Enol-to-keto Tautomerism of Peptide Groups

Katsumasa Kamiya,*,† Mauro Boero,†,‡ Kenji Shiraishi,† and Atsushi Oshiyama†,‡

Institute of Physics and Center for Computational Sciences, University of Tsukuba, Tennodai 1-1-1, Tsukuba, Ibaraki 305-8571, Japan

Received: October 30, 2005; In Final Form: December 21, 2005

Density functional based simulations, performed on polyglycine containing an enol peptide group [-C(OH)N-] which is a structural isomer of a keto form [-CONH-], show that in the enol-to-keto tautomeric reaction, the enol peptide group is less stable than the keto form, and that the enol-to-keto tautomerism is characterized by a cis/trans isomerization of the C−N peptide bond. The rate-limiting step in the cis/trans isomerization is a hydrogen migration from O to N atoms in the peptide group with a transition state consisting of a four-membered ring in the cis configuration. An analysis of the cis/trans isomerization pathway shows that the mechanisms for the cis/trans isomerization are essentially different between the enol and keto forms.

1. Introduction

Proteins are polypeptides whose building blocks are amino acids. A protein is formed by polymerization reactions of a carboxylic acid group of one amino acid with an amino group of another amino acid to produce a chain of peptide groups. The amide groups can form several isomers as shown in Figure 1: enol forms are structural isomers of keto forms, and both forms have trans and cis stereoisomers. Among all these amide isomers, the keto-trans is the most stable form;[1,2] hence, it is not surprising that the majority of the peptide groups takes this form in proteins.[3] Other isomers, however, have been detected in simple amides (model peptide systems) such as formamide and *N*-methylacetamide,[4−12] and this has stimulated the investigation of their presence in the peptide groups of proteins.

Indeed, keto-cis forms of peptide groups have been identified in the X-ray crystal structures of proteins.[13−15] Furthermore, the cis-to-trans isomerization of peptide groups is likely to play a role in biochemical processes such as protein folding and cell signaling, as inferred by the discovery of enzymes catalyzing the cis/trans isomerization.[15−17] Given this scenario, the underlying reaction mechanism has been extensively investigated, both experimentally[18−20] and theoretically,[4−8] using simplified model systems, such as simple amides and dipeptides, to disentangle the complexity of the reaction pathway.

The existence of enol-trans peptide groups has been inferred from recent experiments. Namely, the enol form is a product of the deprotonation reaction of an imidic acid, −C(OH)NH-, occurring in a basic environment as[9,10]

$$H^+ + -CONH- \rightleftharpoons -C(OH)NH- \rightleftharpoons -C(OH)N- + H^+$$

X-ray crystallography performed on cytochrome *c* oxidase indicated that a water molecule is located inside the enzyme in the vicinity of a peptide carbonyl group.[21,22] More recently, neutron scattering analysis performed on lysozyme provided further support to this picture, suggesting that a bulk water molecule might accesses a peptide carbonyl group and form a hydrogen bond with it.[23] A general picture that can be drawn is that the water molecule is part of a hydrogen bond network and a proton can reach the carbonyl groups via a proton wire mechanism.[24] The protonation of the carbonyl group thus produces imidic acid. The fact that, in this case, the nitrogen gets deprotonated and not the carbonyl is due to the peculiarity of the cytochrome *c* oxidase environment. In fact, one basic amino acid residue (COO^-) from the Asp is located in the vicinity of the amide proton,[21,22] and this eventually promotes the transition into enol form.

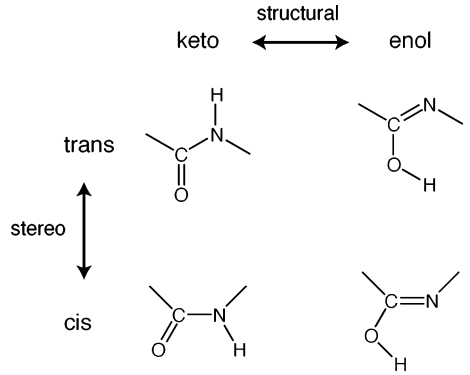

Figure 1. Schematic representation of structural and stereoisomers of amide groups.

The importance of the enol-to-keto tautomerism stems from its basic role in a variety of physiological processes. An example is represented by cytochrome *c* oxidase,[25] the reaction that pumps protons at the final stage of cell respiration. In mutagenesis experiments, the mutated peptide bond is expected to assume the enol form, and as a result, the mutants do not show any proton-pumping activity, suggesting that the enol-to-keto tautomerism might be related to the proton-transfer process. However, despite the extensive experimental outcome about the enol peptide group, its fundamental characteristics are still far from being unraveled.

Density Functional Theory (DFT)[26,27] can provide a computationally efficient approach as shown by its successful applications to a variety of chemical and biochemical systems, often in combination with methods suitable to sample reaction paths.[28−40] More recently, the metadynamics approach[41] com-

* To whom correspondence should be addressed. phone: +81-29-8535921, e-mail: kkamiya@comas.frsc.tsukuba.ac.jp.
† Institute of Physics.
‡ Center for Computational Sciences.

10.1021/jp056250p CCC: $33.50
Published on Web 02/03/2006

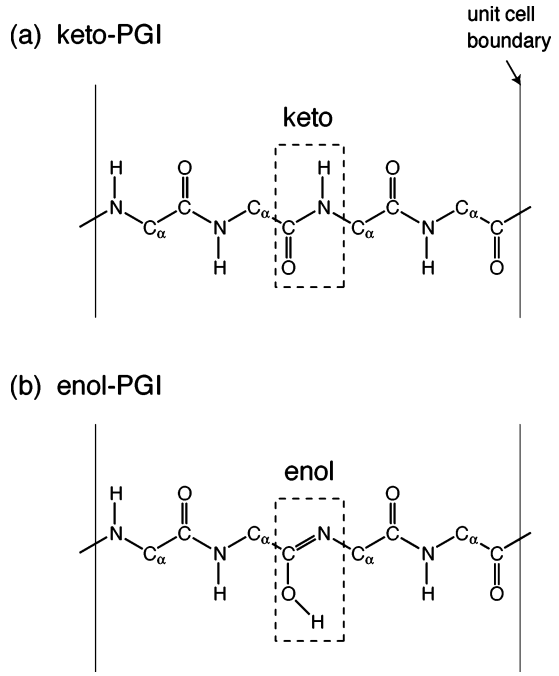

Figure 2. Polyglycine chain systems used in this work: (a) polyglycine with infinite length (PGI) consisting of keto peptide groups, and (b) PGI including one enol peptide group per unit cell. Hydrogen atoms bound to C_α atoms are not shown in this figure for the sake of clarity.

bined with Car−Parrinello molecular dynamics[42] has been shown to be a versatile tool to study also complicated reaction mechanism.[43−47]

The purpose of the present work is to elucidate the characteristics of the enol peptide group and the mechanisms of the enol-to-keto tautomerism via DFT-based calculations. Our model system consists of an isolated polyglycine of infinite length (PGI) containing one enol peptide group. The glycine is the simplest alpha−amino acid, with its side chain being hydrogen; thus, the molecular and electronic structures of the enol peptide group in the PGI display characteristics similar to an enol peptide group in a common peptide backbone of proteins. Furthermore, since no catalyst exists around the enol peptide group in the PGI, its tautomerism to a keto form strongly reflects the intrinsic character of enol peptide groups.

