J. Biochem. Mol. Biol. Vol. 29, No. 4 pp. 380~385 July 31, 1996

Intramolecular Hydrogen Bonds in Proteinase Inhibitor Protein, A Molecular Dynamics Simulation Study

Hye-Shin Chung

Department of Microbiology, Hannam University, Taejon 300-791, Korea

(Received April 17, 1996)

Abstract: Ovomucoid third domain is a serine proteinase inhibitor protein which consists of 56 amino acid residues. A fifty picosecond molecular dynamics (MD) simulation was carried out for ovomucoid third domain protein with 5 Å layer of water molecules. A comparison of main chain atoms in the MD averaged structure with the crystal structure showed that most of the backbone structures are maintained during the simulation. Investigation of the intramolecular hydrogen bondings indicated that most of the interactions between main chain atoms were conserved, whereas those between side chains were reorganized for the period of the simulation. Especially, the side chain interactions around the scissile bond of reactive site P1 (Met18) were found to be more extensive for the MD structures. During the simulation, hydrogen bonds were maintained between the side chains of Glu19 and Arg21 as well as those of Thr17 and Glu19. Extensive side chain interactions observed in the MD structures may shed light on the question of why protein proteinase inhibitors are strong inhibitors for proteinases rather than good substrates.

Key words: molecular dynamics, ovomucoid third domain, proteinase inhibitor.

Proteinase inhibitors are important in regulation of the activities of the proteolytic enzymes in cells or tissues and prevent uncontrolled proteolysis. Inhibition by proteinase inhibitors is essentially irreversible and involves the formation of a tight noncovalent 1 : 1 enzyme-inhibitor complex, which resembles a Michaelis complex (Laskowski and Kato, 1980). Ovomucoid is the major constituent of avian egg white, and is made up of three homologous domains. The three domains can be divided and isolated by limited proteolysis of the linker peptides. Each of the three domains is a strong inhibitor for proteolytic enzymes, among which the third domain is the most extensively studied of the three (Laskowski *et al.*, 1987). Avian ovomucoid third domain (OM3) is a Kazal-type serine proteinase inhibitor, which consists of 56 amino acid residues and three invariable disulfide bonds. Sequences have been determined for ovomucoid third domains from over 100 avian species (Laskowski *et al.*, 1987), and their inhibitory properties compared (Park, 1985; Empie and Laskowski, 1987). The primary structures of some ovomucoid third domains are shown in Fig. 1. Ovomucoid third domain from turkey (OMTKY3) has the same sequence as that from silver pheasant (OMSVP3) except for the P1 residue, whereas that from Japanese quail (OMJPQ3) is more different from OMSVP3. The P1 residue in OMJPQ3 is lysine, which makes it an inhibitor for trypsin-like enzymes, while OMSVP3 and OMTKY3 are inhibitors for chymotrypsin-type proteinases. A third domain binds to the enzyme in the same manner as a sub-

```
                          P12                 P7                  P4
        1                 5                   10                  15
OMSVP3  Leu Ala Ala Val Ser Val Asp Cys Ser Glu Tyr Pro Lys Pro Ala
OMTKY3  Leu Ala Ala Val Ser Val Asp Cys Ser Glu Tyr Pro Lys Pro Ala
OMJPQ3  Leu Ala Ala Val Ser Val Asp Cys Ser Glu Tyr Pro Lys Pro Ala

              P1  P1'  P2'              P7'                 P12'
       16             20                  25                  30
OMSVP3  Cys Thr Met Glu Tyr Arg Pro Leu Cys Gly Ser Asp Asn Lys Thr
OMTKY3  Cys Thr Leu Glu Tyr Arg Pro Leu Cys Gly Ser Asp Asn Lys Thr
OMJPQ3  Cys Pro Lys Asp Tyr Arg Pro Val Cys Gly Ser Asp Asn Lys Thr

                          P17'                P22'                P27'
       31                 35                  40                  45
OMSVP3  Tyr Gly Asn Lys Cys Asn Phe Cys Asn Ala Val Val Glu Ser Asn
OMTKY3  Tyr Gly Asn Lys Cys Asn Phe Cys Asn Ala Val Val Glu Ser Asn
OMJPQ3  Tyr Gly Asn Lys Cys Asn Phe Cys Asn Ala Val Val Gue Ser Asn

                          P32'                P37'
       46                 50                  55  56
OMSVP3  Gly Thr Leu Thr Leu Ser His Phe Gly Lys Cys
OMTKY3  Gly Thr Leu Thr Leu Ser His Phe Gly Lys Cys
OMJPQ3  Gly Thr Leu Thr Leu Asn His Phe Gly Lys Cys
```

Fig. 1. Primary structures of ovomucoid third domain from silver pheasant (OMSVP3), from turkey (OMTKY3) and from Japanese quail (OMJPQ3). Amino acid residues are numbered. Nomanclature of Schechter and Berger (1967) is also used for the numbering of the substrate amino acid residues towards the N-terminal and the C-terminal directions from the scissile bond.

