Conformational dynamics of an alanine dipeptide analog: An *ab initio* molecular dynamics study

Dongqing Wei,[1] Hong Guo,[1,†] and D. R. Salahub[1,2,*]
[1]*Centre de Recherche en Calcul Appliqué, 5160 Boulevard Décarie, Bureau 400, Montréal, Québec, Canada H3X 2H9*
[2]*Département de Chimie, Université de Montréal, Caisse Postale 6128, Succursale Centre-ville, Montréal, Québec, Canada H3C 3J7*
(Received 8 October 1999; revised manuscript received 26 January 2001; published 18 June 2001)

An *ab initio* molecular dynamics (MD) simulation technique employing the Born-Oppenheimer approach in the framework of a Gaussian implementation of Kohn-Sham density functional theory is used to study the gas-phase conformational dynamics of an alanine dipeptide analog. It is found that conformational transformation between $C5$ and $C7_{eq}$ occurs on the picosecond time scale. Classical MD simulations using most of the popular force fields do not yield a transition even after nanoseconds. An analysis is given of the difference, for this small gas-phase system, between *ab initio* MD and traditional MD simulation using force fields.

DOI: 10.1103/PhysRevE.64.011907 PACS number(s): 87.15.He, 31.15.Qg, 03.67.Lx

Classical molecular dynamics (CMD) simulations have played an increasingly important role in our understanding of the structures and functions of biological systems [1–6]. The accuracy and reliability of such MD simulations depend crucially on the empirical potential energy functions that are parametrized to fit to experimental data as well as the results from *ab initio* calculations. Recent advances in *ab initio* molecular dynamics (AIMD) methodologies [7,8] make it possible to study a variety of molecular systems directly, where the forces at each configuration are computed from quantum mechanical calculations. This allows us to make important comparisons of the performances of CMD and AIMD for systems of biological interest. In this article, we report the results of AIMD and CMD simulations on (S)-α-(formylamino)propanamide (an alanine dipeptide analog) which reveal significant differences in the dynamic behavior of this system obtained from these two approaches. The conformational transitions between the two lowest energy conformers ($C5$ and $C7_{eq}$) occur on the picosecond (ps) time scale from the AIMD simulations, while CMD simulations with some of the widely used force fields do not yield a transition even after nanoseconds. Although AIMD has been used previously to study systems of biological interest, to our knowlege such large differences in the conformational transitions have not been reported.

The AIMD simulation method was pioneered by Car and Parrinello (CP) [9]. The CP-AIMD simulation involves the solution of coupled dynamical equations of motion for the nuclear coordinates and molecular orbital expansion coefficients. Our AIMD simulation is carried out at the Born-Oppenheimer (BO) level [8,10–13]. In this approach, it is assumed that the motion of the nuclei is much slower than that of electrons and is always in equilibrium with the electronic structure. Therefore, the classical dynamical equations of motion can be used to give complete trajectories if the potential surface is known. At each time step in the molecular dynamics simulation, the forces on the nuclei are given by the energy gradients and the molecular orbitals are updated by solving a Schrödinger-type equation within the BO approximation. To reduce the computational demand, we use (S)-α-(formylamino)propanamide (an analog of the alanine dipeptide) [14]. The analog is formed by replacing the terminal methyl groups of the alanine dipeptide by hydrogen atoms; a previous study suggests that the effects of such replacements on the (ϕ,ψ) energy map (see below) are rather small [14]. A schematic diagram of the alanine dipeptide analog is shown in Fig. 1. The simulations start with the β-like structure. The velocity is picked from a Maxwell-Boltzmann distribution [15]. The initial velocity norm is given so that the kinetic temperature is about 298 K. The Verlet integrator [16] was used to propagate the equations of motion for the nuclei. The time step is taken to be 50 a.u. (i.e., 1.21 fs) and the simulation is run for about 70 000 time steps. A smaller time step, for example, 20 a.u., was used for a few test runs. This does not significantly change the results. The quantum density functional theory DFT calculations were carried out using the DZVP basis set [17], which is of double-ζ quality and includes field-induced polarization functions. The Becke-Perdew (BP) approximation was used for the exchange and correlation functional [18,19]. Comparison of the BP/DZVP results with those from post-Hartree-Fock calculations for the stable conformations (Frey *et al.* [20]) suggests that the density functional method used in this study is of sufficient accuracy for our purpose of mapping out the general features of the dynamics; the relative energies for the $C7_{eq}$ and $C5$ conformers are 0.0 and 1.7 kcal/mol, respectively, which are nearly identical to those