On the basis of our calculations, we have traced the distinctive features of the enol peptide group; the C−N peptide bond has a double bond character. We determine a reaction path for the enol-to-keto tautomerism, finding that this tautomeric reaction is characterized by a cis/trans isomerization of the C−N peptide bond. The rate-limiting step for the tautomeric reaction is a hydrogen migration from O to N in the peptide group. The transition state (TS) is a four-membered ring structure showing a cis configuration, and the activation energy is about 26 kcal/mol. An analysis of the calculated pathways shows that the mechanisms of the cis/trans isomerization are essentially different between the enol and keto forms. This difference is strongly reflected in the nature of the highest occupied and lowest unoccupied electron states near the energy gap of the PGI.

2. Computational Details

We prepared two different polyglycine chains, as shown in Figure 2, in an orthorombic supercell of sizes $a = 14.510$ Å and $b = c = 9.000$ Å. Since periodic boundary conditions have been adopted, this amounts to a virtually infinite chain with a periodicity equal to a and a separation from the periodically repeated adjacent images equal to b. As can be seen by looking at Figure 2, our models include four peptide groups in a unit cell for each chain; the keto−PGI consists of all keto peptide groups (Figure 2a) and enol−PGI includes one enol peptide group per unit cell (Figure 2b)

In our static DFT calculations,[48] we use the Perdew−Burke−Ernzerhof (PBE) generalized gradient approximation[49] on the exchange and correlation, while core-valence interactions are described by ultrasoft pseudopotentials.[50] Valence electron wave functions are expanded in a plane-wave basis set with a cutoff energy of 36 Ry; two and four **k**-points in the Brillouin zone (BZ) along the chain direction are used for the calculations of the reaction pathways and the analysis of the electronic structures of the intermediate states, respectively. These parameters have been already shown to be sufficient to achieve a good convergence.[36] A conjugated gradient minimization is used for both the electronic-structure calculations and the geometry optimizations. In the geometry optimization, all atoms are relaxed until the residual forces become less than 5.1×10^{-2} eV/Å. In all static calculations, the reaction path is searched by the constrained optimization in a (N-1)-dimensional space, where N is the number of degrees of freedom in one unit cell.[37,38] The transition states and activation (total) energies of the reaction are calculated by the force inversion method.[40]

In dynamical simulations, the reaction paths are sampled by means of the metadynamics approach based on the Car−Parrinello method.[41] Car−Parrinello simulations[51] are performed for the PGI model within the same PBE functional adopted for the static calculations and using norm-conserving pseudopotentials.[52] For this reason, we increased the cutoff energy to 70 Ry, and the BZ was sampled at the Γ point only. The relative total energy differences of the intermediate states in the reaction, as obtained by this dynamical approach, agree with the corresponding values obtained from static DFT calculations within 0.9 kcal/mol. This provides an independent check to our computational results. The ionic temperature is controlled by velocity rescaling and kept at 300 ± 50 K. The MD time step is set to 0.097 fs and a fictitious electron mass of 400 au ensures a good control of the conserved quantities and preserves the Born−Oppenheimer adiabaticity. In the case of metadynamics, the collective variables representing the reaction coordinates are treated as new dynamical variables and added to the Car−Parrinello Lagrangian along with a history-dependent Gaussian potential as described in ref 41. In our specific cases, we selected as collective variables the N−H distance (cv1) and the torsion angle C_α−C−N−C_α (cv2) in the peptide group undergoing the reaction (see Figure 5 for details). These variables represent all the slowly varying degrees of freedom and account for the formation of the N−H bond and the twisting of the peptide backbone, respectively. Two types of simulations were performed to disentangle the contributions of cv1 and cv2: both cv1 and cv2 were used in a first simulation, while only cv1 was used in a second simulation. In the former simulation, fictitious effective masses $M_\alpha^{cv1} = M_\alpha^{cv2} = 50$ au, and harmonic coupling constants $k_\alpha^{cv1} = k_\alpha^{cv2} = 0.4$ were adopted. In the latter, $M_\alpha = 50$ au and $k_\alpha = 1.2$ were used. These parameters ensured a good adiabatic decoupling between fast and slow variables. A new Gaussian is added every 4.85 fs and its width and height are sampled in the intervals [0.042,0.127] Å and [0.027;0.098] eV, respectively.

3. Results and Discussion

3.1. Characteristics of the Enol−Trans Peptide Group. We started with fully relaxed enol−trans and keto−trans peptide

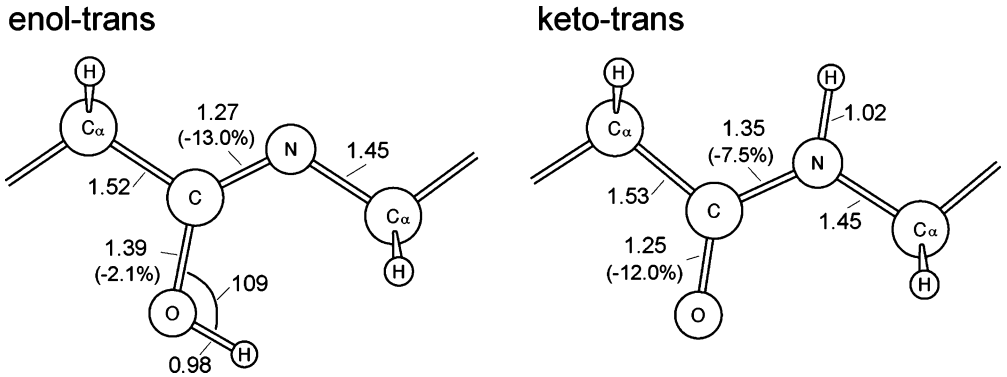

Figure 3. Optimized structures of the enol−trans and keto−trans peptide groups in the enol−PGI and keto−PGI. Distances and angles are expressed in Å and degrees, respectively. The value in each parenthesis is the percent difference in bond length as compared with typical values of single bonds: C−N (1.46) and C−O (1.42).