*To whom correspondence should be addressed.
Tel : 82-42-629-7934, Fax : 82-42-629-7444.

strate, but the energy barrier for hydrolysis is large and unfavorable, resulting in extremely low rates of hydrolysis (Laskowski and Kato, 1980). Crystal structures of the isolated ovomucoid third domains have been determined (Weber *et al.*, 1981; Bode *et al.*, 1985), as well as those complexed with cognate proteinases (Read *et al.*, 1983; Fujinaga *et al.*, 1987; Huang *et al.*, 1995). The structural data in solution, however, are limited to a few 2D-NMR studies (Robertson *et al.*, 1988; Krezel *et al.*, 1994). The structure of an ovomucoid domain contains three disulfide bridges, a long α-helix, and a triple stranded antiparallel β-sheet. The contact area in the inhibitor with the proteinase is usually located on an external loop, and there are many intramolecular interactions that stabilize the binding loop. These intramolecular interactions are thought to reduce the rate of cleavage, which distinguishes the inhibitors from the substrates.

Here, I present some of the results from a molecular dynamics simulation of an ovomucoid third domain. In this study the three-dimensional structure of the ovomucoid third domain from silver pheasant was examined and analyzed using molecular dynamics simulation. The simulation data indicate that the simulated structures retain the interaction between the side chains of Glu19 and Thr17, which exists in the crystal structure, as well as the newly-formed interaction between the side chains of Glu19 and Arg21. Intramolecular interactions in the simulated structure are compared with crystal data and thermodynamic data and the significance of some of hydrogen bonds will be discussed.

Methods

The entire simulation was done on IRIS Indy computer (Silicon Graphics; Mountain View, USA) using the Insight/Discover program provided by Biosym Technology (San Diego, USA). The consistent valence force field (CVFF) parameters (Hagler *et al.*, 1979) were used for the protein atoms in conjunction with the TIP3P water molecules. Nonbonded interactions were calculated for the pairs out to a cutoff distance of 9.5 Å (Kitson and Hagler, 1988), and a dielectric constant of 1 was used throughout the calculation.

The initial coordinates for the third domain of silver pheasant ovomucoid (OMSVP3) were obtained from the crystal structure (Bode *et al.*, 1985) deposited in the Brookhaven Protein Data Bank (2ovo). Solvation was done by adding the equilibrated water molecules so that a solvation layer of 5 Å thickness around the atoms of the protein molecule was formed. Initially, the system was optimized by steepest descent minimization until the maximum derivative was less than 1 kcal/mol/Å, followed by conjugate gradient cycles (Fletcher, 1988) until the maximum derivative was less than 0.001 kcal/mol/Å. Molecular dynamics simulation of the energy-minimized structure was done at 27℃. The simulation was initialized by assigning random velocities of atoms in the molecular system according to the Maxwell-Boltzmann distribution at 300 K. The temperature control during the simulation was achieved by velocity rescaling, and Verlet algorithm was used for integrations with a time step of 1 fs. A total time of 50 ps was covered in the simulation, during which the coordinates, velocities, and energies were saved every 0.5 ps for further analysis.

