JCTC Journal of Chemical Theory and Computation

Letter

pubs.acs.org/JCTC

Sampling of Transition States for Predicting Diastereoselectivity Using Automated Search Method—Aqueous Lanthanide-Catalyzed Mukaiyama Aldol Reaction

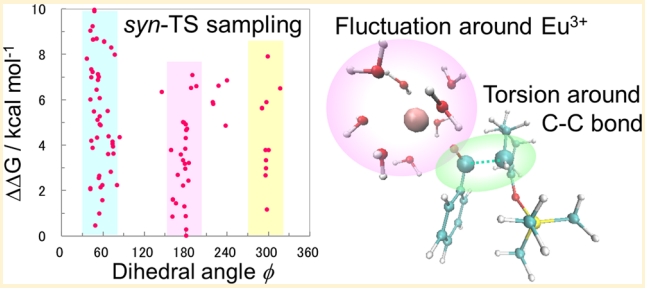
Miho Hatanaka,[†] Satoshi Maeda,[‡] and Keiji Morokuma[†,*]

[†]Fukui Institute for Fundamental Chemistry, Kyoto University, Kyoto 606-8103, Japan

[‡]Department of Chemistry, Faculty of Science, Hokkaido University, Sapporo, 060-0810, Japan

Supporting Information

ABSTRACT: To predict the stereoselectivity of large and flexible reaction systems, structural sampling of many transition states (TSs) is required. We used an automated search method, the artificial force induced reaction (AFIR) method, for TS sampling and found 91 *syn*- and 73 *anti*-TSs for the diastereoselective C−C bond formation step of the aqueous lanthanide-catalyzed Mukaiyama aldol reaction. Among them 11 *syn*- and six *anti*-TSs are found to contribute significantly to the diastereomeric ratio at room temperature.



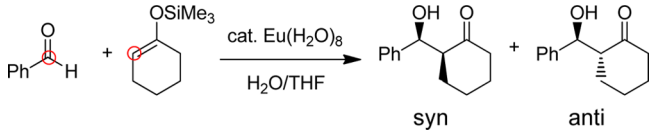
1. INTRODUCTION

The prediction of stereoselectivity is one of the most important topics for theoretical chemistry. In the case of reactions of rigid compounds, stereoselectivity can be discussed by comparing each transition state (TS) producing different stereoproducts.[1,2] However, in the case of large and flexible molecules, because of structural fluctuation, there can be a large number of TSs for a reaction step producing the same stereoproduct. Therefore, appropriate sampling of TSs is needed to predict the stereoselectivity of such a reaction. There are only a few theoretical studies considering a number of TSs for the reaction step producing the same stereoproduct.[3−7] For example, Maeda and Ohno found 52 TSs producing the same stereoproduct for Ru-BINAP catalyzed asymmetric hydrogenation by using the anharmonic downward distortion following (ADDF) method but initially fixing the active site structures.[3] Concerning the Mukaiyama aldol reaction catalyzed by boron trihalide (Lewis acid), several energetically nearly degenerate TSs producing the same stereoproduct were obtained by scanning the dihedral angle around the reactive C−C bond.[7]

The strategy of the present study for sampling TSs is to use a newly developed automated exploration method called the artificial force induced reaction (AFIR) method.[8−10] The basic idea of AFIR is to push reactants together by using the AFIR function, which increases in proportion to the distance between the reactants. By adding an appropriate artificial force, the potential energy surface along the reaction coordinate becomes attractive or monotonically decreases without a barrier. Approximate product and TS structures can be obtained very efficiently by a conventional minimization algorithm on this artificial surface. The real product and TS are then optimized easily from these approximate structures. By taking all possible orientations and approach directions of the reacting molecules, all possible products and TSs in principle can be obtained. The advantage of this method is that no guess of the product or TS is required, and reaction pathways can be found without prejudice.