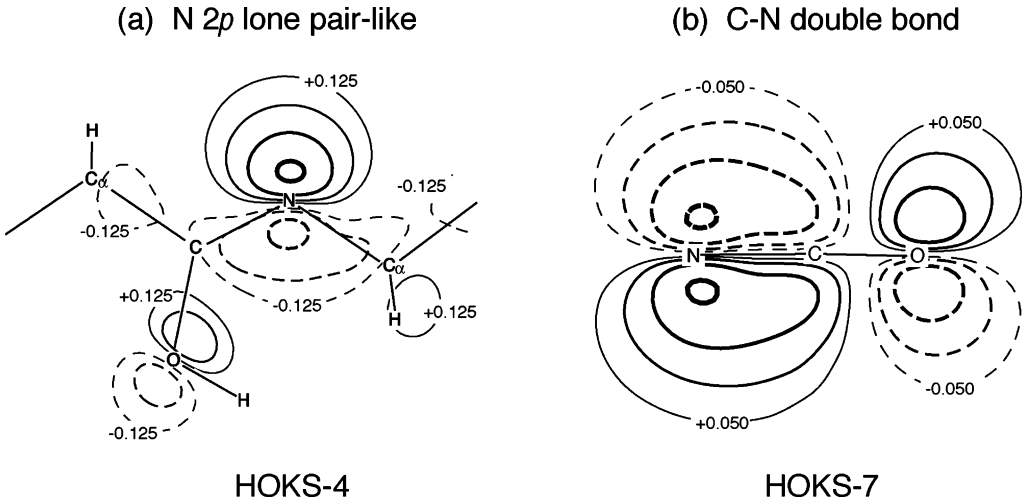

Figure 4. Some representative occupied Kohn−Sham (KS) orbitals of the enol−trans peptide group in the PGI, shown as contour density maps. Each contour denotes twice (or half) the values of the adjacent contour lines: The units of the values are $(e/Å^3)^{1/2}$.

TABLE 1. Mulliken Atomic Charges of the Atoms in the Peptide Groups of the Enol−Trans and Keto−Trans

	enol-trans	keto-trans
C	0.290	0.370
H	0.403	0.324
O	−0.300	−0.272
N	−0.193	−0.268

model systems (Figure 3), hereafter indicated as enol−PGI and keto−PGI. As expected, the structural characteristics of the N−C−O group are different in the enol and keto forms as far as the location of the double bond is concerned: The structures are C═N and C−O with distances of 1.27 and 1.39 Å, respectively, in the enol−trans and C−N and C═O with distances of 1.35 and 1.25 Å, respectively, in the keto-trans form; a simple comparison of typical single and double bonds[53] provides support to this picture.

These structural features are reflected in the electronic states of the enol−trans peptide group. Figure 4 shows several Kohn−Sham (KS) states of enol−trans peptide group in PGI. The KS state of the enol−trans peptide group in PGI in Figure 4a is mainly due to N 2p orbital perpendicular to the backbone, with tails along the adjacent bonds. The KS orbital in Figure 4b is, instead, originated by 2p orbitals of N, C, and O perpendicular to the peptide plane (the plane defined by N−C−O), and can be ascribed to the anitbonding ppπ* of N and C orbitals, plus a contribution from O 2p.

Mulliken charges are summarized in Table 1 for the atoms in the peptide groups of the enol-trans and keto-trans. It can be noticed that the Mulliken charge of C is less positive and that of O is more negative in the enol−trans with respect to the keto-trans, while an opposite trend is seen for H and N. The difference in the bonding of one H atom between the enol and keto (Figure 3), i.e., O−H or N−H, and basic considerations about the electronegativity,[54] allow us to infer that H donates its electron more easily to O or N rather than C when it forms a chemical bond with these atoms. The trend described above is, then, not surprising. As far as the carbon atom is concerned, its positive Mulliken charge originates from the electron withdrawing by N and C via the C−N and C−O bonds. Considering the local electronic structure of the enol (N═C−O) and keto (N−C═O) groups, one can infer that the electronic charge of C moves toward N in the enol case and toward O in the keto case.

3.2. Enol-to-Keto Tautomerism of the Peptide Group in PGI. In this subsection, we focus on the reaction pathways for the enol-to-keto tautomerism in PGI. The peptide group in the PGI is chosen to be in the enol−trans form in the initial state. Our calculations show that the enol−trans−PGI is energetically located at 28 kcal/mol above the keto−trans−PGI, and this is consistent with the observation that the enol−trans is unstable in simple amides.[9−12] To perform a comprehensive study of the reaction pathways for the enol−to−keto tautomerism, we have used both the constrained optimization method and the meta-dynamics approach, independently. Two alternative reaction paths have been identified: the first one leads to a cis/trans isomerization of the C−N peptide bond while the second one