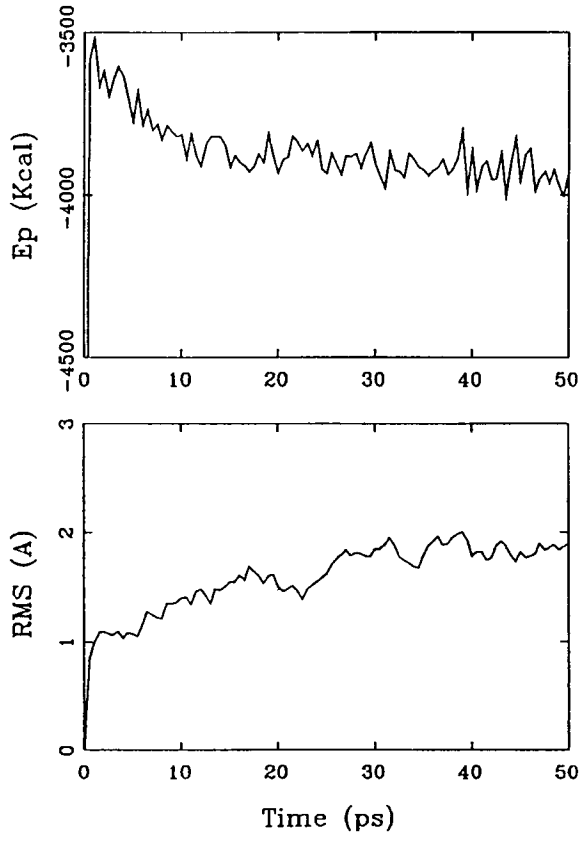

Fig. 2. Energy variations (upper) and rms deviations of all heavy atoms (lower) during the 50 ps molecular dynamics simulation.

Results and Discussion

MD behavior

The crystal structure of OMSVP3 was initially optimized for molecular dynamics simulation by energy minimization. Root mean square (rms) difference between the minimized structure and the crystal structure was 0.76 Å for the main chain atoms and 0.94 Å for the heavy atoms. Potential energy variation during the 50 ps molecular dynamics simulation of the minimized structure is shown in Fig. 2, along with the rms deviations of the heavy atoms in OMSVP3. Fig. 2 indicates that each

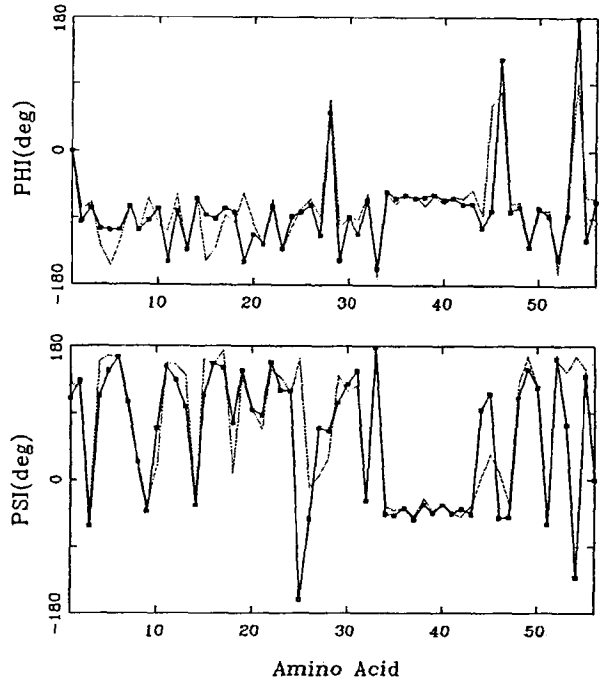

Fig. 3. Main chain dihedral angles (ϕ and ψ) of each amino acid residue in crystal structure (-----), and 25 to 50 ps averaged structure (■-■).

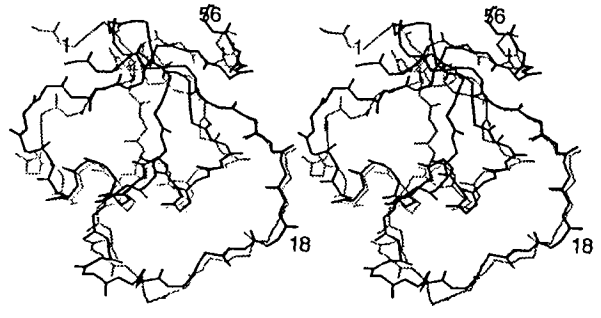

Fig. 4. Comparison of overall structure. The main chains atoms of the MD averaged structure (thick line) are shown along with the crystal structure (thin line). N-terminus, C-terminus and P1(Met 18) positions are labeled.

parameter has reached its plateau value after approximately 25 to 30 ps of MD steps. The system appears to have reached an equilibrium state after 25~30 ps, and the structures from 25 to 50 ps were averaged for further analysis.

