Inhibitory Specificity Change of the Ovomucoid Third Domain of the Silver Pheasant upon Introduction of an Engineered Cys^{14}−Cys^{39} Bond‡

Hikaru Hemmi,§ Takashi Kumazaki,||,⊥ Toshimasa Yamazaki,# Shuichi Kojima,▽ Takuya Yoshida,○ Yoshimasa Kyogoku,△ Masataka Katsu,|| Fumikazu Shinohara,|| Hideyoshi Yokosawa,|| Kin-ichiro Miura,▽ and Yuji Kobayashi*,○

National Food Research Institute, 2-1-12 Kannondai, Tsukuba, Ibaraki 305-8642, Japan, Graduate School of Pharmaceutical Sciences, Hokkaido University, Kita-ku, Sapporo, Hokkaido 060-0812, Japan, National Institute of Agrobiological Sciences, 2-1-2 Kannondai, Tsukuba, Ibaraki 305-8602, Japan, Institute for Biomolecular Science, Gakushuin University, Meguro, Tokyo 171, Japan, Graduate School of Pharmaceutical Sciences, Osaka University, 1-6 Yamadaoka, Suita, Osaka 565-0871, Japan, and Institute for Protein Research, Osaka University, 3-2 Yamadaoka, Suita, Osaka 565-0871, Japan

Received August 26, 2002; Revised Manuscript Received January 14, 2003

ABSTRACT: The ovomucoid third domain from silver pheasant (OMSVP3), a typical Kazal-type inhibitor, strongly inhibits different serine proteases of various specificities, i.e., chymotrypsin, *Streptomyces griseus* protease, subtilisin, and elastase. Structural studies have suggested that conformational flexibility in the reactive site loop of the free inhibitor may be related to broad specificity of the ovomucoid. On the basis of the structural homology between OMSVP3 and ascidian trypsin inhibitor (ATI), which has a cystine-stabilized α-helical (CSH) motif in the sequence, we prepared the disulfide variant of OMSVP3, introducing an engineered disulfide bond between positions 14 and 39 near the reactive site (Met^{18}-Glu^{19}) by site-directed mutagenesis. The disulfide variant P14C/N39C retained potent inhibitory activities toward α-chymotrypsin (CHT) and *S. griseus* proteases A and B (SGPA and SGPB), while this variant lost most of its inhibitory activity toward porcine pancreatic elastase (PPE). We determined the solution structure of P14C/N39C, as well as that of wild-type OMSVP3, by two-dimensional nuclear magnetic resonance (2D NMR) methods and compared their structures to elucidate the structural basis of the inhibitory specificity change. For the molecular core consisting of a central α-helix and a three-stranded antiparallel *β*-sheet, essentially no structural difference was detected between the two (pairwise rmsd value = 0.41 Å). In contrast to this, a significant difference was detected in the loop from Cys^{8} to Thr^{17}, where in P14C/N39C it has drawn approximately 4 Å nearer the central helix to form the engineered Cys^{14}−Cys^{39} bond. Concomitantly, the Tyr^{11}-Pro^{12} *cis*-peptide linkage, which is highly conserved in ovomucoid third domains, was isomerized to the trans configuration. Such structural change in the loop near the reactive site may possibly affect the inhibitory specificity of P14C/N39C for the corresponding proteases.

Ovomucoid in egg white consists of three tandem, homologous, Kazal-type domains (*1*). Comparative studies using a large number of ovomucoid third domains have made contributions to elucidation of the structure−function relationship of the Kazal-type inhibitor (*2*). Some ovomucoids have been analyzed by X-ray crystallography and by two-dimensional nuclear magnetic resonance (2D NMR) methods both free and in complexes with their target proteases (*3*−*8*). The ovomucoid third domains from silver pheasant (OMSVP3) and from Japanese quail (OMJPQ3), as well as from turkey (OMTKY3), are extensively studied third domains. The tertiary structure of OMSVP3 (*7*) is presented in Figure 1. The reactive site loop (P_6 to P'_3) [notation according to Schechter and Berger (*9*)] is linked to the central helix by the Cys^{16}−Cys^{35} bond, together with two hydrogen bonds provided by the side chains of Asn^{33} and Asn^{39}. OMSVP3 with P_1 Met inhibits chymotrypsin (CHT),[1] subtilisin, and elastase (*10*), as is the case for OMTKY3 with P_1 Leu (*11*). Crystallographic analyses of OMJPQ3 and

‡ The atomic coordinates for the 15 best conformers of wild-type OMSVP3 and P14C/N39C, respectively, described in this paper have been deposited with the Protein Data Bank (accession numbers 1IY5 and 1IY6).

* To whom correspondence should be addressed. Telephone: +81-(6)68798220. Fax: +81(6)68798221. E-mail: yujik@protein.osaka-u.ac.jp.

§ National Food Research Institute.

|| Graduate School of Pharmaceutical Sciences, Hokkaido University.

⊥ Present address: Faculty of Engineering, Aomori University, 2-3-1 Kobata, Aomori 030-0943, Japan.