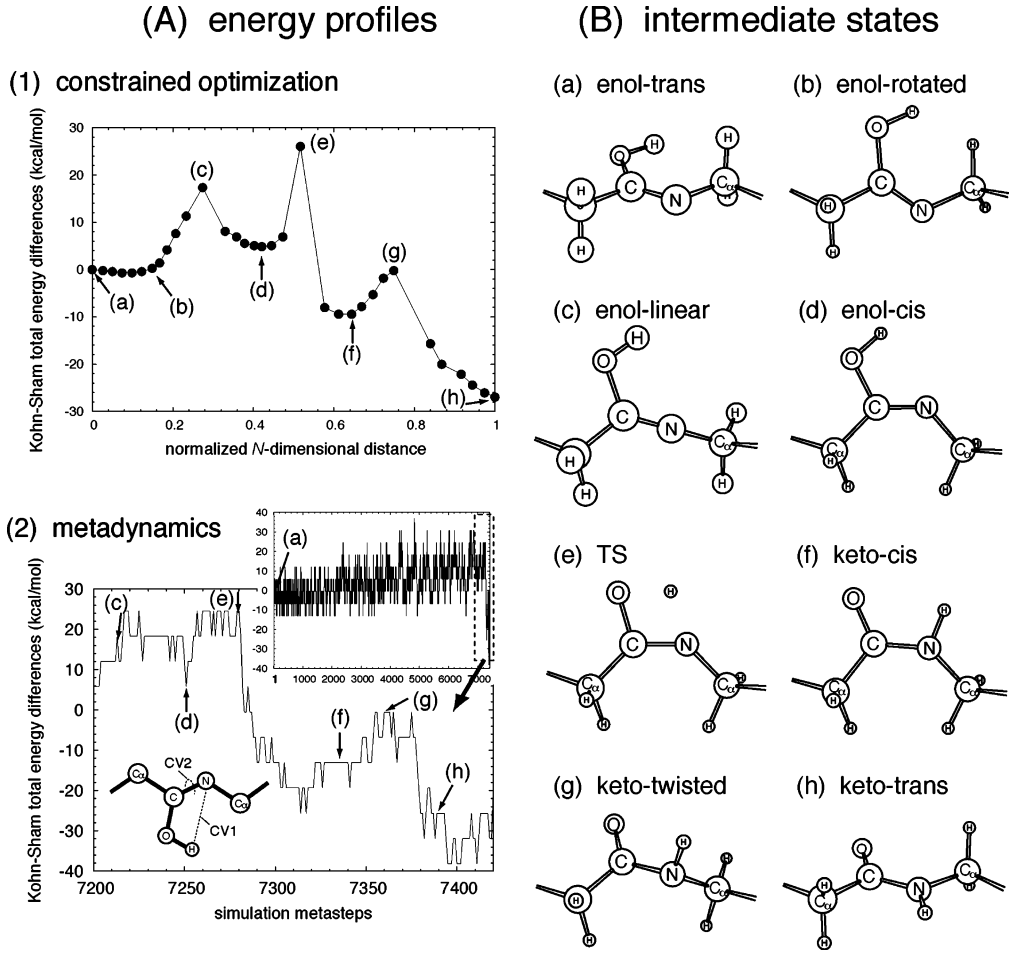

Figure 5. Reaction profiles for the cis/trans isomerization pathway of the enol-to-keto tautomerism. (A-1) Energy profile calculated by constrained optimization. The solid line is intended only as a guide for the eye. (A-2) Energy profile calculated via metadynamics. The origin along the energy axis is set as the Kohn−Sham total energy averaged over the exploration of the reactant local minimum. The lower left inset shows the collective variables selected for the calculation. Labels along each profile refer to the configurations shown in the right panels. (B) Structures of the peptide group in the (meta)stable (a,b,d,f,h) and transition (c,e,g) states.

does not and the two are characterized by different transition states. However, the cis/trans isomerization pathway turns out to be the most favorable.

3.2.1. Cis/Trans Isomerization Reaction Pathway. Figure 5 shows the energy profiles and the intermediate states calculated via constrained optimization and metadynamics, respectively. In the metadynamics simulations, the selected collective variables (inset in Figure 5A2) are the N−H distance (cv1) and the torsion angle $C_\alpha-C-N-C_\alpha$ (cv2). This choice ensures that all the slowly varying degrees of freedom, able to discriminate between reactant and product, are included. As shown in Figure 5, the two methods lead, essentially, to the same reaction pathway: in the case of metadynamics, the system explores, several times, the cis/trans isomerization of the enol peptide group (Figure 5Ba−d) and eventually the enol peptide tautomerizes to the stable keto form, found also in constrained optimizations (Figure 5Bc−h). The free energy landscape, as obtained from the metadynamics simulation, is shown in Figure 6, and, from a qualitative point of view, it shows that the torsion angle gives a certain range of stable minima for the final keto form. We must remark, however, that due to the fact that the polyglycine chain, in our simulations, is kept fixed at the two ends by the periodic boundary conditions and is in a vacuum, the entropy contribution is very small, thus free energy differences and averaged total energy differences do not differ significantly.

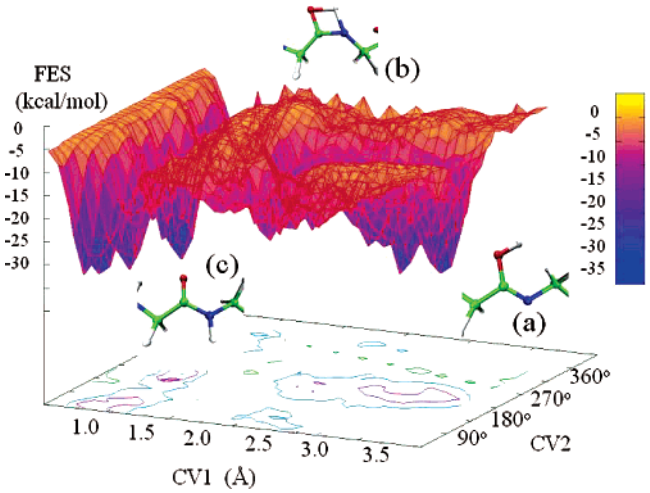

Figure 6. The reconstructed free energy landscape for the metadynamics simulation in which two collective variables are used. The collective variable indicated as CV1 is the N−H distance, while CV2 is the torsion angle as described in the text. The initial (a), final (c), and the highest energy intermediate (b) configurations are shown for the sake of clarity. The color code for atoms is red for O, blue for N, green for C, and white for H.

The cis/trans isomerization process can be summarized in three main elementary steps. The first one (Figure Ba−d) is

