J Mol Model (2006) 12: 793–797
DOI 10.1007/s00894-006-0097-z

ORIGINAL PAPER

D. Joseph Harriman · Ghislain Deslongchamps

Reverse-docking study of the TADDOL-catalyzed asymmetric hetero-Diels–Alder reaction

Received: 16 November 2005 / Accepted: 14 December 2005 / Published online: 23 February 2006

Abstract The reverse-docking of a TADDOL catalyst to rigid transition-state (TS) representations of an asymmetric hetero-Diels–Alder reaction is described. The resulting docking poses represent a simplified geometric model of the TS for the catalyzed reaction. The conformational space of the catalyst in proximity to the catalyst-free TS models is sampled stochastically and the energetically favored poses are subjected to a clustering procedure to highlight structural attributes compatible with organocatalysis. Each pose is scored and ranked based on its molecular-mechanics docking energy. The reverse-docking procedure reveals a clear energetic trend in favor of the experimentally preferred product enantiomers. Analysis of the best poses suggests a geometric model that is consistent with principles of molecular recognition, catalysis, and experimental data.

Keywords Asymmetric synthesis · Organocatalysis · Docking · Reverse-docking · hetero-Diels–Alder

Introduction

The future of asymmetric synthesis lies in the development of reusable catalysts that can induce asymmetric reactions. Metal-free versions of these catalysts, termed organocatalysts, [1–3] can be tolerant to water and aerobic conditions, and are prime examples of "*green chemistry*". Their catalytic properties mimic those of enzymes and catalytic antibodies. The ability to generate only one enantiomer of a chiral compound is a very useful tool in the pharmaceutical industry. In 2003, Rawal established that α,α,α′,α′-tetraaryl-1,3-dioxolan-4,5-dimethanol alcohol (TADDOL) could effectively catalyze asymmetric hetero-Diels–Alder (HDA) reactions [4]. The main factor promoting such reactions appears to be H-bonding patterns. Although these H-bonds are thought to play a critical role in catalysis, a precise 3D transition-state model (TS) for this TADDOL catalyzed HDA reaction remains to be established. In previous work, reverse-docking showed promise for the development and study of asymmetric organocatalyzed reactions [5]. These studies were carried out using AutoDock 3.0.5 for conformational sampling. However, recent efforts to streamline the procedure led to the development of a new docking algorithm based on stochastic sampling and energy minimization-based local search, named EM-Dock [6]. The work presented here highlights the application of EM-Dock to the Rawal HDA system. These reverse-docking studies provide insight into the mechanism of the reaction and may allow for the rational design of more effective organocatalysts.

Rawal HDA organocatalyst H-bond promoted catalysis of the HDA reaction was first reported by Rawal in 2002 [7]. It showed that H-bonding solvents greatly accelerated the rate of the reaction presumably via H-bonding to the carbonyl group of the dienophile. In subsequent work, catalysts having H-bond donating properties were found to be very effective in non-competitive solvents, and the use of TADDOL as an asymmetric catalyst for the HDA reaction was disclosed (Fig. 1) [2].

Transition state of the hetero-Diels–Alder reaction Theoretical studies suggest that the HDA reaction is concerted [8]. For the particular diene-dienophile pairs studied here, ^{1}H NMR shows that the reaction proceeds via an initial *endo*-cycloadduct [2] (in principle, the same final dihydropyran-4-one enantiomer could have been generated via *exo*-addition from the opposite face of the diene prior to elimination of NR_2). Thus for this study, a series of 16 TS structures for the catalyst-free reaction were

D. J. Harriman · G. Deslongchamps (✉)
Department of Chemistry,
University of New Brunswick,
Fredericton, NB, E3B 5A3, Canada
e-mail: ghislain@unb.ca
Tel.: +1-506-4534795
Fax: +1-506-4534981