In this study, we apply the AFIR method to the C−C bond formation step of the Mukaiyama aldol reaction catalyzed by a water-tolerant lanthanide Lewis acid, using the model shown in Scheme 1.[11−13] This step determines the diastereoselectivity

Scheme 1. Lanthanide Triflate-Catalyzed Mukaiyama Aldol Reaction



(*syn*/*anti* selectivity) of the entire reaction. According to experimental studies of the luminescence decay,[14,15] in water solution around eight water molecules coordinate to Eu^{3+}, replacing OTf ligands, and the hydration structure around Eu^{3+} fluctuates. This fluctuation comes from the ionic bonding character of the lanthanide trication. As is well-known, a lanthanide trication does not make covalent bonds with organic ligands because its open-shell 4f electrons are shielded by the closed-shell 5s and 5p electrons from outside.[16−18] Covalent bond character caused by ligand to metal charge transfer, especially to the 5d orbital, was reported for some lanthanide compounds.[17,18] However, virtual 5d orbitals typically do not have a sufficient contribution to create significant covalency.[16]

Received: March 31, 2013

   dx.doi.org/10.1021/ct4002637 | *J. Chem. Theory Comput.* XXXX, XXX, XXX−XXX

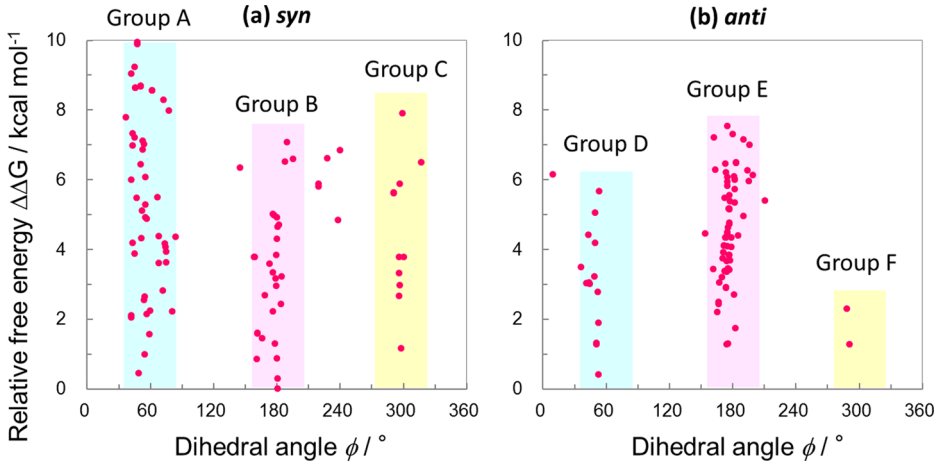


Figure 1. The relative Gibbs free energy $\Delta\Delta G$ (in kcal/mol) and dihedral angle ϕ (in deg) of TSs for the C−C bond forming step. TSs producing *syn*- and *anti*-products are shown in a and b, respectively. Those included in blue, pink, and yellow bars are categorized as groups A, B, and C for *syn*- and D, E, and F for *anti*-products, respectively. The Gibbs free energy barrier of the lowest *syn*-TS is 5.5 kcal/mol, which is used as a reference of relative free energies of other TSs.

Since the present system has very large geometry fluctuation caused by Eu^{3+}, many TSs are expected for the same C−C bond formation step, and a method for searching all possible TSs is required. The AFIR method seems to be ideally suited for the determination of many possible stereoselective TSs of complex and flexible organic reactions and is adopted for a benchmark study of the reaction.

2. COMPUTATIONAL DETAILS

The procedure of AFIR calculation starts from exploring approximate TS and product structures. In a preliminary AFIR search, we put an artificial force between a reactant complex (Cmp1), Eu^{3+} coordinated by eight water molecules (experimentally suggested as discussed above) and a benzaldehyde, and another reactant (Cmp2), trimethyl silyl (TMS) enol ether. In this search, we noticed that the first step of this reaction is the C−C bond formation, and other reactions do not take place simultaneously. Therefore, in order to be more efficient, in the present study we put artificial force only between two reacting carbon atoms, the carbon atom of the aldehyde and the α carbon of the enol ether, as shown by red circles in Scheme 1. AFIR functions were minimized with γ = 50 kcal/mol.[8−10] Thus, this search is expected to find all the reaction pathways with barriers lower than ∼50 kcal/mol. Initial structures of the AFIR optimization were made using the following procedure. First, the relative orientation between Cmp1 and Cmp2, whose geometries are optimized separately, is determined by using random numbers. Second, the coordinates of two reactive carbon atoms are located at the same origin. Then, only Cmp2 is translated to a random direction until the distance between closest atoms in Cmp1 and Cmp2 becomes more than the summation of their covalent radii plus 0.7 Å. This procedure is the same as the general procedure mentioned in ref 9. This is repeated until 10 successive AFIR optimizations gave no new TS. In this case, 437 initial structures were examined. This procedure is intended to cover all the possible initial relative orientations and approach directions of Cmp1 and Cmp2.

