Protein hydrolysis mechanism of HIV-1 protease: Investigation by the *ab initio* MO calculations

Noriaki Okimoto,[*1] Masayuki Hata,[*2] Tyuji Hoshino,[*2] and Minoru Tsuda[*2]
[*1] *Computational Science Division, RIKEN*
[*2] *Laboratory of Physical Chemistry, Faculty of Pharmaceutical Sciences, Chiba University*

The protein hydrolysis mechanism by the HIV-1 protease (HIV-1 PR) was studied using *ab initio* molecular orbital calculations with a model compound. The initial model compound was constructed based on the result of a 100 ps molecular dynamics (MD) simulation of the enzyme-substrate (ES) complex under physiologic conditions. This study suggests that the hydrolysis mechanism of the Phe-Pro peptide bond specific to the HIV-1 PR consists of three elementary reactions: first, the reaction of the formation of the amide hydrate intermediate; second, the reaction of the protonation of the proline nitrogen of the substrate; and third, the reaction of the C-N bond cleavage of the substrate. The rate-determining step is the protonation of the substrate proline nitrogen and its activation energy is 23.95 kcal/mol. This result strongly suggests that protein hydrolysis by the HIV-1 PR occurs *in vivo*.

Overview of protein hydrolysis mechanism of HIV-1 protease

In the process of the replication of HIV type 1 (HIV-1), precursor polyproteins are transformed into structural proteins and replication enzymes by HIV-1 PR. HIV-1 PR belongs to the family of aspartic proteases, which have two aspartic acid residues at the catalytic site. The active form of the HIV-1 PR is a dimer of two identical polypeptide chains, each of which consists of 99 amino acid residues. Previous analysis of the proteolytic processing of the HIV-1 PR revealed that three of the cleavage sites contained Phe-Pro or Tyr-Pro at P1-P1' residues.[1)] Because it is unusual for mammalian endopeptidases to cleave the N-terminal of proline, elucidation of the hydrolysis mechanism of these peptide bonds by the HIV-1 PR is of great interest. This elucidation is also important for the development of effective HIV-1 PR inhibitors. In the present study, on the basis of the fact that the HIV-1 PR cleaves the peptide bond containing proline, we clarify the protein hydrolysis mechanism by HIV-1 PR through *ab initio* molecular orbital calculations.[2)]

In order to obtain the initial model compound for quantum chemical calculations, the MD simulation was performed on the ES complex that was constructed based on the X-ray crystallographic structure of the HIV-1 PR-inhibitor complex (code in Protein Data Bank: 7HVP[3)]). Figure 1(a) shows the average structure of a 100 ps MD simulation under physiologic conditions. The initial model compound for the quantum chemical calculations was constructed by extracting closed and open circle atoms from the structure shown in Fig. 1(a) and replacing open circle atoms with hydrogen atoms. This constructed initial model compound is shown in Fig. 1(b).

The hydrolysis mechanism in this study consists of three elementary reactions that are characterized by each transition state (TS1, TS2, TS3), as seen in the potential energy curve (Fig. 2). The structures of the stable states (S1, S2, S3, S4) and the transition states (TS1, TS2, TS3) that appear in the hydrolysis are shown in Figs. 3 and 4. These elementary reactions are described below.