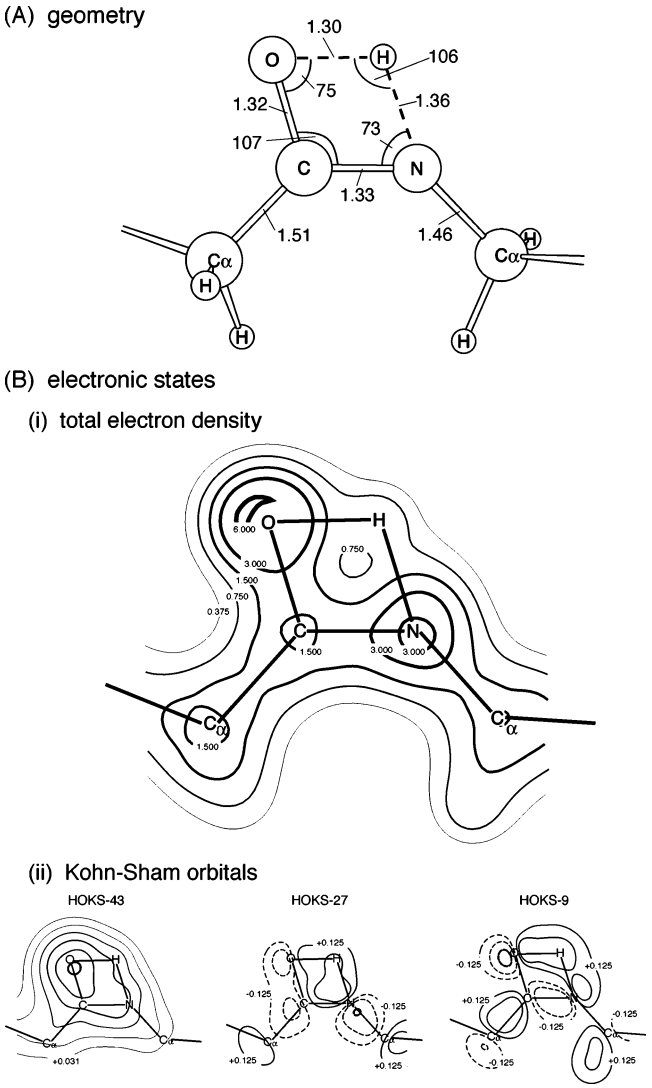

Figure 7. The transition state in the rate-limiting step for the cis/trans isomerization pathway. (A) Geometry of the four-membered ring; distances and angles are expressed in Å and degrees, respectively. (B) Contour maps of (1) the total electron density (e/Å^3) and (2) the corresponding Kohn−Sham orbitals (e/Å^3)$^{1/2}$. Each contour denotes twice (or half) the values of the adjacent contour lines.

the trans-to-cis isomerization of the enol form, starting as a rotation of the peptide group around the adjacent $C_\alpha-C$ and $N-C_\alpha$ bonds with no appreciable barrier (Figure Ba and b); then the enol-trans isomerizes to the enol−cis form (c and d). The second step (Figure Bd−f) is the rate-limiting step of the reaction, involving a hydrogen migration in the cis form of the peptide group. This is the process that practically accounts for the whole activation barrier, estimated to be about 26 kcal/mol. The third step (Figure Bf−h) is the cis-to-trans isomerization of the keto form and occurs by overcoming a rather modest energy barrier.

As mentioned, the rate-limiting step is represented by the migration of one hydrogen from O to N in the cis-configurational peptide group. Figure 7A shows the related geometry of the TS, characterized by a rhombic four-membered ring structure of the atoms belonging to the peptide group. As can be seen, the O−H and N−H distances are very stressed and longer by about 30% when compared to the corresponding O−H and N−H bonds in the stable states. Nonetheless, even in this very stretched position, the H atom displays a nonnegligible electron density around its nucleus, indicating a proper hydrogen transfer rather than a proton transfer; in fact, whenever a H^+ is transferred, the electronic cloud is left behind. Figure 7B shows the total electron density and the corresponding KS states around the ring structure in terms of contour plots. The value of the total electron density around the H atom that undergoes the displacement is still about 1/10 of its maximum, thus supporting the picture of hydrogen migration. The related KS states are spanned by 2s and 2p orbitals of atoms belonging to the ring structure. A Mulliken population analysis, performed on the ring structure, gave a value of +0.263 for this particular H, much smaller than the value for a bare proton. This is a clear indication that the hydrogen is not completely ionized in the four-membered TS.

A four-membered ring structure, very similar to the one found in our study, is also realized in the tautomerization of formamidic acid-to-formamide, which is the simplest peptide group model. The formamidic acid and formamide represent the analogous of the enol-cis (Figure 5d) and keto-cis (Figure 5f) of PGI, respectively. For the sake of completeness, we performed the calculation of the formamdic acid-formamide tautomerization via a force inversion method and verified that this approach is capable of finding the correct transition state. The estimated activation energy amounts to 29 kcal/mol for this reaction and turns out to be comparable to that in the PGI case. These results suggest that the common feature of a four-membered ring TS structure determines the value of the activation energy and represents a general rate-limiting step in this class of processes.

3.2.2. Characteristics of the Cis/Trans Isomerization Pathway. In this subsubsection, we focus the discussion on some relevant characteristics of the cis/trans isomerizations that show remarkable differences between the enol and keto peptide groups. In general, the cis/trans isomerization of the C−N peptide bond in simple amides taking the keto form is characterized both by modifications of the torsion angle $C_\alpha-C-N-C_\alpha$ and by changes in the C−N peptide bond: these two parameters vary with the isomerization, leading to twisting and breaking of the C−N (partial) double bond[4−7,16] and are crucial for an exhaustive description of the reaction mechanism. Indeed, this served as a guideline in the selection of the collective variables for the metadynamics simulations.

The distance of the C−N bond length does not undergo significant changes during the enol cis/trans isomerization. On the contrary, it changes dramatically in the keto case (Figure 8). Thus, we can argue that the breaking of the C−N (partial) double bond does not occur in the enol stage but in the keto one. The transition states for each isomerization show corresponding peculiar features. Namely, in the enol case, the $C-N-C_\alpha$ moiety assumes a linear conformation (Figure 5c), while the keto peptide group takes a pyramidal-twisted conformation (Figure 5g).

These differences in the cis/trans isomerizations are reflected in the KS occupied and unoccupied energy levels near the energy gap. Figure 9 shows these energy levels and the atomic-orbital characters of the corresponding KS states near the energy gap in the intermediate states. It is clear from the figure that the energy level of the N 2p-like KS state represents the highest occupied KS state (HOKS) at the transition state for the trans-to-cis isomerization of the enol (enol−linear). Instead, the level of the lowest unoccupied KS state (LUKS), having a ppπ* antibonding character, shifts down significantly along the energy axis at the transition state for the cis-to-trans isomerization of the keto (keto−twisted). As a result, a shrink of the energy gap is observed. Of course, these level shifts are related to the structural modifications affecting the system at each TS.

