BIOORGANIC CHEMISTRY 11, 322–327 (1982)

The Role of the Environment around the Catalytic Triad in β-Trypsin: A Molecular Orbital Study

SETSUKO NAKAGAWA AND HIDEAKI UMEYAMA[1]

School of Pharmaceutical Sciences, Kitasato University, Shirokane, Minato-ku, Tokyo 108, Japan

Received December 21, 1981

In the enzymatic reaction of β-trypsin the role of environment around the catalytic triad is studied by means of *ab initio* molecular orbital calculations. The triple ion form of the catalytic triad (Asp 102(−)–His 57(+)–Ser 195(−)) is considerably more stable than the double proton-transferred form (Asp 102(neutral)–His 57(neutral)–Ser 195(−)), due to the environment around it, rather than its nature. The "electrostatic mechanism" is more favorable than the "charge relay mechanism" owing to the nature of the enzyme as a biopolymer.

The presence of a precisely arranged catalytic triad consisting of the carboxyl group of Asp, the imidazole group of His, and the hydroxyl group of Ser have been found for chymotrypsin (*1*), trypsin (*2*), elastase (*3*), streptomyces griseus protease B (*4*), and α-lytic protease (*5*). Although Blow *et al.* (*6*) had proposed a "charge relay mechanism" in which the aspartate acts as the ultimate base holding a proton during the catalysis, the experiments of low-field ^{1}H nmr (*7*), C$^{\epsilon 1}$H nmr of the imidazole ring (*8*), and ^{15}N nmr (*9*) have not supported this hypothesis (*6*).

The theoretical *ab initio* molecular orbital (MO) calculations at the double-zeta level (*10–13*) also have not supported the charge relay hypothesis. On the basis of the potential energy map of the proton transfers from Ser 195 to His 57 and from His 57 to Asp 102 (α and β proton transfers) in β-trypsin, Nakagawa *et al.* reported that the triple ion form (Asp 102(−)–His 57(+)–Ser 195(−)) is more stable than the double-proton-transferred form (Asp 102(neutral)–His 57(neutral)–Ser 195(−)) (*10*). This result was supported by the calculation for a model system consisting of methanol, imidazole, and formic acid, in which the optimizations of the monomers and the intermolecular distances of the dimeric complex were carried out with an STO-3G basis set (*13, 14*). Umeyama *et al.* proposed an "electrostatic mechanism" in which the anion form of Asp 102 plays a role in lowering the barrier height of α-proton transfer due to the electrostatic interaction (*12*). It is expected that the change of the charge state of the catalytic triad with the proton transfers is greatly affected by the environment around it. In order to clarify the charge state in the real enzyme system, therefore, the hydrogen bond and van der Waals' interactions around the catalytic triad and the electrostatic effects in the enzyme should be included in the calculation (Fig. 1). In addition to

[1] To whom correspondence should be addressed.

0045-2068/82/030322-06$02.00/0

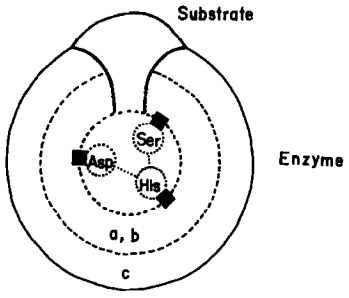

FIG. 1. Schematic representation of interactions between the catalytic triad and its environment. a, Hydrogen bond; b, van der Waals' interaction; c, electrostatic interaction.

the effects of the hydrogen bonds around the catalytic triad and 26 ionic amino acid residues (*12, 13*), the van der Waals' interactions and the electrostatic effect of the main chain for the proton transfers in the catalytic triad were estimated individually in this paper.

Molecular orbital calculations were carried out within the closed-shell LCAO SCF approximation using *ab initio* MO method. The IMSPACK program (*15, 16*) which was modified for the large number of point charges by the authors was used for the calculations. A 4-31G basis set (*17*) was used.