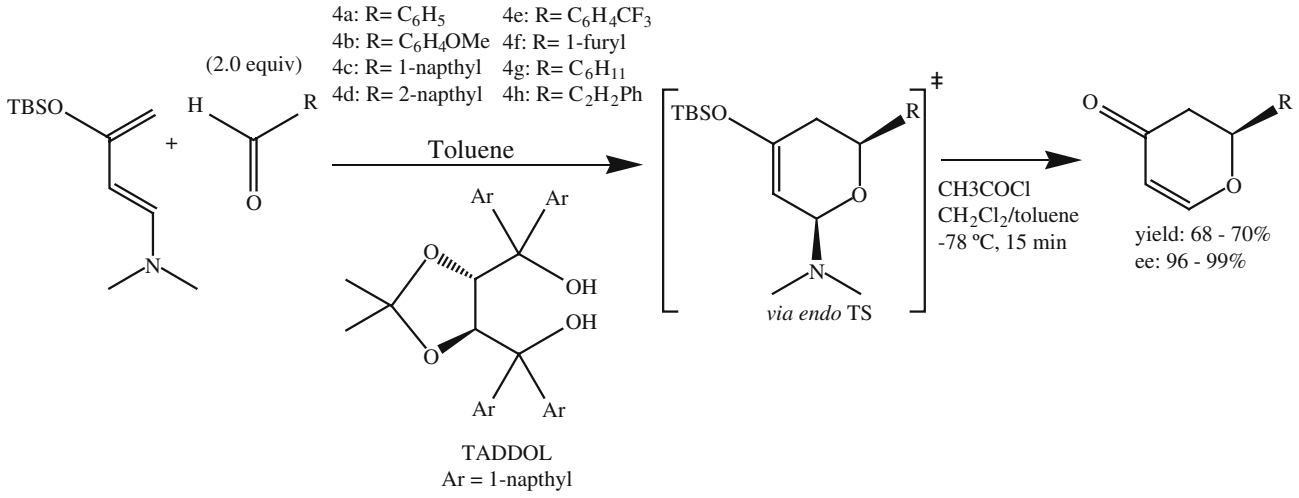

Fig. 1 TADDOL catalyzed HDA reaction

modeled, corresponding to eight enantiomer pairs of *endo*-TS for each of the substrates used in the original Rawal study (Fig. 1).

Transition state of TADDOL catalyzed reaction Rawal proposed a working model for the TADDOL catalyzed Diels–Alder (DA) reaction (Fig. 2) [9]. By analogy, this model may also be applicable to the HDA reaction, providing intramolecular H-bonding in TADDOL, H-bonding between the free hydroxyl proton of the catalyst and the carbonyl of the dienophile, and stacking of the dienophile against the pseudo-equatorial naphthyl ring of TADDOL. This would pre-organize an asymmetric environment for cycloaddition to the diene.

Reverse-docking As previously reported, a reverse-docking methodology was devised to rationalize the stereoselectivity of asymmetric reactions. While traditional docking explores the configurational space of a small molecule within the confines of a large receptor, reverse-docking explores the configurational space of a flexible organocatalyst around a catalyst-free TS representation of the asymmetric reaction (Fig. 3). Evaluating the most energetically favorable poses resulting from the reverse-docking simulation can provide insight into the observed stereoselectivity of the reaction. The interest in this computational approach is that it allows for the study of highly flexible catalysts, and to search through vast configurational spaces. Here, we report the reverse-docking of TADDOL, a less flexible catalyst, to TS models of the HDA reaction, and show that the most energetically favorable docking poses are consistent with principles of molecular recognition, catalysis, and experimental data.