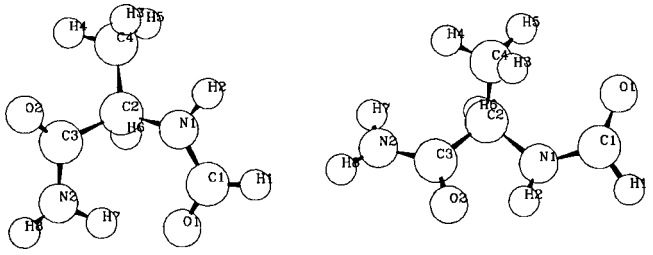

FIG. 1. Schematic diagram of alanine dipeptide analog. The left and right figures refer to $C7_{eq}$ and $C5$ conformations, respectively.

*Present address: Steacie Institute for Molecular Sciences, National Research Council of Canada, 100 Sussex Drive, Ottawa, Ontario, Canada K1A 0R6.

†Present address: Department of Chemistry and Chemical Biology, Harvard University, Cambridge, MA 02138.

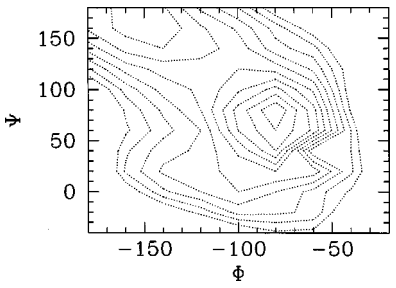

FIG. 2. The DFT [ϕ (deg), ψ (deg)] potential energy contours (every 0.5 kcal/mol), obtained from a grid with 20° spacing.

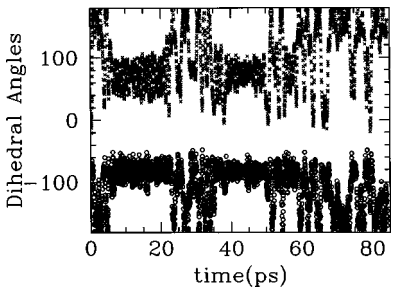

FIG. 4. Time evolution of the dihedral angles. The open circles and crosses are for ϕ (deg) and ψ (deg), respectively.

from MP2/6-311G** calculations with complete geometry optimization (the energy difference between $C7_{eq}$ and $C5$ is 1.66 kcal/mol [20]).

The conformational space of the alanine dipeptide is conveniently described by the backbone dihedral angles ϕ [$\angle$(C1-N1-C2-C3)] and ψ [$\angle$(N1-C2-C3-N2)]. *Ab initio* studies [14,20–23] have shown that the internally hydrogen-bonded conformations, the cyclic hydrogen-bonded $C7_{eq}$ structure [O1$\cdots$H7] and the extended $C5$ [O2$\cdots$H2] structure, are of lowest energy. The $C7_{eq}$ and $C5$ structures have (ϕ,ψ) angles of about ($-80°$, 80°) and ($-160°$, 160°), respectively. To have a clear view of the energy landscape near $C7_{eq}$ and $C5$ conformations we show a contour plot of the potential energy surface in Fig. 2, which was obtained with full geometry optimization by fixing the corresponding dihedral angles for each point with a grid spacing of 20°. The barrier from $C7_{eq}$ to $C5$ is 2.5 kcal/mol which is much lower than the value of 4.86 kcal/mol obtained previously from Hartree-Fock calculations with a 3-21G basis set [14], but reasonably close to the MP2/6-31 +G**/HF/6-31+G* barrier (1.82 kcal/mol) [20] (without geometry optimization).