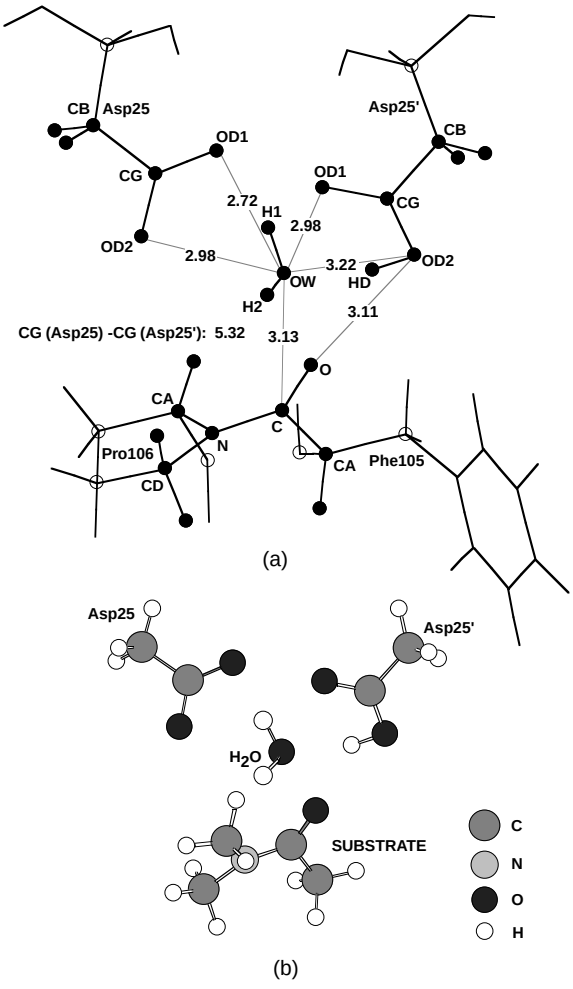

Fig. 1. (a) Catalytic site in the average structure of a 100 ps MD simulation for the ES complex at physiologic temperature (300 K). The used substrate molecule was Ac-Thr-Leu-Asn-Phe-Pro-Ile-Ser-NMe. The MD simulation was performed with the program package AMBER, Version 4.1. Numerals are the inter-atomic distances in Å. (b) Initial model compound. The initial model compound consists of an acetate/acetic acid pair for two catalytic-site Asp residues, a water molecule acting on the hydrolysis, and a dimethylacetamide for the substrate.

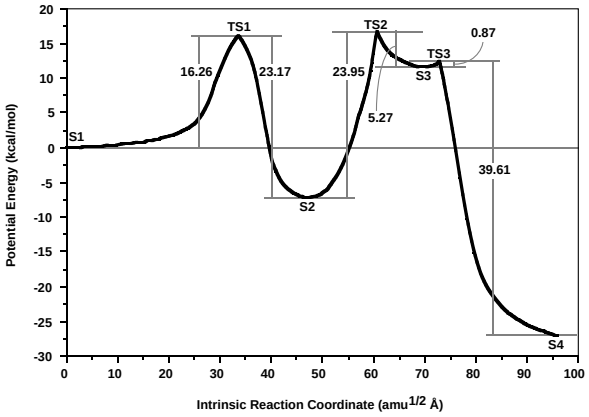

Fig. 2. Total potential energy curve for the hydrolysis of the Phe-Pro peptide bond by the HIV-1 PR. The ordinate is the potential energy differences (kcal/mol) and the abscissa is the intrinsic reaction coordinate (amu$^{1/2}$ Å). The Schrödinger equation of the model compound was solved with the Hartree-Fock method using the 6-31G** basis set. The model structures at the minimum points and transition states on the potential energy hypersurface were fully optimized using the energy gradient method. The computational program used was Gaussian 94. In all optimized structures, some constraints were applied to prevent the corruption of the basic conformation of the ES complex.

The first step of the hydrolys mechanism is the reaction of the formation of the amide hydrate intermediate from ES complex. The structure of the transition state TS1 (Fig. 3) was determined from the initial model compound constructed based on the result of the MD simulation of the ES complex under physiologic conditions. This structure is obtained theoretically through the geometry optimization to search a saddle point on the potential energy hypersurface by solving the Schrödinger equation. The normal vibrational mode of the imaginary frequency indicates the attack of the water molecule on the carbonyl carbon of the substrate. A steep energy drop in the direction of the separation of the OW atom of the water molecule and the C atom (carbonyl carbon) of the substrate transforms the structure of TS1 to the structure of the initial state of the hydrolysis S1; that is, the ES complex (Fig. 3) The structure of the ES complex (S1, Fig. 3) obtained theoretically from the structure of TS1 is very similar to the structure shown in Fig. 1(a). This means that the model compound represents the catalytic site of ES complex. In contrast, a steep energy drop in the direction of the attack of the OW atom of the water molecule on the C atom of the substrate transforms the structure of TS1 to the structure of the intermediate S2, that is, the enzyme-bound amide hydrate intermediate (Fig. 3). This reaction is supported by the isotope kinetic experiments.[4] The activation energy of this reaction is 16.26 kcal/mol (Fig. 2).