National Institute of Agrobiological Sciences.

▽ Institute for Biomolecular Science, Gakushuin University.

○ Graduate School of Pharmaceutical Sciences, Osaka University.

△ Institute for Protein Research, Osaka University.

[1] Abbreviations: CSH motif, cystine-stabilized α-helical motif; DQF-COSY, double-quantum-filtered correlation spectroscopy; HLE, human leukocyte elastase; ^{1}H NMR, proton nuclear magnetic resonance; CHT, α-chymotrypsin; PPE, porcine pancreatic elastase; NOE, nuclear Overhauser effect; NOESY, nuclear Overhauser effect spectroscopy; OMXXX3, ovomucoid third domain from species XXX (code of bird species); CHI, chicken; JPQ, Japanese quail; TKY, turkey; SVP, silver pheasant; PAGE, polyacrylamide gel electrophoresis; rmsd, root-mean-square deviation; SGPA and SGPB, *Streptomyces griseus* proteases A and B; TOCSY, total correlation spectroscopy; TPPI, time-proportional phase increment.

10.1021/bi026727c CCC: $25.00
Published on Web 02/14/2003

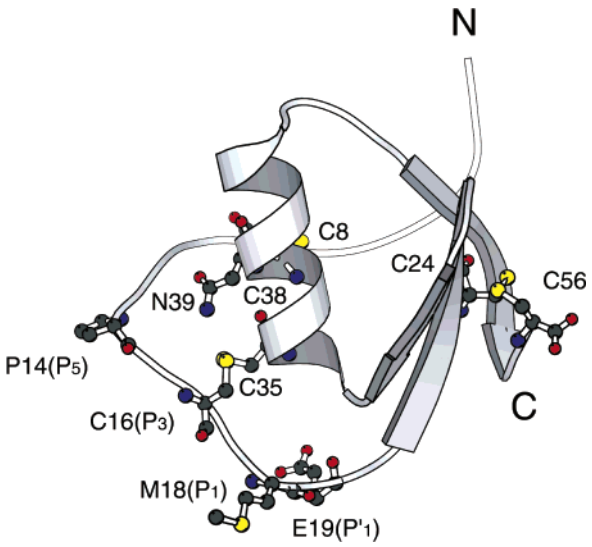

FIGURE 1: Molecular structure of natural OMSVP3 in a crystal state (*7*). The structure is drawn with MOLSCRIPT (*44*). Three disulfide bridges (Cys^{8}−Cys^{38}, Cys^{16}−Cys^{35}, and Cys^{24}−Cys^{56}), two mutation site residues (Pro^{14} and Asn^{39}), and the reactive site residue (Met^{18}) are also shown. N and C indicate amino and carboxyl termini, respectively. The reactive site loop (P6 to P′3) [notation according to Schechter and Berger (*9*)] is fixed to the central helix by the Cys^{16}−Cys^{35} bond, together with two hydrogen bonds provided by the side chains of Asn^{33} and Asn^{39}.

OMSVP3 in the free state indicate that the reactive site loop has relatively large *B* values (*6*, *7*). Furthermore, a conformational change occurs in the reactive site loop between OMTKY3 structures bound to CHT and bound to *Streptomyces griseus* protease B (SGPB). Thus, it has been suggested that some conformational flexibility in the reactive site loop of the free inhibitor might be related to the broad specificity of ovomucoid (*12*). This is in marked contrast to a Kunitz-type inhibitor, such as bovine pancreatic trypsin inhibitor, which has a rigid reactive site and shows narrower specificity. So far, inhibitory specificity has been studied by using naturally occurring variants of ovomucoid (*2*). However, the structural basis of the broad specificity of ovomucoid has not been satisfactorily elucidated. Natural variants do not necessarily possess suitable amino acid substitutions to relate structure and specificity. Site-directed mutagenesis to create mutated proteins is a powerful technique for obtaining more information on the structure−function relationship of ovomucoids. In particular, introduction of the nonnatural disulfide bridge into the reactive site loop and removal of the highly conserved disulfide bridge near the reactive site should provide valuable information on the relationship between conformational flexibility and specificity of protease inhibitors. This line of approach has been applied to serpins (*13*) or *Streptomyces* subtilisin inhibitor (*14*) but not to a small-sized and compact inhibitor such as the ovomucoid third domain. In this study we introduced a nonnatural disulfide bridge on the reactive loop of OMSVP3 and examined the effect of the mutation on its inhibitory specificity. OMSVP3 was used because this inhibitor, as a parent inhibitor molecule, has the broadest inhibitory specificity, and thus the effect of the mutation on the inhibitory specificity should be easily detected.