As the model of Michaelis–Menten complex, the X-ray data for the bovine pancreatic trypsin inhibitor (PTI)-trypsin complex given by Huber and co-workers (*18*) were used. Atomic coordinates, except for the hydrogens, were obtained from the Protein Data Bank (*19*). The model structures used for the calculations are shown in Table 1. Methanol (Me), imidazole (Im), and formic acid (Fo) were used as the side-chain models of Ser 195, His 57, and Asp 102, respectively. The coordinates of the hydrogens were determined using the optimized geometries of these molecules (*12*). The proton transfer from Ser 195 O^{γ} to His 57 $N^{\epsilon 2}$ is called α-proton transfer, and that from His 57 $N^{\delta 1}$ to Asp 102 $O^{\delta 2}$ is called β-proton transfer. Three states of the catalytic triad are labeled A (^{195}Me (neutral)–^{57}Im (neutral)–^{102}Fo(−)), B (triple ion form), and D (Double-proton-transferred form). The relative energies of model I, $dE_{BA(I)}$ and $dE_{DB(I)}$, are defined as

$$dE_{BA(I)} = E_B - E_A \quad \text{and} \quad dE_{DB(I)} = E_D - E_B,$$

where E_A, E_B, and E_D are total energies at the A, B, and D states, respectively. Since the model structure is oversimplified in model I, the effects of the remaining moieties of Asp 102, His 57, and Ser 195 for the proton transfers are corrected in model II by using larger molecules. The molecules represented by Mep, Imp, and Fop(−) in Table I correspond to $NH_2COCH(CH_2OH)NHCOH$, $HCOCH(CH_2N_2C_3H_3)NH_2$, and $HCOCH(CH_2CO_2^-)NH_2$, respectively. The relative energies in which the corrections by the larger molecules are performed are estimated in an assumption of additivity as

$$dE_{BA(II)} = dE_{BA(I)} + dE_{BA(II-2)} + dE_{BA(II-3)} - 2dE_{BA(II-1)}$$

and

$$dE_{DB(II)} = dE_{DB(I)} + dE_{DB(II-5)} + dE_{DB(II-6)} - 2dE_{DB(II-4)}.$$

TABLE 1

VARIOUS MODEL STRUCTURES

Model	Molecules[a]	
I[b]	^{195}Me–^{57}Im–^{102}Fo(−)	Side chain of Ser 195, His 57, and Asp 102
II-1	^{195}Me–^{57}Im–^{102*}Fo(−)	Correction of the remaining moiety of Ser 195, His 57, and Asp 102
II-2	^{195}Me–^{57}Imp–^{102*}Fo(−)	
II-3	^{195}Mep–^{57}Im–^{102*}Fo(−)	
II-4	^{195*}Me(−)–^{57}Im(+)–^{102}Fo(−)	
II-5	^{195*}Me(−)–^{57}Imp(+)–^{102}Fo(−)	
II-6	^{195*}Me(−)–^{57}Im(+)–^{102}Fop(−)	
III-1	I, $^{214}CH_3OH$	Side chain of Ser 214 and —NH— of Ala 56
III-2	I, $^{56}NH_3$	
IV-1	II-1,$^{42}CH_3S$—$^{58}SCH_3$	Disulfide bridge of Cys 42–Cys 58 and peptide moieties of Ser 213–Trp 214 and Val 214–Ser 215
IV-2	II-1, ^{213}HCO—$^{214}NH_2$	
IV-3	II-1, ^{214}HCO—$^{215}NH_2$	
IV-4	II-4, $^{42}CH_3S$—$^{58}SCH_3$	
IV-5	II-4, ^{213}HCO—$^{214}NH_2$	
IV-6	II-4, ^{214}HCO—$^{215}NH_2$	
V-1	III-1, #IAA, $^{15(I)}$#Lys	Ionic amino acid residues (IAA), Lys 15 (I), and main chain (MCR)
V-2	II-1, *MCR	
V-3	II-4, *MCR	
VI	I, $^{15(I)}$HCO—$^{16(I)}NH_2$	Peptide moiety of Lys 15 (I)–Ala 16 (I)
VII[c]	CH_3OH—$C_3N_2H_4$—HCO_2(−)	Geometry optimized system

[a] The fragments labeled by * and # are approximated by point fractional charges and integral charges in the calculation, respectively.

[b] The distances between Ser 195 O^{γ} and His 57 $N^{\epsilon 2}$ (R_1) and between His 57$^{\delta 1}$ and Asp 102 $O^{\delta 2}$ (R_2) are 2.68 and 2.72 Å, respectively.

[c] R_1 in A, B, and D states are 2.82, 2.49, and 2.56 Å, respectively. R_2 in A, B, and D states are 2.73, 2.53, and 2.72 Å, respectively.