Conformation of the MD averaged structure

The overall conformation of the polypeptide structure can be expressed by backbone dihedral angles, ϕ and ψ. Fig. 3 compares the backbone dihedral angles of the crystal structure and those of the MD averaged structure. The backbone dihedral angles in the MD averaged structure are reminiscent of those in the crystal structure, indicating that the overall structures during the simulation are similar to the crystal structure. N-terminal amino acid residues are highly flexible, because the site for limited proteolysis of ovomucoid is at the N-terminus of the third domain and the third domain proper starts at residue 6. Thus, the two structures were transformed so that the main chain atoms between residues 6 to 56 are superimposed. As shown in Fig. 4, the backbone structure of the MD averaged molecule is similar to the crystal structure.

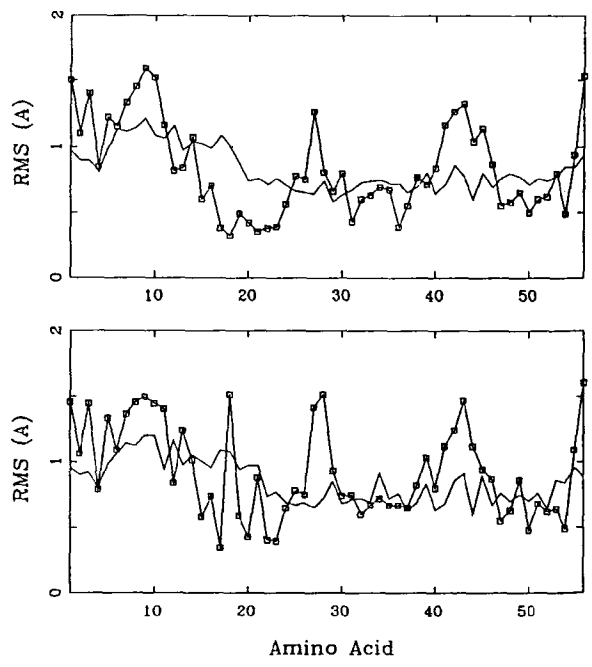

Fig. 5. Comparison of fluctuaions of main chain atoms (upper) and all heavy atoms (lower) in crystal structure (—) and during MD (□-□).

Positional fluctuations

The averaged rms deviations of main chain atoms and all heavy atoms of each amino acid residue during the last 5 ps simulation are compared with the positional fluctuations calculated from the crystallographic thermal factors (Fig. 5). During the simulation, there are three regions fluctuating the most; residues 7 to 11, residues 26 to 29, and residues 41 to 47. Crystal structure of the molecule indicates that these residues with high fluctuation form turns. On the other hand, movements of the backbone atoms around the scissile bond (Met18) are small (upper panel in Fig. 5), although the side chain atoms in the same residue are highly mobile (lower panel in Fig. 5). This may be related to the fact that the reactive site P1 amino acid appears to change during evolution so that P1 residue in OM3 consists of a hypervariable region. This extraordinary phenomenon of the OM3 evolution was exten-

Table 1. Intramolecular hydrogen bonds during MD simulation compared with those in crystal structure

Donor		Acceptor		Distance in crystal (Å)	MD frequency (%)
atom	residue	atom	residue		
N	Tyr 11	O	Cys 8	3.4	9
N	Arg 21	O	Gly 32	2.7	97
N	Leu 23	O	Tyr 31	2.9	92
N	Cys 24	O	His 52	2.8	78
N	Gly 25	O	Lys 29	2.8	74
N	Ser 26	O	Thr 49	2.8	83
N	Asn 28	O	Gly 25	2.9	19
N	Tyr 31	O	Leu 23	2.8	96
N	Phe 37	O	Asn 33	3.0	96
N	Cys 38	O	Lys 34	2.8	98
N	Asn 39	O	Cys 35	3.0	86
N	Ala 40	O	Asn 36	3.1	24
N	Val 41	O	Phe 37	2.9	91
N	Val 42	O	Cys 38	3.1	77
N	Glu 43	O	Asn 39	2.9	80
N	Ser 44	O	Ala 40	2.9	94
N	Asn 45	O	Val 42	2.9	11
N	Gly 46	O	Val 41	2.8	16
N	Thr 47	O	Ser 44	3.2	13
N	Ser 51	O	Cys 24	2.9	95
N	His 52	O	Cys 24	3.2	28
N	Cys 54	O	Pro 22	3.1	80
N	Lys 13	OD1	Asn 39	3.1	39
N	Asp 27	OD1	Asp 27	3.1	0
N	Lys 29	OD2	Asp 27	2.7	90
N	Asn 36	OD1	Asn 33	3.0	63
N	Leu 48	OG	Ser 44	3.0	1
N	Thr 49	OG	Ser 26	3.0	41
N	Cys 56	OG1	Thr 30	3.0	83
OG	Ser 26	O	Thr 49	3.2	3
ND2	Asn 33	O	Thr 17	2.8	73
ND2	Asn 33	O	Glu 19	3.0	64
NZ	Lys 34	O	Asp 7	3.1	7
ND2	Asn 39	O	Lys 13	3.1	51
OG	Ser 44	O	Ala 40	3.1	98
OG	Ser 44	O	Val 41	2.9	6
OG1	Thr 47	O	Ser 44	2.9	21
OG	Ser 9	OD1	Asp 7	2.5	5
OG1	Thr 17	OE1	Glu 19	3.0	97
OG	Ser 26	OG1	Thr 49	3.2	2
OH	Tyr 31	OD1	Asp 27	2.7	30
NZ	Lys 55	OH	Tyr 20	3.2	0