Materials and methods

Transition-state modeling The catalyst-free *endo*-transition states were modeled initially using MMFF94x, a modification of the MMFF94s [10] force field available in MOE 2004.03 [11]. As an initial estimate for the TS structures, the C–C addition bonds were restrained to 2.5 Å during the energy minimization. This was done for all eight aldehyde structures reported in Fig. 1, reacting with the

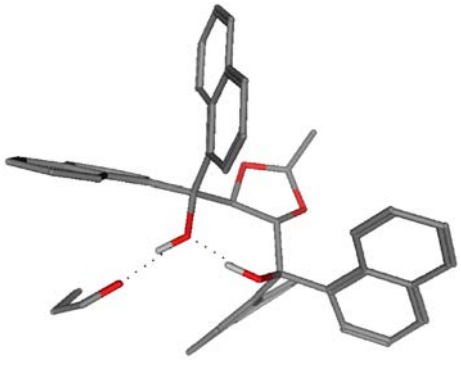

Fig. 2 Proposed TADDOL-dienophile complex for the Diels–Alder reaction (interpreted from Rawal)

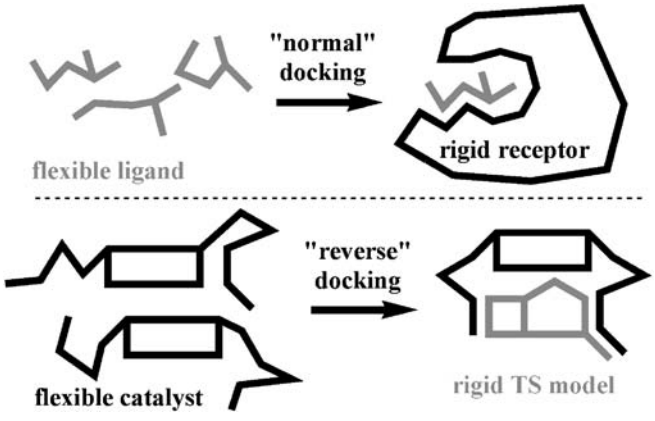

Fig. 3 Reverse-docking vs normal docking

same diene. The resulting structures were used as input for subsequent TS searches at the HF/6–31G* level with Gaussian03 [12]. Initial force constants at the starting geometry were calculated using the same basis set as the optimization. All optimized structures were subjected to a frequency calculation to determine whether they corresponded to stationary points on the potential energy surface. In addition, the number of negative eigenvalues was examined to determine if geometries corresponded to a first-order saddle point. RESP charges [13] were subsequently calculated using Antechamber in Amber7 [14] and used for subsequent docking energy calculations. Enantiomers of each TS model were also generated by simple imaging, resulting in a total of 16 catalyst-free TS models.

TADDOL catalyst modeling TADDOL was modeled in MOE using the MMFF94x forcefield. The structure was then optimized at the HF/6–31G* level and RESP charges were fitted in Antechamber.

Reverse-docking The reverse-docking simulations were carried out using EM-Dock in MOE. This new algorithm was trained on a large number of ligand/receptor complexes using RESP charges. The solvent dielectric constant was set to 2.4 to incorporate the bulk electrostatic dampening of toluene into the docking environment. The following parameters were used in the simulations: number of runs: 400; population size: 400; number of generations: 1,600. All reverse-docking simulations were carried out in triplicate and reported as rows A, B, C in Table 1. The catalyst poses in all 3×16 databases were energy-minimized (MMFF94x) with toluene solvation (ε=2.4) around the rigid TS models and scored based on internal and non-bonded interactions in MOE. This ensured that all poses obtained from the reverse-docking simulation were scored based on geometry-optimized local energy minima.

Docking energy The reverse-docking energies were calculated in MOE as the sum of the internal energy of the catalyst+intermolecular interactions [electrostatic+van der Waals]. Reverse-docking poses were sorted and ranked according to their energy.

Pose clustering Reverse-docking poses were clustered based on the distances of both hydroxyl protons of TADDOL to the carbonyl oxygen of the aldehyde dienophile. Poses with no H-bonding between catalyst and TS model (i.e. both H O distances≥4 Å) were discarded leaving only catalytically relevant poses in the databases. In addition, docking ranks determined before the clustering were recorded for each entry. Thus, the overall rank of a clustered pose, relative to all the docking poses (i.e. prior to clustering), was retained for analysis and discussion. This would serve as an indicator of whether the pose clustering excluded any energetically favorable docking poses.