At this stage, the B3LYP method with small basis sets, which is (7s6p5d)/[2s1p1d] on Eu^{3+} and 6-31G for others, was used. For Eu^{3+}, the large core relativistic effective core potential (RECP) of the Stuttgart−Dresden group was used throughout,[19,20] which explicitly considers the 5s, 5p, 5d, and 6s electrons as the valence shell and puts six 4f electrons in the core potential because 4f electrons are considered not to participate in bonding with ligands.[16−18] The effect of solvation was also considered by the polarized continuum model (PCM)[21] with a dielectric constant of 78.4 (water).

All of the approximate TSs obtained by AFIR were reoptimized (using a standard optimizer) with the dispersion-corrected B3LYP-D[22] with the (7s6p5d)/[5s4p3d] RECP basis set for Eu^{3+} and 6-31+G* for others. After geometry optimizations, to improve energetics, single point calculations were performed at the dispersion-corrected B3LYP-D3[23] level with the (8s7p6d)/[6s5p5d] RECP basis set augmented with f- and g-polarization functions for Eu^{3+} and cc-pVTZ for others. AFIR optimizations and subsequent geometry optimizations of TSs were performed using a developmental version of the GRRM program[24] using energies and gradient vectors computed by the Gaussian 09 program.[25] The dispersion correction "D3" was calculated with the Gaussian developmental version.

3. RESULTS AND DISCUSSION

Starting from 437 randomly selected initial structures with different orientations and approach directions of Cmp2, the AFIR optimization evaluated the gradient 89 775 times and the Hessian 2081 times and resulted in 91 and 74 unique TSs that produce *syn*- and *anti*-products, respectively. Figure 1 shows the distribution of the relative Gibbs free energy and the C−C−C−O dihedral angle ϕ defined in Figure 2 of all the calculated TSs for the C−C bond formation. The Gibbs energy is calculated relative to the most stable TS producing *syn*-product in group B, which has a Gibbs free energy barrier of 5.5 kcal/mol relative to the most stable reactant complex.

Most TSs can be categorized into six groups, A, B, and C for *syn*-structures and D, E, and F for *anti*-structures based on the C−C−C−O dihedral angle ϕ because they are localized only around 60°, 180°, and 300°, respectively. These three dihedral angles show that all the TSs in these groups have the "staggered" conformations, which reduce strain in the molecules. Several TSs not included in these groups have eclipsed conformations, whose dihedral angles are around 0°, 120°, and 240°. However, their energies are 5−8 kcal/mol higher than that of the lowest TS, and so their existence

dx.doi.org/10.1021/ct4002637 | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