In Fig. 3, we plot the bond lengths as a function of time. The solid and dotted lines correspond to the O$\cdots$H bond lengths for the two internally hydrogen-bonded conformers $C7_{eq}$ and $C5$, respectively. An O$\cdots$H bond length of about 2.0 Å or less is taken to mean that a hydrogen bond has been formed. By monitoring the forming and breaking of the two types of hydrogen-bonds we can follow the conformational dynamics in the simulation. We clearly see the conversion from the $C7_{eq}$ to the $C5$ conformation. From Fig. 3, we observe two distinct time scales. For about 15 ps, the

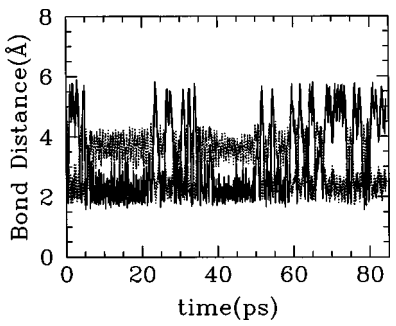

FIG. 3. Time evolution of the H-bond lengths. The solid and dotted lines are for O1$\cdots$H7 and O2$\cdots$H2, respectively.

$C7_{eq}$ O$\cdots$H bond vibrates around its equilibrium value, i.e., $C7_{eq}$ is sampled for a significant duration, then we see a few consecutive transformations between $C7_{eq}$ and $C5$, and each transformation takes about 3 ps to complete. The $C7_{eq}$ comformation is of lower energy than $C5$ and the energy difference is about 1.7 kcal/mol. This is consistent with the fact that the O$\cdots$H bond length for the $C7_{eq}$ conformation reaches a lower value than that of the $C5$ conformation because the $C7_{eq}$ O$\cdots$H bond is stronger. Our results indicate that $C7_{eq}$ is heavily populated in the phase space at room temperature. In Fig. 4, we show a plot of the dihedral angles ϕ and ψ as a function of time. Again we see the interchange of $C7_{eq}$ and $C5$. In addition, we see some population of the β-like conformation ($-134°$,57°). The α-like conformation ($-60°$, $-40°$) is also visited, but less frequently.

This time scale of picosecond conformational transformation is an interesting discovery. We have carried out classical simulations using the CHARMM [24], MM3 [25], AMBER [26], and ESFF [27] force fields. We did not see any transition between $C5$ and $C7_{eq}$ conformations at 300 K for simulations lasting nanoseconds. Two factors might be considered to explain the different behavior of CMD and AIMD. First, the energy barrier in these force fields may be higher than that of DFT-BP. Second, parameters in the force fields may not be appropriate for the gas phase. For example, the rotational barrier corresponding to the dihedral angle ω may be too high, or in some force fields ω is kept constant. It has been assumed that the peptide unit [with a peptide (C-N) bond at the center] has a rigid, planar structure, i.e, the ω angle, [$\angle$(C2-C3-N2-H8(C))]$=180°$ and ω' [$\angle$(C2-C3-N2-H7)]$=0°$. However, significant distortion from planarity has been observed by experiments [28] and theoretical studies [14]. In some region of ϕ,ψ space, it might be as large as 40°.