sively discussed by Laskowski *et al.* (1987).

Backbone to backbone interactions

For each coordinate set, every potential donor-acceptor pair was tested and a hydrogen bond was considered to form if the hydrogen to acceptor distance was within 2.5 Å and the donor-hydrogen-acceptor angle was between 120° and 180°. Table 1 lists the intramolecular hydrogen bonds in the crystal structure and the frequency of their occurrence in the MD simulation. Most of the hydrogen bonds initially found in the crystal structure are retained during the simulation. Hydrogen bonds between main chain atoms are important in maintaining the secondary structure and most of them are conseved in the MD simulation. A majority of the backbone hydrogen bonds were observed for more than 10% of the structures during the 50 ps simulation. Hydrogen bonds between main chain atoms of residues 33 to 44 which consist of α-helix are all conserved. Main chain hydrogen bond conservation is related to the observation that the MD structures have very similar backbone dihedral angles to those in the crystal structure (see Fig. 3). The interaction between residue Tyr11 and Cys8 is less conserved, which may be due to their location near the N-terminal end.

Backbone to side chain interactions

Several interactions between backbone atoms and side chain atoms appears to be weak compared with those within backbone atoms (Table 1). One of the important interactions is between the side chain of Asn 33 and the main chains of Thr17 and Glu19. The crystal structure shows that the side chain amide nitrogen of Asn33 donates hydrogen bonds to the carbonyl groups of Thr17 and Glu19 (Bode *et al.*, 1985). These interactions are known to be one of the factors that contribute to the stabilization of the intact form of the inhibitor (Ardelt and Laskowski, 1991; Musil *et al.*, 1991). As shown in Table 1, hydrogen bonds between the side chain of Asn33 and the main chains of Thr17 and Glu19 are maintained most of the time (73% and 64%, respectively), which helps to fix the conformation at P1.

Side chain to side chain interactions

The notable strong side chain to side chain interaction is between Thr17 and Glu19. Hydrogen bonds between OG1 of Thr17 and OE1 of Glu19 are retained through the simulation time (Table 1 and Fig. 6). In addition, a new hydrogen bond was formed between the side chains of Glu19 and Arg21 (OE2 of Glu19 and NH1 of Arg21), which was absent in the crystal structure. Fig. 6 shows that not only the interaction between Thr17 and Glu19 is conserved, also the hydrogen bond between the side chains of Glu19 and Arg21 is formed and remained through the 50 ps MD simulation. Therefore, it was found that there are hydrogen bonds form-