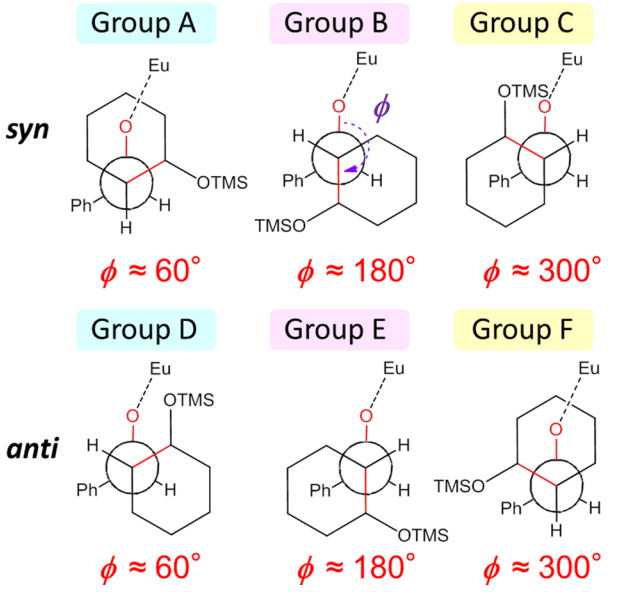


Figure 2. Definition of dihedral angle ϕ and Newman projections of TSs in groups A, B, and C for *syn*- and D, E, and F for *anti*-products, respectively.

probabilities are less than 0.02% at room temperature. It should be noted that any TS with both oxygen atoms in aldehyde and enol ether coordinated to Eu^{3+} was not obtained. Such a TS, called cyclic TS,[26−29] was used to explain the *anti*-preference for the Mukaiyama aldol reaction catalyzed by other Lewis acids, such as $TiCl_4$ and BCl_3. Several TSs where aldehyde coordinates to the hydrogen of coordination water instead of Eu^{3+} were obtained. However, their energies are about 12 kcal/mol higher than the lowest TS.

Next, we focus on the lowest TSs producing *syn*- and *anti*-products. Among all the TSs, the lowest *syn*- and *anti*-TSs are in the groups B and D, respectively. Comparing these two TSs, the potential energy of the *syn*-TS is 1.1 kcal/mol higher than that of the *anti*-TS. The *anti*-TS structure has hydrogen bonding between one of the coordinating water molecules and the oxygen in the enol ether group and is stabilized. However, this hydrogen bond makes the reactive C−C bond shorter and the structure closer to that of the product (shown in TSs 1 and 3 in Table 1). Because of the lesser flexibility of the structure, the entropic contribution $-T\Delta S$ of the *anti*-TS is 1.5 kcal/mol more positive than that of the *syn*-TS. As a result, the free energy of the *syn*-TS becomes 0.4 kcal/mol lower than the *anti*-TS.

Comparing different groups, there are one to several TSs in each group whose free energies are less than 2 kcal/mol higher from the lowest TS. This wide distribution of conformations (syn/anti and ϕ) is due to small steric repulsion between two substrates. One notices that groups C and E lack in low energy TS structures with $\Delta\Delta G$ less than 1 kcal/mol. This is due to a modest steric repulsion between the phenyl group in aldehyde and the cyclohexene group in enol ether. It should be noted that only two TSs are found in group F, although many more TSs are obtained in other groups. This is caused by a larger steric repulsion between TMS and the phenyl part and between cyclohexene and coordinated water molecules. Even when the initial dihedral angle ϕ was about 300°, Cmp2 rotates, and the resultant TSs end up in group D ($\phi \sim 60°$) or E (~180°) to reduce the strain.

Table 1. Relative Free Energies, Key Structural Parameters, and Existence Probability of Lower TSs, with $\Delta\Delta G$ Less than 2.00 kcal/mol

TS	$\Delta\Delta G$ (kcal/mol)	ϕ (deg)	group	product	C−C distance (Å)	existence probability (%)
1	0.00	180.6	B	*syn*	2.09	20.85
2	0.28	180.6	B	*syn*	2.06	12.94
3	0.40	53.1	D	*anti*	1.97	10.59
4	0.45	49.0	A	*syn*	1.99	9.73
5	0.85	161.3	B	*syn*	2.07	4.96
6	0.86	180.2	B	*syn*	2.10	4.89
7	0.97	55.0	A	*syn*	2.11	4.02
8	1.16	298.2	C	*syn*	1.95	2.92
9	1.27	51.2	D	*anti*	1.99	2.46
10	1.28	174.5	E	*anti*	2.06	2.41
11	1.28	291.1	F	*anti*	2.07	2.41
12	1.28	176.0	E	*anti*	1.99	2.38
13	1.30	178.1	B	*syn*	2.06	2.32
14	1.44	166.4	B	*syn*	2.09	1.84
15	1.56	59.2	A	*syn*	2.16	1.50
16	1.58	161.6	B	*syn*	2.06	1.44
17	1.73	182.9	E	*anti*	2.06	1.12