In Fig. 5, we plot a statistical profile of ω and ω'. The population looks similar to the result of a survey using the Cambridge Structural Database of small molecules [28]. It is a Gaussian distribution with a standard deviation of 13° and 19°, respectively, which is larger than the value of the survey, about 6°. The difference could come from the fact our simulation is for a dipeptide analog in the gas phase, where the CH_3 group of the dipeptide is replaced by a hydrogen atom. In the same figure, we also plot the pyramid angle θ, which is defined as $\theta=\omega'-\omega+180.0$. The distribution of θ does not have a Gaussian form. It may deviate from its average value of 0° by as much as 50° (with a fluctuation of

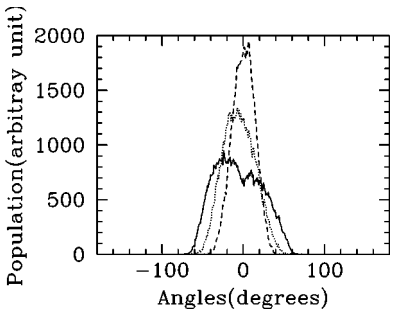

FIG. 5. Population profile of the ω and ω' dihedral angles and θ angle (in degrees). The solid, dotted, and dashed lines are for θ, ω [$\angle$(C2-C3-N2-H8(C))], and ω' [$\angle$(C2-C3-N2-H7)], respectively.

28°). We also calculated the statistical profiles of other dihedral angles, for example, [$\angle$(O1-C1-N1-H2)], [$\angle$(H7-N2-C3-O2)], [$\angle$(H2-N1-C1-H1)], and [$\angle$(O1-C1-N1-C2)], which are all treated as ''improper'' torsion angles in the CMD simulations using the CHARMM force field. Significant fluctuations have been seen for all these angles. The time dependence of these angles is rather interesting. They all fluctuate at a much faster pace than the backbone dihedral angles ϕ and ψ, i.e., with a period of 0.1 ps compared with 0.5 ps for ϕ and ψ.

If we increase the temperature to, say, 375 K, a force field such as CHARMM [24] does show conformational transformations between $C5$ and $C7_{eq}$(with dominant population), and the time scale is about 20 ps.

Figure 6 presents a Ramachandran-like plot of the trajectories of the dihedral angles ϕ and ψ which may be compared with the result of CHARMM shown in Fig. 7. If the simulation is carried out for a long enough time, the map of the trajectory should reflect the landscape of the potential energy surface, which does not depend upon the dynamic details. The DFT trajectory reaches a much wider range of angles than does CHARMM. In fact, the potential surfaces of DFT and CHARMM in terms of ϕ and ψ are quite similar in the region our system explores, as is clear from Fig. 8, which shows a complete ϕ-ψ energy map of the CHARMM force field. It was obtained with full geometry optimization by fixing the corresponding dihedral angles for each point with a grid spacing of 15°. The barrier between $C7_{eq}$ and C_5 is

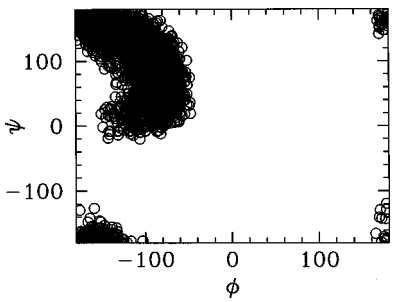

FIG. 6. Ramachandran-like plot of the DFT trajectories for the dihedral angles ϕ (deg) and ψ (deg).

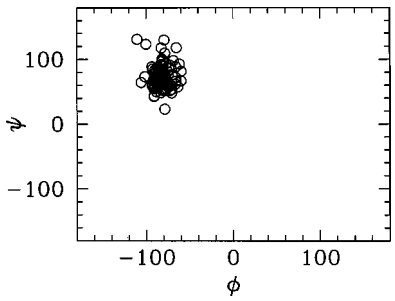

FIG. 7. Ramanchandran-like plot of the CHARMM trajectories for the dihedral angles ϕ (deg) and ψ (deg).

about 2.4 kcal/mol which is quite close to the DFT value of 2.5 kcal/mol. As we pointed out earlier, other dihedral angles, some of which are defined as improper in CHARMM, undergo very significant change during our simulation and should contribute to the dynamical and thermodynamical behavior of the dipeptide analog. Presumably, less flexibility with respect to the dihedral angles in the CHARMM force field prevents the system from escaping from $C7_{eq}$ at this temperature, i.e., 300 K.