The effects of the hydrogen bond between $O^{\delta 1}$ of Asp 102 and —NH— group of His 57 and the van der Waals' contacts of the backbone moieties are included in $dE_{BA(II)}$ and $dE_{DB(II)}$. In model III the effects of the hydrogen bonds between O^{γ} of Ser 214 and $O^{\delta 2}$ of Asp 102 and between —NH— group of Ala 56 and $O^{\delta 1}$ of Asp 102 are estimated (*12*). The relative energies in which the effects of the hydrogen bonds are included are estimated in the assumption of additivity as

$$dE_{BA(III)} = dE_{BA(II)} + dE_{BA(III-1)} + dE_{BA(III-2)} - 2dE_{BA(I)}$$

and

$$dE_{DB(III)} = dE_{DB(II)} + dE_{DB(III-1)} + dE_{DB(III-2)} - 2dE_{DB(I)}.$$

In model IV the effects of the disulfide bridge of Cys 42–Cys 58, the peptide moiety of Ser 214–Trp 215, and the peptide moiety of Val 213–Ser 214 which

interact with His 57 near van der Waals' distances are estimated. In model V the electrostatic effects of the ionic amino-acid residues and the backbone are estimated. Twenty-six ionic amino-acid residues were included in the SCF calculations as integral charges (*12*). The backbones which were not included in the paper (*12*) were included in the SCF calculations as the point fractional charges, as shown by Nakagawa and Umeyama (*20*). In model VI the effect of the peptide moiety of Lys 15(I)–Ala 16(I) as a substrate is estimated. $dE_{BA(IV)}$, $dE_{DB(IV)}$, $dE_{BA(V)}$, $dE_{DB(V)}$, $dE_{BA(VI)}$, and $dE_{DB(VI)}$ were calculated by using the same method described above. Model VII is a complex consisting of methanol, imidazole, and formic acid, which is independent of the structure of β-trypsin. The 4-31G geometry-optimized molecules of formic acid and methanol were used for the calculations (*21, 22*). For imidazole r(NH) and r(CH) were optimized with the 4-31G basis set (*21*). The optimizations of the intermolecular distances of the dimeric and trimeric complexes were carried out with the 4-31G basis set.

The potential energy levels for the B and D states relative to the A state are shown in Fig. 2. When only the side chains of the catalytic triad were included in the calculation (model I), the energy level of B is higher than that of A by 25 kcal/mol, and the energy level of D is higher than that of B by 13 kcal/mol (*10*). In model II the energy level of B decreases; the perturbed effect of the peptide moiety of Ser 195 dominantly facilitates α-proton transfer by 6 kcal/mol (24.6–19.0 kcal/mol). On the other hand, β-proton transfer is dominantly inhibited by the perturbed effect of the peptide moiety of His 57 by 5 kcal/mol (18.2–13.3 kcal/mol). When the hydrogen bonds of Ser 214 and Ala 56 for the side chain of Asp 102 are also included (model III), the difference between the energy levels of B and D increases; and when van der Waals' interactions around His 57 and the electrostatic interactions are additionally included, this difference increases again (models IV and V). When the model substrate of formamide is also included, the energy level of D is higher than that of B by 37 kcal/mol (model VI). On the basis of the results in models I–V, when the effects of various moieties of the environment around the catalytic triad are gradually included in the calculations, the difference between the energy levels of B and D increases step by step. Therefore the double-proton-transferred form is considerably more unstable than the triple ion form owing to the environment of the enzyme.

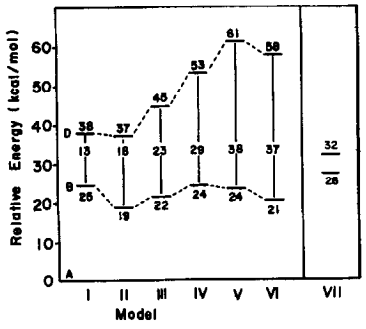

FIG. 2. Energy levels of B and D states relative to A state in various models.

The difference between the energy levels of B and D indicates which mechanism of the "charge relay mechanism" and the "electrostatic mechanism" is appropriate in the enzymatic reaction. If the "charge relay system" is supported, the result will be $D < B$; and if the "electrostatic mechanism" is supported, it will be $D > B$. The result of $D \gg B$ in this paper supports the "electrostatic mechanism" more strongly than that reported by Umeyama *et al.* (*12*).