We next discuss the reason for different energies of TSs in the same group. As you can see in Figure 1, the energetics of TSs in the same group show a wide distribution even though their key structural parameters, such as dihedral angles ϕ, are similar. It is mainly caused by the structural fluctuation of the coordination water molecules around Eu^{3+}. The structures of four lower TSs in groups B (*syn*) and D (*anti*) are shown in Figure 3. The structures of the aldehyde part almost perfectly

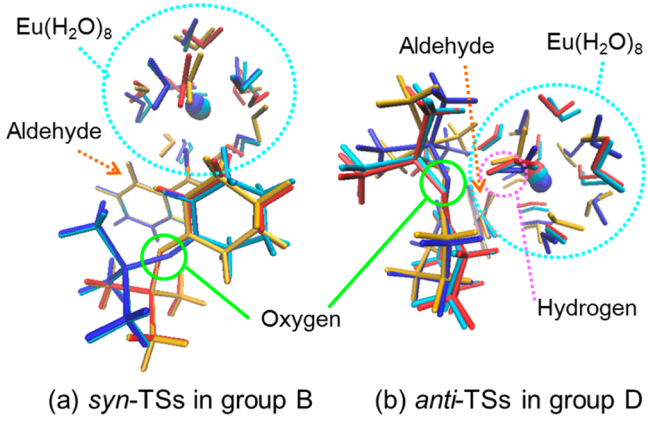


Figure 3. Comparison of the four lower *syn*-TSs in group B (a) and *anti*-TSs in group D (b). The lowest, second, third, and fourth lower TSs are in blue, red, light blue, and yellow. The oxygen atom in the enol ether and $Eu(H_2O)_8$ part are shown in green and light blue circles, respectively. The hydrogen atom in the pink circle in *anti*-TSs forms H-bonds with the oxygen in the enol ether.

overlap, and those of the enol ether part are slightly different because of the different dihedral angle ϕ and the rotation of C−O bonds in enol ether (this oxygen atom shown in green circle in Figure 3). Despite these similarities of aldehyde as well as the enol ether part, the structures of the water coordination shell are substantially different for different TSs, as shown in the light blue circle in Figure 3. This fluctuation of coordinating water structures, coming from the ionic bond character between Eu^{3+}

dx.doi.org/10.1021/ct4002637 | J. Chem. Theory Comput. XXXX, XXX, XXX−XXX

and organic ligands, gives rise to the energy difference of TSs. The present AFIR search successfully sampled various conformations of water molecules and the TMS group as seen in Figure 3. This would sound strange as AFIR does not explicitly consider conformational rearrangements. Nevertheless, the bias forcing generation of the C−C bond effectively induced conformational changes depending on initial geometries of the AFIR calculations. This is not surprising because the bias imposed by the AFIR function with γ = 50 kcal/mol is large enough to induce such low barrier rearrangements.

Finally, we calculate the diastereomeric ratio by using obtained samples of TSs. Table 1 shows the relative energies, key structural parameters, and the existence probabilities at 298 K calculated by the Boltzmann distribution of lower TSs. The sum of existence probabilities of the most stable *syn* and *anti*-TSs is only 31%, which means it is not enough to consider the most stable TSs producing *syn*- and *anti*-products for calculation of the diastereomeric ratio. Actually, the diastereomeric ratio (*syn*/*anti*) calculated by the existence probabilities of all the calculated TSs is 75:25%, while that considering only the most stable *syn*- and *anti*-TSs is 66:34%. The former reproduces the experimental ratio of 73:27%[13] quantitatively. Because the existing probabilities of TSs whose relative energies are more than 2 kcal/mol are less than 1%, the diastereomeric ratio is determined by about 11 *syn*- and six *anti*-lowest TSs, and the other 148 TSs can be ignored at room temperature.