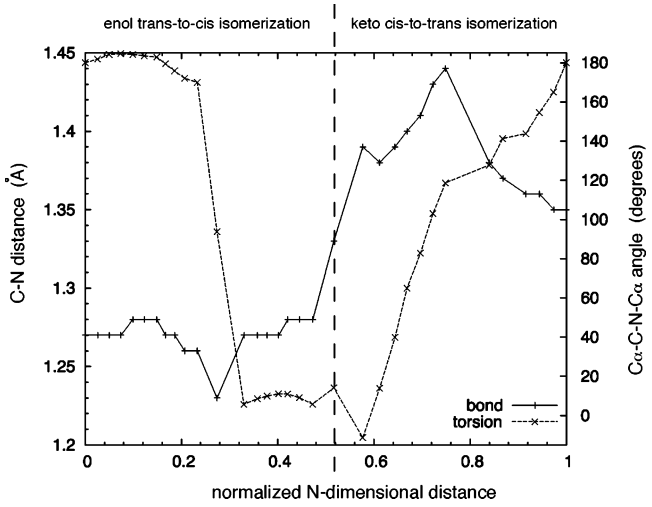

Figure 8. The variation of the C−N bond distance and the $C_\alpha-C-N-C_\alpha$ torsion angle along the cis/trans isomerization path. The dotted vertical line crosses the point corresponding to the transition state of the enol-to-keto transformation. The value 180° of the torsion angle is defined as a trans, while 0° is defined as a cis.

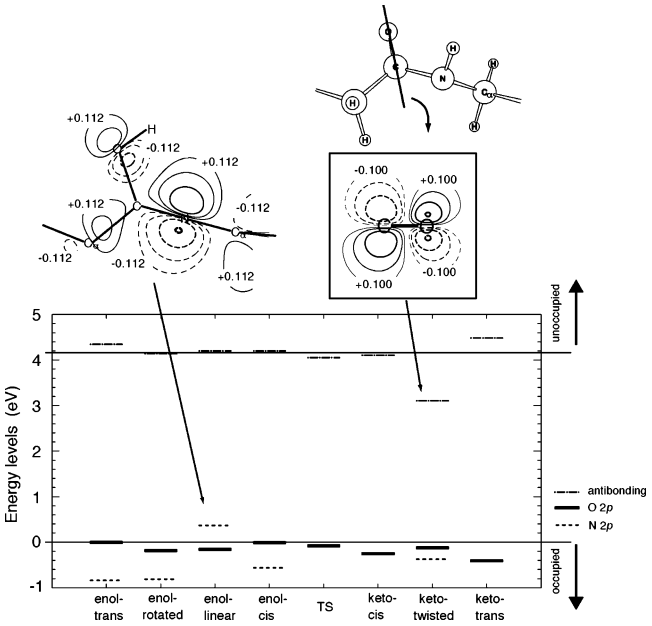

Figure 9. Energy levels and atomic-orbital characters of the KS states near the energy gap. The labels below the horizontal axis indicate the states shown in Figure 5 (B). All the O 2p- and N 2p-like states are occupied, while all the ppπ*-like antibonding states are unoccupied. The KS orbitals at the Γ point are shown as contour maps $(e/Å^3)^{1/2}$. Each contour denotes twice (or half) the values of the adjacent contour lines.

Let us first consider the enol case. As mentioned in the previous subsection, the electronic structure of the enol−trans peptide group originates mainly from the occupied KS state whose main character is the N 2p atomic orbital. However, this state is partly delocalized around the $C-N-C_\alpha$ moiety (Figure 4a) and its energy level (Figure 9) located below the energy gap. As a consequence, the HOKS state has the character of a nonbonding lone pair orbital belonging to the O atom in the adjacent keto peptide group.[36] The situation is the same in the rotated enol−trans and in the enol−cis. In the transition state of the enol cis/trans isomerization, however, the N 2p-like state is localized on the N atom in the $C-N-C_\alpha$ moiety, and it becomes the HOKS state (Figure 9).

In the (meta)stable states, the $C-N-C_\alpha$ takes a triangular conformation (Figure 5a, b, and d), and the N 2p atomic-orbital is thus hybridized with either the C or the C_α atomic orbital along the two bonds; thus, the N 2p-like states are delocalized along the bonds. This delocalization contributes to the decrease of the orbital energies of the N 2p-like states, so that they appear as states below the energy gap. In the transition state, on the other hand, the $C-N-C_\alpha$ is in a linear conformation (Figure 5c), and the C or C_α atomic-orbital does not hybridize with the N 2p orbital. Consequently, the N 2p-like state has a lone pair character.

Moving now to the keto case, we can observe that in all the intermediate states for the keto cis/trans isomerization, the LUKS orbitals have nonnegligible amplitudes on the N−C−O moieties, and they show a ppπ*-like antibonding states character. The compositions in terms of atomic orbitals are, however, different in the (meta)stable structures and in the transition states: In the keto−cis and keto−trans, p orbitals of N, C, and O atoms compose the LUKS state,[36] whereas, in the TS, p orbitals of C and O atoms constitute the LUKS (Figure 9). Furthermore, as expected, the energy level of the LUKS state in the TS is lower in energy than that in each (meta)stable state and accounts for the gap shrinking observed at the transition states.

In the (meta)stable states, the keto peptide groups including the N−C−O keep a planar configuration (Figure 5f and h), and thus the p orbital of C can interact with those of N and O. In the transition states, on the contrary, the keto peptide group assumes a pyramidal-twisted conformation (Figure 5g): The C−N distance is longer than those of the keto−cis and keto−trans by about 3.6% and 6.7%, respectively, while the C−O distance remains practically unchanged. In this structure, the p orbital of C interacts with that of O. Therefore, the LUKS orbitals in the (meta)stable states are composed by p orbitals of N, C, and O, while in the TS, the LUKS is formed by p orbitals of C and O only. For further insight, we examined, on the basis of the Hückel theory, the shift in the orbital energy level of the ppπ* antibonding LUKS state between the (meta)-stable and transition states.[55] We found that the orbital energy of the ppπ* antibonding state has a tendency to decrease its value as the absolute value of the resonance integral between the N and C p orbitals decreases. Hence, the energy level of the LUKS state consisting of the π orbitals of C and O is lower in energy than that of the LUKS including π orbitals of N, C, and O.

3.2.3. An Alternative Pathway for the Enol-to-Keto Tautomerism. Beside the pathway discussed in the previous subsubsections, we identified another possible reaction path via metadynamics simulations. Contrary to the former cis/trans isomerization study, in this simulation we selected only one collective variable. Dropping the torsion angle $C_\alpha-C-N-C_\alpha$, we used the N−H distance as a unique reaction coordinate. Since this collective variable accounts for the H migration and, hence, the rate-limiting step of the reaction, we aimed at checking whether it would be sufficient to describe the known reaction pathway. It turned out that this is not the case.