Table 1 Reverse-docking results following database screening

Substrate[a]		"S" TS model		"R" TS model	
		E[b]	Rank[c]	E	Rank
4a	A	218.6	(1)	223.4	(1)
	B	217.7	(2)	222.4	(1)
	C	217.1	(1)	223.4	(1)
4b	A	220.3	(1)	226.9	(7)
	B	222.5	(1)	225.9	(1)
	C	220.2	(1)	224.5	(1)
4c	A	221.7	(1)	225.6	(1)
	B	221.2	(1)	224.2	(1)
	C	221.4	(1)	224.9	(1)
4d	A	222.8	(1)	226.1	(6)
	B	224.1	(1)	225.8	(1)
	C	215.8	(1)	226.5	(5)
4e	A	218.9	(1)	225.7	(1)
	B	220.3	(1)	225.2	(1)
	C	222.4	(1)	224.7	(1)
4f	A	225.3	(1)	227.1	(3)
	B	223.7	(1)	226.0	(1)
	C	223.0	(1)	226.9	(3)
4g	A	225.1	(1)	228.9	(2)
	B	225.3	(1)	225.3	(1)
	C	224.5	(1)	226.6	(2)
4h	A	221.4	(1)	223.0	(1)
	B	219.3	(1)	224.1	(1)
	C	219.4	(1)	224.9	(2)

[a]Substrates according to Fig. 1. Rows A–C refer to triplicate reverse-docking runs
[b]Docking energy of lowest-energy entry as calculated by MOE (kcal mol^{-1})
[c]Overall docking rank based on MOE energy prior to clustering (1=best rank)

Results

Reverse-docking energies As previously reported, the enantioselectivity of the TADDOL-catalyzed HDA reaction always favored the formation of the *S*-enantiomer product. Based on this notion, one would expect the docking poses of the catalyst around the TS model leading to the *S*-enantiomer to have lower energies than to those around the *R*-enantiomer TS model. The data shown in Table 1 indicate a clear energetic trend favoring the formation of the product *S*-enantiomers, in agreement with experiment. Indeed, for each of the 24 pairs of docking runs, the highest-ranking *S*-enantiomer model shows lower energy than that of the corresponding *R*-enantiomer model. The average docking energy of the best poses to the preferred *S*-enantiomer (221.3 kcal mol^{-1}) is ≤4.0 kcal mol^{-1} lower than that for the *R*-enantiomer (225.3 kcal mol^{-1}). Furthermore, the *S*-enantiomer poses emerge within the top two ranks prior to clustering, indicating that the H-bonding clustering scheme did not exclude relevant poses. As for the *R*-enantiomer poses, we note

much variability in the rank, suggestive of a certain degree of mismatch between the chirality of the catalyst and that of the asymmetric reaction, and that lower energy docking poses may exist but are inconsistent with a unified mode of catalysis.

Although the reverse-docking simulations were able to predict the trend of the lowest-energy poses for the experimentally favored enantiomer, they currently only offer a qualitative correlation as the energy differences between the *S* and *R* enantiomers are too large (≤4.5 kcal mol^{-1}) for the quantitative correlation of enantiomeric excess values. This is not surprising due to the simplicity of the approach, which after all, treats the catalyst-free TS model as a rigid component, uses a simple MM-based approach to compute molecular interactions, and neglects explicit solvent and entropic effects.

Discussion

3D transition state model for asymmetric organocatalysis The best ranked TS models for forming the *S*-enantiomers all share structural features similar to those present in Rawal's original Diels–Alder model. Intramolecular H-bonding within the TADDOL catalyst is present in all structures; this would likely increase the acidity of the free hydroxyl proton and allow for a stronger intermolecular H-bond with the carbonyl oxygen of the incoming dienophile. Not surprisingly due to the pose clustering, analysis of the *S*-enantiomer poses reveals an intermolecular H-bond to the dienophile carbonyl. More importantly, these poses suggest π-stacking between the dienophile and the pseudo-equatorial naphthyl ring of the catalyst. Interestingly, the highest ranking poses position the dienophile's aldehyde hydrogen near one of the naphthyl rings of TADDOL, possibly suggesting some electrostatic interaction. Indeed, H-bonding of the aldehyde to TADDOL's free hydroxyl proton should increase slightly the acidity of the aldehyde proton itself. Figure 4a shows the highest-ranking pose of the TADDOL catalyst around the rigid *S*-enantiomer TS model 4a.