4. CONCLUSIONS

We have demonstrated that the AFIR method can efficiently sample all the TSs that may contribute to the diastereoselectivity of a complex and flexible reaction system. In the past, in most cases, only some preselected TSs have been determined, and many other TSs that could be important have been neglected. In the present AFIR approach, no assumption on reaction pathways and TS structures is needed, and all important TSs can be determined without bias.

In the present example of the water-tolerant lanthanide Lewis-acid-catalyzed Mukaiyama aldol reaction, we found that as many as 17 TSs contribute to the diastereomeric ratio at room temperature. The diastereoselectivity is low because many different TSs giving either *syn*- or *anti*-products exist within the 2 kcal/mol range of the lowest TS. The obvious reason for the low diastereoselectivity of this reaction is the small steric repulsion between the two substrates; there is too much freedom for TSs to be stereoselective. Definitely stronger steric constraints are necessary. An additional complication comes from the structural fluctuation of coordination water molecules around Eu^{3+}. Lanthanide cations have large atomic radii and allow many water molecules to coordinate, giving rise to many possible coordination structures. However, the structural fluctuation due to water coordination can be controlled by more rigid TS structures. Knowing all the TSs and finding the reason for poor selectivity would lead rational design of efficient catalysts. Such a study in lanthanide-catalyzed Mukaiyama aldol reactions is in progress and will be reported in due course.

■ ASSOCIATED CONTENT

Supporting Information

The relative free energies; key structural parameters of all the TSs; and Cartesian coordinates of Cmp1, Cmp2, and the lowest TSs producing *syn* and *anti*-products. This information is available free of charge via the Internet at http://pubs.acs.org.

■ AUTHOR INFORMATION

Corresponding Author

*E-mail: morokuma@fukui.kyoto-u.ac.jp.

Notes

The authors declare no competing financial interest.

■ ACKNOWLEDGMENTS

This work was in part supported by a CREST (Core Research for Evolutional Science and Technology) grant in the Area of High Performance Computing for Multiscale and Multiphysics Phenomena from the Japanese Science and Technology Agency (JST) and in part by a grant from the Japan Society for the Promotion of Science (Grants-in-Aid for Scientific Research <KAKENHI> No. 24245005) at Kyoto University. M.H. acknowledges the Fukui Fellowship of Kyoto University. The Computer resources at the Academic Center for Computing and Media Studies (ACCMS) at Kyoto University and Research Center of Computer Science (RCCS) at the Institute for Molecular Science are also acknowledged.