Figure 10 shows this new pathway as obtained in this simulation. The first thing that can be noticed is that the cis conformation does not occur. Nonetheless, similar to the case of the cis/trans isomerization, the rate-limiting step for this reaction is, again, the hydrogen migration from O to N, and the transition state takes a rhombic four-membered ring structure. Yet, it is realized in a trans conformation, and this makes the process energetically more demanding; in fact, the activation barrier is about 62 kcal/mol. For comparison, this

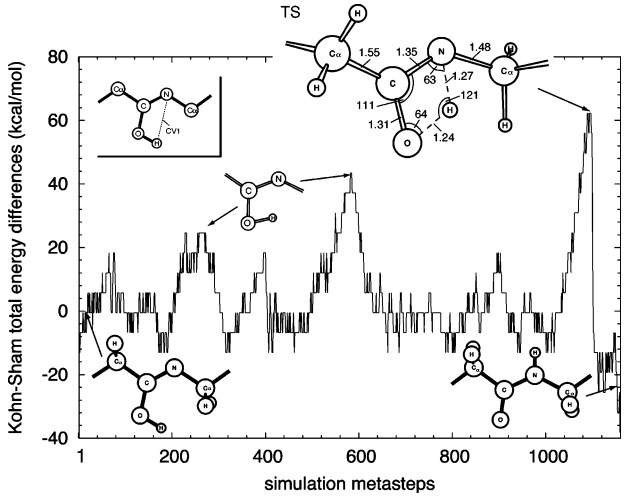

Figure 10. Energy profile and geometries of the intermediate states in the alternative pathway described in the text. In the energy profile, the assumed origin of energy axis is analogous to what is reported in Figure 5A-2). The upper left inset shows the collective variable chosen in this calculation. Distances and angles are expressed in Å and degrees, respectively. The peaks at ∼260 and ∼580 represent some of the attempts of the system to overcome the activation barrier during the exploration of the energy landscape.

value is more than a factor of two larger than in the previous cis/trans isomerization pathway. It is, then, difficult for this trans-pathway to become a favorable mechanism for the enol-to-keto tautomerism. Yet, it is a viable alternative, although its probability is very low due to the large energy barrier that has to be overcome. Looking at it from a more general perspective, this result provides further support to the discussed importance of the angular torsion in the tautomerization.

4. Conclusions

Summarizing, we inspected, via first-principle calculations, the tautomerization of an isolated polyglycine with infinite length (PGI) containing an enol peptide group (−C(OH)N-), which is a structural isomer of the keto peptide group (−CONH-). We found that the geometrical differences between the enol and keto forms are responsible for a switch of the double bond in the N−C−O moieties from N═C−O to N−C═O. The reaction pathways for the enol-to-keto tautomerism were sampled independently by constrained optimization and metadynamics with essential agreement. A distinctive feature is the cis/trans isomerization of the C−N peptide bond. The rate-limiting step of the reaction is a hydrogen (not proton) migration from an O to a N atom belonging to the peptide group, that proceeds upon the formation of a four-membered ring TS in the cis configuration, with an activation energy of about 26 kcal/mol. The mechanisms of the cis/trans isomerization result essentially different in the enol and in the keto forms. This difference is strongly reflected in the electron states close to the energy gap (HOKS and LUKS levels) of PGI. The present calculation, besides clarifying the distinctive features of the enol peptide groups, can provide some useful hints to understand the roles of the various phases of the process in physiological environments.

Acknowledgment. We gratefully acknowledge the fruitful discussions and precious suggestions from S. Okada, S. Berber, M. Otani, K. Chiba, M. Iannuzzi, M. Tateno, K. Muramoto, T. Tsukihara, and S. Yoshikawa. Computations were done at the Science Information Processing Center, University of Tsukuba, Institute for Solid State Physics, University of Tokyo, and at Research Center of the Computational Science, Okazaki National Institute. K.K. was supported by Research Fellowships of the Japan Society for the Promotion of Science for Young Scientists.