The second step of the hydrolysis mechanism is the reaction of the protonation of the proline nitrogen of the substrate. Regarding the transition state TS2 (Fig. 3), the normal vibrational mode of the imaginary frequency indicates the transfer

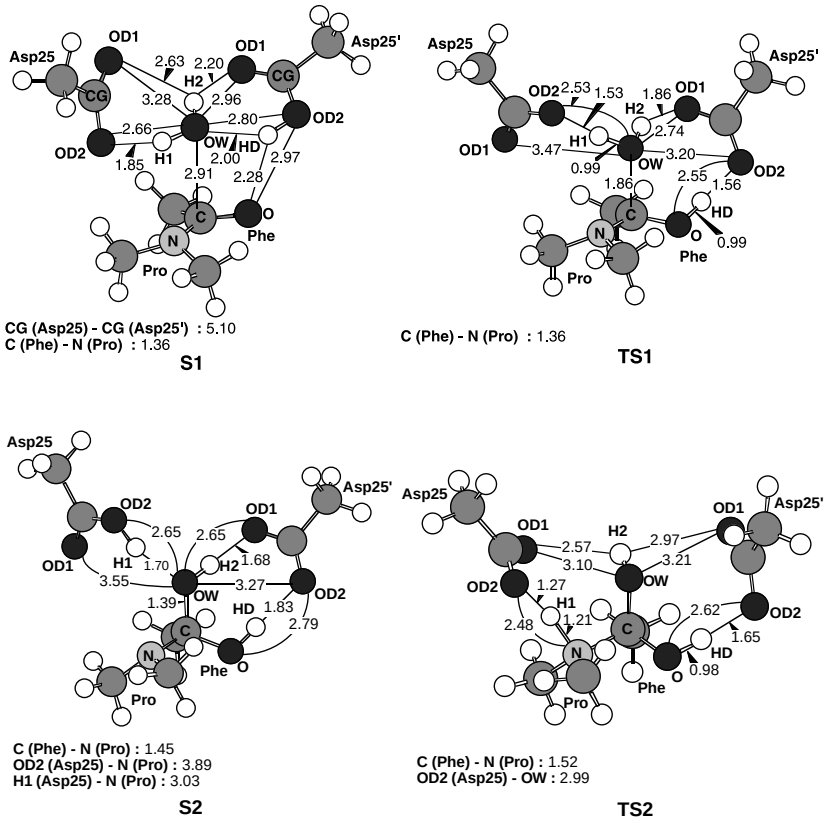

Fig. 3. The structures of the stable state (S1 and S2) and the transition state (TS1 and TS2) which appear in the hydrolysis. Numerals are the inter-atomic distances in Å.