■ REFERENCES

(1) Houk, K. N.; Cheong, P. H.-Y. *Nature* **2008**, *455*, 309−313.
(2) Balcells, D.; Maseras, F. *New J. Chem.* **2007**, *31*, 333−343.
(3) Maeda, S.; Ohno, K. *J. Am. Chem. Soc.* **2008**, *130*, 17228−17229.
(4) Donoghue, P. J.; Helquist, P.; Norrby, P.-O.; Wiest, O. *J. Am. Chem. Soc.* **2009**, *131*, 410−411.
(5) Jensen, F.; Norrby, P.-O. *Theor. Chem. Acc.* **2003**, *109*, 1.
(6) Weill, N.; Corbeil, C. R.; Schutter, J. W. D.; Moitessier, N. *J. Comput. Chem.* **2011**, *32*, 2878−2889.
(7) Lee, J. M.; Helquist, P.; Wiest, O. *J. Am. Chem. Soc.* **2012**, *134*, 14973−14981.
(8) Maeda, S.; Morokuma, K. *J. Chem. Phys.* **2010**, *132*, 241102−241106.
(9) Maeda, S.; Morokuma, K. *J. Chem. Theory Comput.* **2011**, *7*, 2335−2345.
(10) Maeda, S.; Ohno, K.; Morokuma, K. *Phys. Chem. Chem. Phys.* **2013**, *15*, 3683−3701.
(11) Kobayashi, S.; Sugiura, M.; Kitagawa, H.; Lam, W. W. L. *Chem. Rev.* **2002**, *102*, 2227−2302.
(12) Kobayashi, S.; Nagayama, S.; Busujima, T. *J. Am. Chem. Soc.* **1998**, *120*, 8287−8288.
(13) Kobayashi, S. *Synlett* **1994**, 689−701.
(14) Dissanayake, P.; Allen, M. J. *J. Am. Chem. Soc.* **2009**, *131*, 6342−6343.
(15) Averill, D. J.; Dissanayake, P.; Allen, M. J. *Molecules* **2012**, *17*, 2073−2081.
(16) Maron, L.; Eisenstein, O. *J. Phys. Chem. A* **2000**, *104*, 7140−7143.
(17) Clark, D. L.; Gordon, J. C.; Hay, P. J.; Poli, R. *Organometallics* **2005**, *24*, 5747−5758.
(18) Krogh-Jespersen, K.; Romanelli, M. D.; Melman, J. H.; Emge, T. J.; Brennan, J. G. *Inorg. Chem.* **2010**, *49*, 552−560.
(19) Dolg, M.; Stoll, H.; Savin, A.; Preuss, H. *Theor. Chim. Acta* **1989**, *75*, 173−194.
(20) Dolg, M.; Stoll, H.; Preuss, H. *Theor. Chim. Acta* **1993**, *85*, 441−450.
(21) Tomasi, J.; Mennucci, B.; Cammi, R. *Chem. Rev.* **2005**, *105*, 2999−3093.
(22) Grimme, S. *J. Comput. Chem.* **2004**, *25*, 1463−1473.
(23) Grimme, S.; Antony, J.; Ehrlich, S.; Krieg, H. *J. Chem. Phys.* **2010**, *132*, 154104−154119.
(24) Maeda, S.; Osada, Y.; Morokuma, K.; Ohno, K. *GRRM*, developmental version; Kyoto University: Kyoto, Japan.
(25) Frisch, M. J.; Trucks, G. W.; Schlegel, H. B.; Scuseria, G. E.; Robb, M. A.; Cheeseman, J. R.; Scalmani, G.; Barone, V.; Mennucci, B.; Petersson, G. A.; Nakatsuji, H.; Caricato, M.; Li, X.; Hratchian, H.

P.; Izmaylov, A. F.; Bloino, J.; Zheng, G.; Sonnenberg, J. L.; Hada, M.; Ehara, M.; Toyota, K.; Fukuda, R.; Hasegawa, J.; Ishida, M.; Nakajima, T.; Honda, Y.; Kitao, O.; Nakai, H.; Vreven, T.; Montgomery, J. A., Jr.; Peralta, J. E.; Ogliaro, F.; Bearpark, M.; Heyd, J. J.; Brothers, E.; Kudin, K. N.; Staroverov, V. N.; Kobayashi, R.; Normand, J.; Raghavachari, K.; Rendell, A.; Burant, J. C.; Iyengar, S. S.; Tomasi, J.; Cossi, M.; Rega, N.; Millam, J. M.; Klene, M.; Knox, J. E.; Cross, J. B.; Bakken, V.; Adamo, C.; Jaramillo, J.; Gomperts, R.; Stratmann, R. E.; Yazyev, O.; Austin, A. J.; Cammi, R.; Pomelli, C.; Ochterski, J. W.; Martin, R. L.; Morokuma, K.; Zakrzewski, V. G.; Voth, G. A.; Salvador, P.; Dannenberg, J. J.; Dapprich, S.; Daniels, A. D.; Farkas, O.; Foresman, J. B.; Ortiz, J. V.; Cioslowski, J.; Fox, D. J. *Gaussian 09*, revision A.2; Gaussian, Inc.: Wallingford, CT, 2009.

(26) Mukaiyama, T.; Kobayashi, S.; Uchiro, H.; Shiina, I. *Chem. Lett.* **1990**, 129−132.

(27) Yamamoto, Y.; Maruyama, K. *J. Am. Chem. Soc.* **1983**, *105*, 6963−6965.

(28) Yamamoto, Y.; Maruyama, K. *J. Am. Chem. Soc.* **1982**, *104*, 2323−2325.

(29) Mukaiyama, T.; Banno, K.; Narasaka, K. *J. Am. Chem. Soc.* **1974**, *96*, 7503−7509.