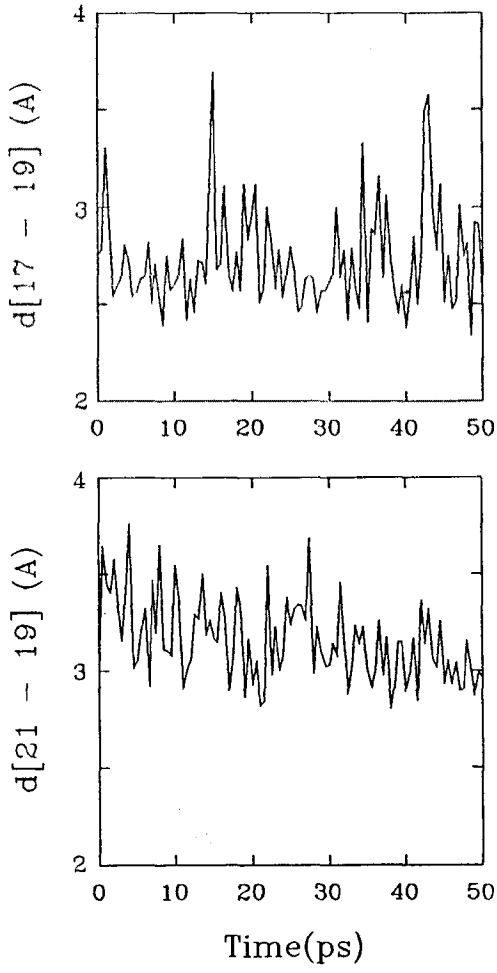

Fig. 6. Distance variations during MD simulation. Distances are between OG of Thr17 and OE1 of Glu19 (upper), and between OE2 of Glu19 and NH1 of Arg21 (lower).

ing Thr17 -- Glu19 -- Arg21 triad.

Comparison with structural and thermodynamic data

Ardelt and Laskowski (1991) investigated the effect of single amino acid replacements in ovomucoid third domains on the degree of reactive site hydrolysis. Their thermodynamics data indicate that replacing Thr for Arg at P2 position (residue 17) increases K_{hyd} by a factor of 2.9, while replacing Glu for Leu at P1′ (residue 19) leads to a 1.8-fold increase in K_{hyd}. They concluded that Thr17 and Glu19 are hydrogen bonded to each other most of the time. Musil *et al.* (1991) showed that this interaction present in virgin inhibitor is disrupted in the modified inhibitor (reactive site cleaved inhibitor). Moreover, crystal structures of the complex between *Streptomyces griseus* proteinase B (SGPB) and ovomucoid third domain from turkey (OMTKY3) show that the hydrogen bond between the side chains of Thr17 and Glu19 becomes even stronger (distance of 2.6 Å) compared with that in free inhibitor (distance

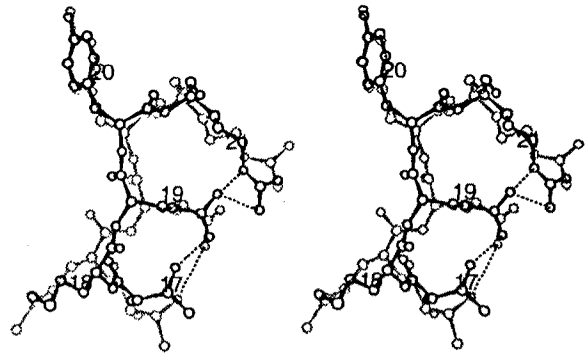

Fig. 7. Hydrogen bonds in the neighbor of P1(Met18). Both the MD averaged structure (thick line) and the crystal structure (thin line) are dipicted.

of 3.0 Å) (Read *et al.*, 1983; Huang *et al.*, 1995). Those structural data are consistent with the values of equilibrium constant for the association of inhibitor with most enzymes (Park, 1985). These observations imply that Thr17 and Glu19 interaction is important in inhibitor structure, becomes stronger on complex formation with enzyme, and contributes to the resistance to reactive site cleavage. MD results are consistent with the above observation showing that the side chain interaction between Thr17 and Glu19 is maintained. It also shows that not only the interaction between Thr17 and Glu19 is conserved, but also the hydrogen bond between the side chains of Glu19 and Arg21 is formed and remained throughout the simulation. Interaction between Glu19 and Arg21 is not seen in the OMSVP3 crystal structure; the distance between the OE2 of Glu 19 and NH2 of Arg21 is 5.1 Å (Bode *et al.*, 1985). In OMJPQ3, however, P2 position is Pro17 instead of Thr, and P1′ is Asp17 instead of Glu, and the crystal structure showed that OD1 of Asp19 is hydrogen-bonded to NH1 of Arg21 (Papamokos, 1982). In the enzyme-inhibitor complex, the side chain atoms of Arg21 interact with the backbone atoms of enzymes. In the OMTKY3-chymotrypsin complex, NH2 of Arg21 has a hydrogen bond with the carbonyl oxygen of residue 58 in chymotrypsin (Fuginaga *et al.*, 1987), while it is hydrogen-bonded with the carbonyl oxygen of residue 39 in the enzyme in the OMTKY3-SGPB complex (Huang *et al.*, 1995).