We have obtained a potential map in the area around the internal hydrogen-bonded conformations $C5$ and $C7_{eq}$. A Ramachandran plot derived from gas-phase *ab initio* molecular dynamics simulation shows dipeptide conformational changes on the picosecond time scale. The classical force fields do not yield any observable transformation even for nanosecond simulations. Given that the classical force fields have reasonably similar (ϕ,ψ) energy maps, this observation may seem rather surprising. However, one must not forget that the real surface has 42 dimensions and the energy fluctuation depends in principle on all of these. Instantaneous quantum electronic effects, for example, polarization, correlation, and charge transfer, have a very significant effect on the dynamics behavior of biologically interesting systems. This certainly creates a challenge to develop new force fields. The conformational transformation on the picosecond

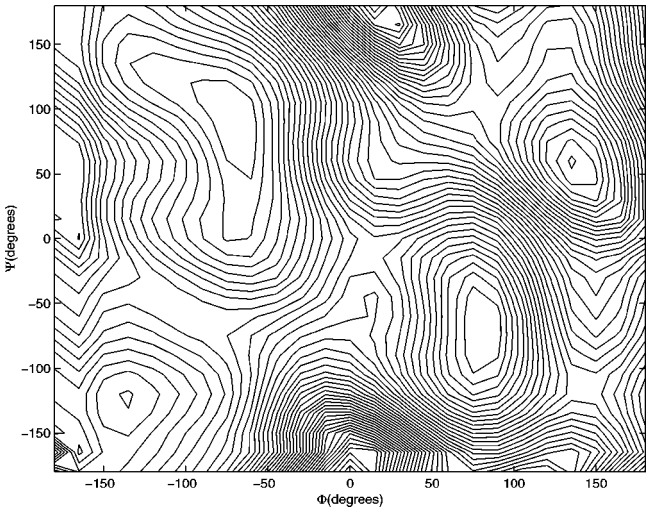

FIG. 8. The CHARMM [ϕ (deg), ψ (deg)] potential energy contours (every 0.5 kcal/mol), obtained from a grid with 15° spacing.

time scale could perhaps be observed by future gas-phase experiments.

We are grateful to Dr. Glenn J. Martyna for helping us with the CHARMM calculations and also to Dr. A. D. Maskerell for providing the data of the $\phi-\psi$ map. The financial support of the Natural Sciences and Engineering Research Council of Canada is gratefully acknowledged.

[1] C. L. Brooks III, M. Pettit, and M. Karplus, *Proteins, A Theoretical Perspective of Dynamics, Structure, and Thermodynamics*, Vol. 71 in *Advances in Chemical Physics* (John Wiley and Son, New York, 1988).

[2] J. A. McCammon and S. Harvey, *Dynamics of Proteins and Nucleic Acids* (Cambridge University Press, Cambridge, 1987).

[3] C. L. Brooks III and D. Case, Chem. Rev. **93**, 2487 (1993).

[4] P. E. Smith, B. M. Pettit, and M. Karplus, J. Phys. Chem. **97**, 6907 (1993).

[5] F. Fraternali and W. F. Van Gunsteren, Biopolymers **34**, 347 (1994).

[6] T. J. Marrone, M. K. Gilson, and J. A. McCammon, J. Phys. Chem. **100**, 1439 (1996).

[7] M. E. Tuckerman, P. J. Ungar, T. von Rosenvinge, and M. L. Klein, J. Phys. Chem. **100**, 12 878 (1996), and references therein.

[8] D. Q. Wei and D. R. Salahub, J. Chem. Phys. **106**, 6086 (1997).

[9] R. Car and M. Parrinello, Phys. Rev. Lett. **55**, 2471 (1985).