In connection with the influence of the environment around the catalytic triad, the energy levels of B and D in the isolated formic acid–imidazole–methanol system independent of the PTI–trypsin structure were calculated in the process of geometry optimization. The energy level of D was higher than that of B (higher by 4 kcal/mol[2]), as shown in model VII in Fig. 2. The value is smaller than that of model I because of the optimizations of the monomers and the intermolecular distances. The value of 4 kcal/mol is very small in comparison with the value of 38 kcal/mol in model V. Thus, the triple ion form which is significant in the enzymatic reaction is more stable than the double-proton-transferred form, owing to the influence of the environment around the catalytic triad, not due to only its intrinsic nature. In other words, the "electrostatic mechanism" is more favorable than the "charge relay mechanism" owing to the nature of the enzyme as a biopolymer.

ACKNOWLEDGMENTS

The authors are grateful to Professor I. Moriguchi of this university for his support. Numerical calculations were carried out with a HITAC M-200H computer at the Computer Center of Institute for Molecular Science.

REFERENCES

1. B. W. MATTHEWS, P. B. SIGLER, R. HENDERSON, AND D. M. BLOW, *Nature* (*London*) **214,** 652 (1967).
2. H. FEHLHAMMER AND W. BODE, *J. Mol. Biol.* **98,** 683 (1973).
3. H. C. WATSON, D. M. SHOTTON, J. M. COX, AND H. MUIRHEAD, *Nature* (*London*) **225,** 806 (1970).
4. L. T. J. DELBAERE, W. L. B. HUTCHEON, M. N. G. JAMES, AND W. E. THIESSEN, *Nature* (*London*) **257,** 758 (1975).
5. G. D. BRAYER, L. T. J. DELBAERE, AND M. N. G. JAMES, *J. Mol. Biol.* **131,** 743 (1979).
6. D. M. BLOW, J. J. BIRKTOFT, AND B. S. HARTLEY, *Nature* (*London*) **221,** 337 (1969).
7. G. ROBILLARD AND R. G. SHULMAN, *J. Mol. Biol.* **86,** 541 (1974).
8. J. L. MARKLEY AND M. A. PORUBCAN, *J. Mol. Biol.* **102,** 487 (1976).
9. W. W. BACHOVCHIN AND J. D. ROBERTS, *J. Amer. Chem. Soc.* **100,** 8041 (1978).
10. S. NAKAGAWA, H. UMEYAMA, AND T. KUDO, *Chem. Pharm. Bull.* **28,** 1342 (1980).
11. H. UMEYAMA AND S. NAKAGAWA, *Chem. Pharm. Bull.* **28,** 2292 (1980).
12. H. UMEYAMA, S. NAKAGAWA, AND T. KUDO, *J. Mol. Biol.* **150,** 409 (1981).
13. P. A. KOLLMAN AND D. M. HAYES, *J. Amer. Chem. Soc.* **103,** 2955 (1981).
14. D. M. HAYES AND P. A. KOLLMAN, "Catalysis in Chemistry and Biochemistry: Theory and Experiment" (B. Pullman, Ed.), pp. 77. Reidel, Dor, Boston, Mass., 1979.
15. K. MOROKUMA, S. KATO, K. KITAURA, I. OHMINE, AND S. SAKAI, An *ab initio* program package consisting of Gaussian 70 and many additional routines, unpublished.

[2] 7 kcal/mol was reported by Kollman and Hayes (*13*).

16. W. J. HEHRE, W. A. LATHAN, R. DITCHFIELD, M. D. NEWTON, AND J. A. POPLE, Gaussian 70, Quantum Chemistry Program Exchange, Indiana Univ. (1974).
17. R. DITCHFIELD, W. J. HEHRE, AND J. A. POPLE, *J. Chem. Phys.* **54,** 724 (1971).
18. R. HUBER, D. KUKLA, W. BODE, P. SCHWANGER, K. BARTELS, J. DEISENHOFER, AND W. STEIGEMANN, *J. Mol. Biol.* **89,** 73 (1974).
19. F. C. BERNSTEIN, T. F. KOETZLE, G. J. B. WILLIAMS, E. F. MEYER, JR., M. D. BRICE, J. R. RODGERS, O. KENNARD, T. SHIMANOUCHI, AND M. TASUMI, *J. Mol. Biol.* **112,** 535 (1977).
20. S. NAKAGAWA AND H. UMEYAMA, *J. Theoret. Biol.* **96,** in press (1982).
21. S. NAKAGAWA AND H. UMEYAMA, *Chem. Pharm. Bull.* **29,** 1 (1981).
22. H. UMEYAMA AND S. NAKAGAWA, *Chem. Pharm. Bull.* **25,** 1671 (1977).