References and Notes

(1) Zabicky, J. *The Chemistry of Amides*; Wiley-Interscience: New York, 1970.
(2) Patai, S.; Rappoport, Z. *The Chemistry of amidines and imidates*, Volume 2; Wiley-Interscience: New York, 1991.
(3) Ramachandran, G. N.; Sasisekharan, V. *Adv. Prot. Chem.* **1968**, *23*, 283−437.
(4) Li, Y.; Garrell, R. L.; Houk, K. N. *J. Am. Chem. Soc.* **1991**, *113*, 5895−5896.
(5) Luque, F. J.; Orozco, M. *J. Org. Chem.* **1993**, *58*, 6397−6405.
(6) Olson, L. P.; Li, Y.; Houk, K. N.; Kresge, A. J.; Schaad L. J. *J. Am. Chem. Soc.* **1995**, *117*, 2992−2997.
(7) Lauvergnat, D.; Hiberty, P. C. *J. Am. Chem. Soc.* **1997**, *119*, 9478−9482.
(8) Kang, Y. K.; Park, H. S. *J. Mol. Struct. (THEOCHEM)* **2004**, *676*, 171−176.
(9) Perrin, C. L.; Lollo, C. P.; Johnston, E. R. *J. Am. Chem. Soc.* **1984**, *106*, 2749−2753.
(10) Perrin, C. L. *Acc. Chem. Res.* **1989**, *22*, 268−275.
(11) Wang, X. C.; Nichols, J.; Feyereisen, M.; Gutowski, M.; Boatz, J.; Haymet, A. D. J.; Simons, J. *J. Phys. Chem.* **1991**, *95*, 10419−10424.
(12) Aplincourt, P.; Bureau, C.; Anthoine, J.-L.; Chong, D. P. *J. Phys. Chem. A* **2001**, *105*, 7364−7370.
(13) Stewart, D. E.; Sarkar, A.; Wampler, J. E. *J. Mol. Biol.* **1990**, *214*, 253−260.
(14) Jabs, A.; Weiss, M. S.; Hilgenfeld, R. *J. Mol. Biol.* **1999**, *286*, 291−304.
(15) Fischer, G.; Aumüller, T. *Rev. Physiol. Biochem. Pharmacol.* **2003**, *148*, 105−150.
(16) Fischer, G. *Chem. Soc. Rev.* **2000**, *29*, 119−127.
(17) Svensson, A.-K. E.; O'Neill, J. C., Jr.; Matthews, C. R. *J. Mol. Biol.* **2003**, *326*, 569−583.
(18) Scherer, G.; Kramer, M. L.; Schutkowski, M.; Reimer, U.; Fischer, G. *J. Am. Chem. Soc.* **1998**, *120*, 5568−5574.
(19) Holtz, J. S. W.; Li, P.; Asher, S. A. *J. Am. Chem. Soc.* **1999**, *121*, 3762−3766.
(20) Schiene-Fischer, C.; Fischer, G. *J. Am. Chem. Soc.* **2001**, *123*, 6227−6231.
(21) Tsukihara, T.; Shimokata, K.; Katayama, Y.; Shimada, H.; Muramoto, K.; Aoyama, H.; Mochizuki, M.; Shinzawa-Itoh, K.; Yamashita, E.; Yao, M.; Ishimura, Y.; Yoshikawa, S. *Proc. Natl. Acad. Sci. U.S.A.* **2003**, *100*, 15304−15309.
(22) Yoshikawa, S.; Shinzawa-Itoh, K.; Nakashima, R.; Yaono, R.; Yamashita, E.; Inoue, N.; Yao, M.; Fei, M. J.; Libeu, C. P.; Mizushima, T.; Yamaguchi, H.; Tomizaki, T.; Tsukihara, T. *Science* **1998**, *280*, 1723−1729.
(23) Chiba, K. *Bioimaging* **2005**, *14*, 29−34 (in Japanese).
(24) Pommès, R.; Roux, B. *J. Phys. Chem.* **1996**, *100*, 2519−2527.
(25) Shimokata, K.; Katayama, Y.; Shimada, H. presented at the 42nd annual meeting of the Biophysical Society of Japan (unpublished).
(26) Hohenberg, P.; Kohn, W. *Phys. Rev.* **1964**, *136*, B864−B871.
(27) Kohn, W.; Sham, L. J. *Phys. Rev.* **1965**, *140*, A1133−A1138.
(28) Parr, R. G.; Yang, W. *Density Functional Theory of Atoms and Molecules*; Oxford University Press: New York, 1989.
(29) Payne, M. C.; Teter, M. P.; Allan, D. C.; Arias, T. A.; Joannopoulus, J. D. *Rev. Mod. Phys.* **1992**, *64*, 1045−1097.
(30) Miyamoto, Y.; Rubio, A.; Blase, X.; Cohen, M. L.; Louie, S. G. *Phys. Rev. Lett.* **1995**, *74*, 2993−2996.
(31) Okada, S.; Oshiyama, A.; Saito, S. *Phys. Rev. B* **2000**, *62*, 7634−7638.
(32) Takeda, K.; Shiraishi, K. *J. Phys. Soc. Jpn.* **1996**, *65*, 421−438.
(33) Sato, F.; Yoshihiro, T.; Era, M.; Kashiwagi, H. *Chem. Phys. Lett.* **2001**, *341*, 645−651.
(34) Improta, R.; Barone, V.; Kudin, K. N.; Scuseria, G. E. *J. Am. Chem. Soc.* **2001**, *123*, 3311−3322.
(35) Boero, M.; Terakura, K.; Tateno, M. *J. Am. Chem. Soc.* **2002**, *124*, 8949−8957.
(36) Kamiya, K.; Shiraishi, K.; Oshiyama, A. *J. Phys. Soc. Jpn.* **2004**, *73*, 3198−3208.
(37) Jeong, S.; Oshiyama, A. *Phys. Rev. Lett.* **1998**, *81*, 5366−5369.
(38) Otani, M.; Shiraishi, K.; Oshiyama, A. *Phys. Rev. B* **2003**, *68*, 184112−1−184112−7.
(39) Henkelman, G.; Uberuaga, B. P.; Jónsson, H. *J. Chem. Phys.* **2000**, *113*, 9901−9904.

(40) Tateyama, Y.; Ogitsu, T.; Kusakabe, K.; Tsuneyuki, S. *Phys. Rev. B* **1996**, *54*, 14994−15001.
(41) Iannuzzi, M.; Laio, A.; Parrinello, M. *Phys. Rev. Lett.* **2003**, *90*, 238302−1−238302−4.
(42) Car, R.; Parrinello, M. *Phys. Rev. Lett.* **1985**, *55*, 2471−2474.
(43) Boero, M.; Ikeshoji, T.; Liew, C. C.; Terakura, K.; Parrinello, M. *J. Am. Chem. Soc.* **2004**, *126*, 6280−6286.
(44) Boero, M.; Tateno, M.; Terakura, K.; Oshiyama, A. *J. Chem. Theory Comput.* **2005**, *1*, 925−934.
(45) Micheletti, C.; Laio, A.; Parrinello, M. *Phys. Rev. Lett.* **2004**, *92*, 170601−1−170601−4.
(46) Gervasio, F.; Laio, A.; Iannuzzi, M.; Parrinello, M. *Chem. Eur. J.* **2004**, *10*, 4846−4852.
(47) Ensing, B.; Klein, M. L. *Proc. Natl. Acad. Sci. U.S.A.* **2005**, *102*, 6755−6759.
(48) Codes used in the present work are based on Tokyo Ab initio Program Package (TAPP); Yamauchi, J.; Tsukada, M.; Watanabe, S.; Sugino, O. *Phys. Rev. B* **1996**, *54*, 5586−5603; Kageshima, H.; Shiraishi, K. *Phys. Rev. B* **1997**, *56*, 14985−14992; Sugino, O.; Oshiyama, A. *Phys. Rev. Lett.* **1992**, *68*, 1858−1861.
(49) Perdew, J. P.; Burke, K.; Ernzerhof, M. *Phys. Rev. Lett.* **1996**, *77*, 3865−3868.
(50) Vanderbilt, D. *Phys. Rev. B* **1990**, *41*, 7892−7895.
(51) CPMD, Copyright IBM Corp. 1990−2001, Copyright MPI für Festkörperforschung, Stuttgart, 1997−2004.
(52) Troullier, N.; Martins, J. L. *Phys. Rev. B* **1991**, *43*, 1993−2006.
(53) Lide, D. R. *CRC Handbook of Chemistry and Physics;* CRC Press: Boca Raton, FL, 2005.
(54) The electronegativity increases in the order of H, C, N, and O.
(55) We consider the simplest model consisting of a N−C−O linear configuration: The 2p atomic orbitals orthogonal to the atomic array are used as a basis set, and all the Coulomb integrals are set to match their related atomic levels. The resonance integral between the N and C π orbitals is treated as a variable, whereas the resonance integral between the C and O π orbitals is assumed to be constant: the C−O distance does not vary with the keto cis-to-trans isomerization, as stated in the text.