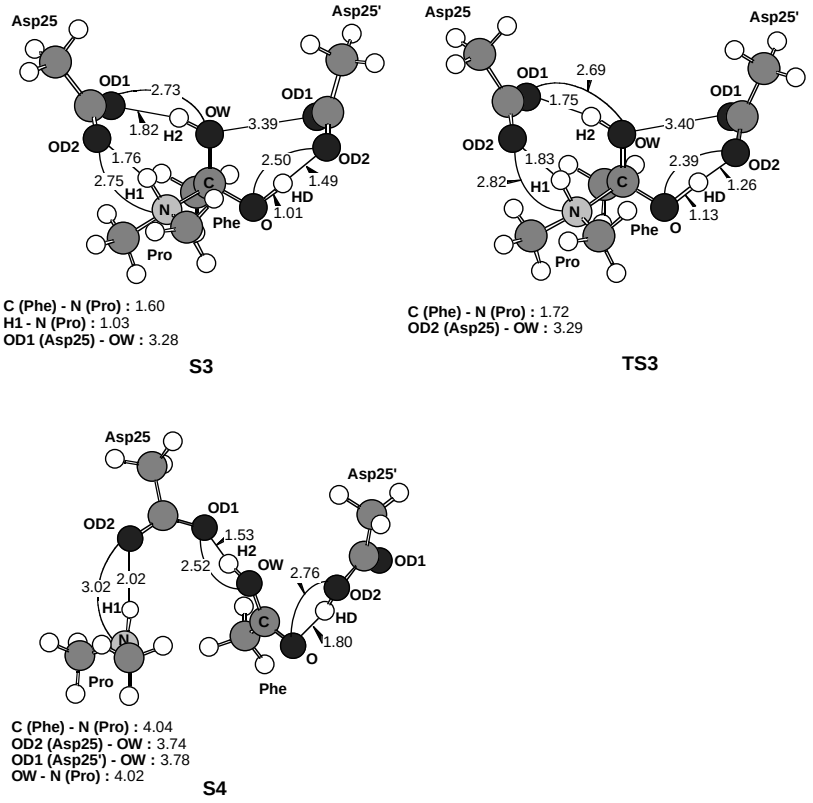

Fig. 4. The structures of the stable state (S3 and S4) and the transition state (TS3) which appear in the hydrolysis. Numerals are the inter-atomic distances in Å.

of the H1 atom to the N atom of the substrate and the H2 atom approaching the OD1 atom of Asp25. A steep energy drop in the direction of the separation of the H1 atom and the N atom of the substrate transforms the structure of TS2 to the structure of the intermediate S2, the product of the first elementary reaction (Fig. 3). This result reveals that TS2 is connected to the first elementary reaction. In contrast, a steep energy drop in the direction of the transfer of the H1 atom to the N atom of the substrate transforms the structure of TS2 to the structure of the intermediate S3 (Fig. 4), in which the N atom of the substrate is protonated. This reaction is supported by isotope kinetic experiments.[4] This reaction is the rate-determining step for the hydrolysis mechanism and its activation energy is 23.95 kcal/mol (Fig. 2). One reason that this reaction is the rate-determining step might be the presence of the proline in the amide hydrate intermediate, that is, the proline would cause the steric hindrance for the protonation reaction.

The third step of the hydrolysis mechanism is the reaction of the C-N bond cleavage of the substrate. In the structure of the transition state TS3 (Fig. 4), the normal vibrational mode of the imaginary frequency indicates the C-N bond cleavage and the transfer of the HD atom to the OD2 atom of Asp25'. A steep energy drop in the direction of the shrinkage of the C-N distance transforms the structure of TS3 to the structure of the intermediate S3, the product of the second elementary reaction (Fig. 4). The result reveals that the transition state TS3 is connected to the second elementary reaction. In contrast, a steep energy drop in the direction of the C-N bond cleavage transforms the structure of TS3 to the structure of the final state of the hydrolysis S4 (Fig. 4). The activation energy of this reaction is 0.87 kcal/mol (Fig. 2). In the reaction of the C-N bond cleavage of the substrate, the final state of the hydrolysis S4 is more stable than the transition state TS3 by 39.61 kcal/mol so that this reaction is not readily reversed. The irreversibility of the reaction of the C-N bond cleavage is supported by the isotope kinetic experiments.[5]

We conclude the following from this study. First, the hydrolysis mechanism of the Phe-Pro peptide bond specific to the HIV-1 PR consists of three elementary reactions. These reactions are supported by the experimental data. Next, the rate-determining step of the protein hydrolysis mechanism by the HIV-1 PR is the process of the protonation of the proline nitrogen of the peptide bond. This activation energy is 23.95 kcal/mol, which suggests that this hydrolysis occurs *in vivo*.

References

1) P. L. Darke et al.: Biochem. Biophys. Res. Commun. **156**, 297 (1988).
2) N. Okimoto et al.: J. Am. Chem. Soc. **121**, 7349 (1999).
3) A. L. Swain et al.: Proc. Natl. Acad. Sci. USA **87**, 8805 (1990).
4) T. D. Meek et al.: Methods Enzymol. **241**, 127 (1994).
5) L. J. Hyland et al.: Biochemistry **30**, 8441 (1991).