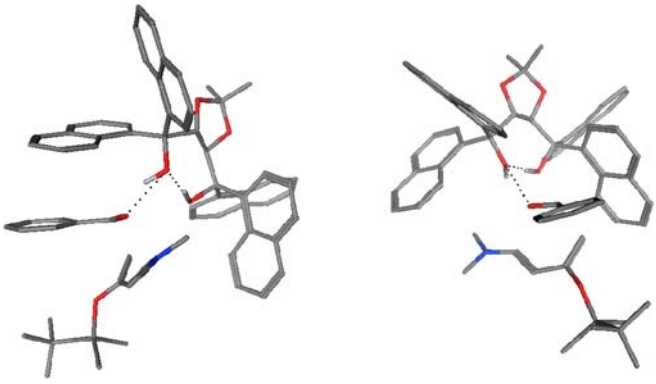

Fig. 4 Best reverse-docking poses of the TADDOL catalyst to the TS model for substrate 4a. **a** *S*-enantiomeric TS model (*E*=217.1 kcal mol^{-1}, rank=1); **b** *R*-enantiomeric TS model (*E*=223.4 kcal mol^{-1}, rank=1). Non-polar hydrogens omitted for clarity

The docking poses for the "*R*" TS model also display inter- and intramolecular H-bonding patterns but, in contrast to the "*S*" docking results, do not show a consistent stabilization pattern of π-stacking of the dienophile to the naphthyl rings. This is illustrated in Fig. 4b, where the energy of this best pose to the *R*-enantiomer TS model is significantly higher (6 kcal mol^{-1}) than that of the corresponding *S*-enantiomer pose. This difference in docking energy could be attributed to the lack of π-stacking in the *R*-enantiomeric pose. Based on these observations, it appears that the catalyst is able to preferentially bind to *S*-enantiomer TS poses based on intra- and intermolecular H-bonds and the preorganization of the dienophile based on stacking interactions. This serves to create an asymmetric environment conducive to preferential formation of the *S*-enantiomer.

In summary, the TS for the TADDOL-catalyzed asymmetric HDA reaction was studied using the reverse-docking technique. With EM-Dock, the configurational space of TADDOL around rigid catalyst-free TS models of the hetero-Diels–Alder reaction was explored stochastically. The resulting lowest-energy docking poses were scored and ranked based on MM-derived energies. Analysis revealed lowest-energy docking scores with the TS models leading to formation of the *S*-enantiomer product, in complete agreement with the available experimental data for this reaction. Structural analysis of the docking poses to the *S*-enantiomer TS models suggests a mechanism for enantioselective organocatalysis which is consistently in agreement with Rawal's proposed TS model for the TADDOL-catalyzed Diels–Alder reaction. The model comprises intramolecular H-bonding in TADDOL, intermolecular H-bonding to the dienophile carbonyl, and stacking of the dienophile against one of the pseudo-equatorial rings of TADDOL, exposing only one aldehyde face for the ensuing hetero-Diels–Alder reaction, all in agreement with principles of molecular recognition and experimental data. In contrast, results with the *R*-enantiomer TS models do not show a consistent stabilization pattern of stacking of the dienophile to the naphthyl ring and do not suggest any consistent preorganization pattern of the dienophile to the catalyst. To date, results give only a qualitative prediction of enantioselectivity since the reverse-docking process is based on a standard molecular-mechanics force fields. High-level ab initio calculations of the best transition-state models of the entire catalyst–substrate complexes may lead to more accurate predictions of enantioselectivities. Nevertheless, the reverse-docking approach is appropriate for studying larger, more flexible organocatalysts, with vast numbers of potential binding configurations at the transition state of a selected asymmetric reaction (a similar analysis using ab initio methods would be prohibitive). Ultimately, the TS models derived