Implications for inhibitor mechanism

The importance of a hydrogen bond between Thr17 and Glu19 and interactions between the side chain of Asn33 and the main chains of Thr17 and Glu19 has already been discussed (Ardelt and Laskowski, 1991; Musil *et al.*, 1991). These interactions are considered to be the major force stabilizing the virgin form of inhibitor compared with modified inhibitor. On the

other hand, the interaction between the side chains of Asp19 and Arg21 was found in the OMJPQ3 crystal structure, where P2 is Pro so that the hydrogen bond between Thr17 and Glu19 cannot be made. The present simulation study showed that the following hydrogen bonds around P1 site are maintained; hydrogen bond between the main chain atoms of Asn33 and the side chain of Thr17, between the main chain atoms of Asn33 and the side chain of Glu19, between the side chains of Thr17 and Glu19, and between the side chains of Glu19 and Arg21. Hydrogen bonds between the side chains of amino acids near the reactive site (Thr17 -- Glu19 -- Arg21) appear to create a strong hydrogen bonding network without disturbing the backbone structure around P1. These interactions around P1 might be one of the factors contributing to the resistence to proteolytic cleavage. Further investigations would include extending the simulation time and the analysis of water molecules during the simulation. In the mean time, solution experimental work may find the existence of the hydrogen bonding network around the reactive site.

Acknowledgement

This work has been partly supported by the grant from Korea Research Institute of Bioscience and Biotechnology (95-G-08-01-A-45-B).

References

Ardelt, W. and Laskowksi, Jr., M. (1991) *J. Mol. Biol.* **220**, 1041.

Bode, W., Epp, O., Huber, R., Laskowksi, Jr., M. and Ardelt, W. (1985) *Eur. J. Biochem.* **147**, 387.

Empie, M. W. and Laskowski, Jr., M. (1987) *Biochemistry* **21**, 2274.

Fletcher, R. (1988) Practical Methods of Optimization, Vol. 1, *John Wiley & Sons, New York.*

Fujinaga, M., Sielecki, A. R., Read, R. J., Ardelt, W., Laskowski, Jr., M. and James, M. N. G. (1987) *J. Mol. Biol.* **195**, 397.

Hagler, A. T., Stern, P. S., Lifson, S. and Ariels, S. (1979) *J. Am. Chem. Soc.* **101**, 813.

Huang, K., Lu, W., Anderson, S., Laskowski, Jr., M. and James, M. N. G. (1995) *Prot. Sci.* **4**, 1985.

Kitson, D. H. and Hagler, A. T. (1988) *Biochemistry* **27**, 5246.

Krezel, A. M., Darba, P., Robertson, A. D., Fejzo, J., Macura, S. and Markley, J. (1994) *J. Mol. Biol.* **242**, 203.

Laskowski, Jr., M. and Kato, I. (1980) *Annu. Rev. Biochem.* **49**, 593.

Laskowski, Jr., M., Kato, I., Ardelt, W., Cook, J., Denton, A., Empie, M., Kohr, W., Park, S. J., Parks, K., Shatzley, B., Schoenberger, M., Tashiro, M., Vichot, G., Watley, H. E., Wieczorak, A., Wieczorak, M. (1987) *Biochemistry* **26**, 202.

Musil, D., Bode, W., Huber, R., Laskowski, Jr., M., Lin, T.-Y. and Ardelt, W. (1991) *J. Mol. Biol.* **220**, 739.

Papamokos, E., Weber, E., Bode, W., Huber, R., Empie, M. W., Kato, I. and Laskowski, Jr., M. (1982) *J. Mol. Biol.* **158**, 515.

Park, S. J. (1985) Ph.D. thesis, Purdue University.

Read, R. J., Fujinaga, M., Sielecki, A. R. and James, M. N. G. (1983) *Biochemistry* **22**, 4420.

Robertson, A., Westler, W. M. and Markley, J. L. (1988) *Biochemistry* **27**, 2519.

Schechter, I. and Berger, A. (1967) *Biochem. Biophys. Res. Comm.* **27**, 157.

Weber, E., Papamokos, E., Bode, W., Huber, R., Kato, I., Laskowski, Jr., M. (1981) *J. Mol. Biol.* **149**, 109.