[10] D. Berard, D. Wei, and D. R. Salahub, in *Proceedings of Pacific Symposium on Biocomputing*, Hawaii, 1997, edited by R. B. Altman, A. K. Dunker, L. Hunker, and T. E. Klein (World Scientific, Singapore, 1997).

[11] R. Barnett and U. Landman, Phys. Rev. B **48**, 2081 (1993).

[12] X. Jing, N. Troullier, D. Dean, N. Binggeli, J. R. Chelikowsky, K. Wu, and Y. Saad, Phys. Rev. B **50**, 12 234 (1994).

[13] D. Gibson, I. Ionova, and E. Carter, Chem. Phys. Lett. **240**, 261 (1995).

[14] T. Head-Gordon, M. Head-Gordon, M. J. Frisch, C. L. Brooks III, and John A. Pople, J. Am. Chem. Soc. **113**, 5989 (1991).

[15] M. P. Allen and D. J. Tildesley, *Computer Simulation of Liquids* (Oxford University Press, London, 1989).

[16] L. Verlet, Phys. Rev. **159**, 98 (1967).

[17] A. St-Amant and D. R. Salahub, Chem. Phys. Lett. **169**, 387 (1990); Alain St-Amant, Ph.D. thesis, Université de Montréal, 1992; J. W. Andzelm, M. E. Casida, A. Koester, E. Proynov, A. St-Amant, D. R. Salahub, H. Duarte, N. Godbout, J. Guan, C. Jamorski, M. Leboeuf, V. Malkin, O. Malkina, F. Sim, and A. Vela, deMon-KS version 1.2, University of Montreal, 1995.

[18] J. P. Perdew, Phys. Rev. B **33**, 8822 (1986).

[19] A. D. Becke, Phys. Rev. A **38**, 3098 (1988).

[20] R. F. Frey, J. Coffin, S. Q. Newton, M. Ramek, V. K. W. Cheng, F. A. Momany, and L. Shäfer, J. Am. Chem. Soc. **114**, 5396 (1991).

[21] M. D. Beachy, D. Chasman, R. B. Murphy, T. A. Halgren, and R. A. Friesner, J. Am. Chem. Soc. **119**, 5908 (1997).

[22] I. R. Gould, W. D. Cornell, and I. H. Hillier, J. Am. Chem. Soc. **116**, 9250 (1994).

[23] W. D. Cornell, I. R. Gould, and P. A. Kollman, J. Mol. Struct.: THEOCHEM **392**, 101 (1997).

[24] A. D. MacKerell, Jr., D. Bashford, M. Bellott, R. L. Dunbrack, Jr., J. Evanseck, M. J. Field, S. Fischer, J. Gao, H. Guo, S. Ha, D. Joseph-McCarthy, L. Kuchnir, K. Kuczera, F. T. K. Lau, C. Mattos, S. Michnick, T. Ngo, D. T. Nguyen, B. Prodhom, W.E. Reiher III, B. Roux, M. Schlenkrich, J. Smith, R. Stote, J. Straub, M. Watanabe, J. Wiorkiewicz-Kuczera, D. Yin, and M. Karplus, J. Phys. Chem. B **102**, 3586 (1998). A time step of 1.5 fs was used in the CMD simulation, and minimum 100 000 steps were requested. No bonds and bending angles were fixed; some torsion dihedral angles were constrained according to criteria defined in CHARMM2.

[25] N. L. Allinger, Y. H. Yuh, and J.-H. Lii, J. Am. Chem. Soc. **111**, 8551 (1989).

[26] W. D. Cornell, P. Cieplak, C. I. Bayly, I. R. Gould, K. M. Merz, Jr., D. M. Ferguson, D. C. Spellmeyer, T. Fox, J. W. Caldwell, and P. A. Kollman, J. Am. Chem. Soc. **117**, 5179 (1995).

[27] *User Guide*, Biosym/MSI, 9685 Scranton Road, San Diego, CA 92121-3752.

[28] M. W. MacArthur and J. M. Thornton, J. Mol. Biol. **264**, 1180 (1996).