.
[7] D. J. Harriman, G. Deslongchamps, *J. Comput.-Aided Mol. Des.* **2004**, *18*, 303.
[8] D. J. Harriman, A. Lambropoulos, G. Deslongchamps, *Tetrahedron Lett.* **2007**, *48*, 689.
[9] S. Chavali, B. Lin, D. C. Miller, K. V. Camarda, *Comput. Chem. Eng.* **2004**, *28*, 605.
[10] B. Lin, S. Chavali, K. Camarda, D. C. Miller, *Comput. Chem. Eng.* **2005**, *29*, 337.
[11] S. Sciabola, A. Alex, P. D. Higginson, J. C. Mitchell, M. J. Snowden, I. Morao, *J. Org. Chem.* **2005**, *70*, 9025.
[12] J. C. Ianni, V. Annamalai, P. W. Phuan, M. Panda, M. C. Kozlowski, *Angew. Chem.* **2006**, *118*, 5628; *Angew. Chem. Int. Ed.* **2006**, *45*, 5502.
[13] J. Huang, J. C. Ianni, J. E. Antoline, R. P. Hsung, M. C. Kozlowski, *Org. Lett.* **2006**, *8*, 1565.
[14] R. J. Deeth, N. Fey, *Organometallics* **2004**, *23*, 1042.
[15] F. Jensen, P. O. Norrby, *Theor. Chem. Acc.* **2003**, *109*, 1.
[16] P. O. Norrby, *J. Mol. Struct. THEOCHEM* **2000**, *506*, 9.
[17] J. E. Eksterowicz, K. N. Houk, *Chem. Rev.* **1993**, *93*, 2439.
[18] P. T. Olsen, F. Jensen, *J. Chem. Phys.* **2003**, *118*, 3523.
[19] F. Jensen, *J. Chem. Phys.* **2003**, *119*, 8804.
[20] A. Warshel, R. M. Weiss, *J. Am. Chem. Soc.* **1980**, *102*, 6218.
[21] J. Åqvist, A. Warshel, *Chem. Rev.* **1993**, *93*, 2523.
[22] D. G. Truhlar, *J. Comput. Chem.* **2007**, *28*, 73.
[23] N. Moitessier, F. Chrétien, Y. Chapleur, B. Maigret, *Eur. J. Org. Chem.* **2000**, 995.
[24] N. Moitessier, C. Henry, C. Len, Y. Chapleur, *J. Org. Chem.* **2002**, *67*, 7275.
[25] P. Fristrup, G. H. Jensen, M. L. N. Andersen, D. Tanner, P. O. Norrby, *J. Organomet. Chem.* **2006**, *691*, 2182.
[26] C. Gennari, E. Fioravanzo, A. Bernardi, A. Vulpetti, *Tetrahedron* **1994**, *50*, 8815.
[27] T. Rasmussen, P. O. Norrby, *J. Am. Chem. Soc.* **2003**, *125*, 5130.
[28] A. Bernardi, C. Gennari, J. M. Goodman, I. Paterson, *Tetrahedron: Asymmetry* **1995**, *6*, 2613.
[29] M. C. Kozlowski, S. P. Waters, J. W. Skudlarek, C. A. Evans, *Org. Lett.* **2002**, *4*, 4391.
[30] C. Gennari, C. T. Hewkin, F. Molinari, A. Bernardi, A. Comotti, J. M. Goodman, I. Paterson, *J. Org. Chem.* **1992**, *57*, 5173.
[31] R. J. Deeth, *Inorg. Chem.* **2007**, *46*, 4492.
[32] S. Shi, L. Yan, Y. Yang, J. Fischer-Shaulsky, T. Thacher, *J. Comput. Chem.* **2003**, *24*, 1059.
[33] C. R. Corbeil, P. Englebienne, N. Moitessier, *J. Chem. Inf. Model.* **2007**, *47*, 435.
[34] V. Branchadell, *Int. J. Quantum Chem.* **1997**, *61*, 381.
[35] K. N. Rankin, J. W. Gauld, R. J. Boyd, *J. Phys. Chem. A* **2002**, *106*, 5155.
[36] C. Allemann, R. Gordillo, F. R. Clemente, P. H. Y. Cheong, K. N. Houk, *Acc. Chem. Res.* **2004**, *37*, 558.
[37] H. Goto, E. Osawa, *J. Am. Chem. Soc.* **1989**, *111*, 8950.