from these studies may serve as templates for the rational in silico design of new catalysts, and other molecular devices. Current work is focused on the development of TS models for other organocatalyzed reactions, including Strecker, Mannich and other Diels–Alder reactions [7]. A high-throughput method for the virtual screening of large organocatalyst databases is also being developed. We will report on these advances in due course.

Acknowledgements Financial support from the Natural Sciences and Engineering Research Council, the New Brunswick Innovation Foundation, and the University of New Brunswick is gratefully acknowledged. We are also grateful to the Chemical Computing Group Inc. for software licenses and SVL support.

References

1. Houk KN, List B (2004) Acc Chem Res 37:487–631
2. Schreiner PR (2003) Chem Soc Rev 32:289–296
3. Pihko PM (2004) Angew Chem Int Ed Engl 43:2062–2064
4. Huang Y, Unni AK, Thadani AN, Rawal VH (2003) Nature 424:146
5. Harriman DJ, Deslongchamps GD (2004) J Comp Aided-Mol Des 18:303–308
6. Wiley EA, MacDonald M, Lambropoulos A, Harriman DJ, Deslongchamps G Can J Chem, (in press)
7. Huang Y, Rawal VH (2002) J Amer Chem Soc 124:9662–9663
8. Houk KN, González J, Li Y (1995) Acc Chem Res 28:81–90
9. Thadani AN, Stankovic AR, Rawal VH (2004) PNAS 101:5846–5850
10. Halgren TA (1999) J Comp Chem 20:720–729
11. Molecular Operating Environment, version 2004.03 (2004) Chemical Computing Group Inc, Montreal, Canada
12. Frisch MJ, Trucks GW, Schlegel HB, Scuseria GE, Robb MA, Cheeseman JR, Montgomery JA Jr, Vreven T, Kudin KN, Burant JC, Millam JM, Iyengar SS, Tomasi J, Barone V, Mennucci B, Cossi M, Scalmani G, Rega N, Petersson GA, Nakatsuji H, Hada M, Ehara M, Toyota K, Fukuda R, Hasegawa J, Ishida M, Nakajima T, Honda Y, Kitao O, Nakai H, Klene M, Li X, Knox JE, Hratchian HP, Cross JB, Adamo C, Jaramillo J, Gomperts R, Stratmann RE, Yazyev O, Austin AJ, Cammi R, Pomelli C, Ochterski JW, Ayala PY, Morokuma K, Voth GA, Salvador P, Dannenberg JJ, Zakrzewski VG, Dapprich S, Daniels AD, Strain MC, Farkas O, Malick DK, Rabuck AD, Raghavachari K, Foresman JB, Ortiz JV, Cui Q, Baboul AG, Clifford S, Cioslowski J, Stefanov BB, Liu G, Liashenko A, Piskorz P, Komaromi I, Martin RL, Fox DJ, Keith T, Al-Laham MA, Peng CY, Nanayakkara A, Challacombe M, Gill PMW, Johnson B, Chen W, Wong MW, Gonzalez C, Pople JA (2004) Gaussian 03. Gaussian Inc, Wallingford, Connecticut
13. Bayly CI, Cieplak P, Cornell WD, Kollman PA (1993) J Phys Chem 97:10269–10280
14. Case DA, Pearlman DA, Caldwell JW, Cheatham III TE, Wang J, Ross WS, Simmerling CL, Darden TA, Merz KM, Stanton RV, Cheng AL, Vincent JJ, Crowley M, Tsui V, Gohlke H, Radmer RJ, Duan Y, Pitera J, Massova I, Seibel GL, Singh UC, Weiner PK, Kollman PA (2002) Amber 7. University of California, San Francisco