When a disulfide bridge is introduced, it is important that the engineered bridge does not destroy the core structure of the inhibitor. A cystine-stabilized α-helical (CSH) motif (*15*) would be an ideal scaffold for the purpose. The CSH motif has been found in the sequences of various biologically active peptides. It is well characterized by an α-helical segment spanning the Cys-X-X-X-Cys sequence portion that is cross-linked by two disulfide bridges to the sequence portion Cys-X-Cys that is folded in an extended *β*-strand-type structure (*17*). Thus, it is expected that an appropriate introduction of disulfide bonds in the ovomucoid third domain to form a CSH motif consisting of its central helix and its reactive loop forces the reactive loop to adapt to a rigid extended structure without any effect to the central helix. We prepared an OMSVP3 variant, P14C/N39C, by site-directed mutagenesis that was designed to form a novel disulfide bond between positions 14 and 39 (Figure 1). The mutation sites for creation of a disulfide bridge in OMSVP3 were chosen on the basis of structural homology with ascidian trypsin inhibitor (ATI) that has a CSH motif in the sequence (*15*). ATI with P1 Lys strongly inhibits mammalian trypsin and plasmin (*16*). In ATI, the two segments 37−41 and 12−14 correspond to the respective ones of the CSH motif, where Cys^{12}−Cys^{41} and Cys^{14}−Cys^{37} are formed. The solution structure of ATI determined by 2D NMR methods is very similar to that of OMSVP3, and the Cys^{14}−Cys^{37} bond in ATI, corresponding to the Cys^{16}−Cys^{35} bond in OMSVP3, fixes the reactive site loop of ATI (*15*). Residues Cys^{12} and Cys^{41}, forming an additional disulfide bond in ATI, then correspond to Pro^{14} and Asn^{39} in OMSVP3, respectively (Figure 1).

First, in the present study we describe the preparation and characterization of P14C/N39C, including the disulfide locations. Second, the effect of introducing a non-native disulfide bridge on the inhibitory activity for different proteases is described. Third, the determination of the solution structure of P14C/N39C, as well as that of wild-type OMSVP3, by 2D ^{1}H NMR spectroscopy is described. Finally, their structures are compared, and the structural basis of the inhibitory specificity change of OMSVP3, caused by introducing the engineered Cys^{14}−Cys^{39} bond, is discussed.

EXPERIMENTAL PROCEDURES

Design of OMSVP3 Variants. In this study, the wild-type OMSVP3 gene is constructed first, and then the genes for three variants such as P14C/N39C, P12A, and P14A were prepared. The mutation sites are shown on the tertiary structure of native OMSVP3 (see also Figure 1). P14C/N39C was designed on the basis of the structural homology between OMSVP3 and ATI (*15*) to examine the relationship between the inhibitory specificity and the conformational flexibility of the reactive site region. P14A was prepared to check whether the mutation changes the inhibitory specificity due to the mere amino acid substitution at the P_5 site. As will be described later, the Tyr^{11}−Pro^{12} peptide bond in P14C/N39C is found to convert to the trans configuration. Thus, P12A was designed to examine the effect of configuration change in the highly conserved Tyr^{11}−Pro^{12} *cis*-peptide linkage on inhibitory activity. The configuration in the respective peptide linkage of this variant is expected to adopt the trans form by the replacement of Pro to Ala at the P_7 site. The numbering system used in this study is according to Laskowski et al. (*2*).

Site-Directed Mutagenesis. PCR kits were purchased from Perkin-Elmer, Chiba. The restriction enzymes and ligation

kits were purchased from Boehringer Mannheim, Takara Shuzo, and Toyobo. Kojima et al. (*18*) had already constructed the expression plasmid pMKAPOV for the OMCHI3 gene (OV3). The amino acid sequence of OMSVP3 is almost the same as that of OMCHI3, except for three residues. Thus Asp^{15}, Ala^{18}, and Asp^{20} at the P_4, P_1, and P'_2 sites in OMCHI3 were genetically substituted for Ala, Met, and Tyr, respectively, to construct the wild-type OMSVP3 gene. The *Pst*I−*Eco*RI fragment of an OV3 gene having Met at the P_1 site [OV3 (P_1 Met)] was inserted into the respective site of pTZ18U, and *Escherichia coli* CJ236 was transformed by the constructed plasmid. The single-stranded DNA containing uracil bases was prepared by infection of helper phage M13KO7. Codon substitution of Asp → Tyr at the P'_2 site was carried out using the single-stranded DNA and a mutagenic primer of 5′ ACG-ATG-GAA-T*AC-AGA-CCT-C 3′ (where the asterisk indicates a mismatched base), according to the method of Kunkel (*19*). Subsequent mutation of Asp → Ala at the P_4 site was carried out using a mutagenic primer of 5′ CT-AAG-CCT-GC*C-TGC-ACG-A 3′ and a single-stranded DNA derived from the mutated plasmid encoding OV3 (P_1Met/P'_2Tyr). The expression plasmid for the wild-type OMSVP3 gene thus obtained was named pOVWT. For construction of the P14C/N39C gene, substitution of Pro → Cys and Asn → Cys at the P_5 and P'_{21} sites, respectively, was carried out by successive site-directed mutagenesis of the gene of OV3 (P_1Met/P'_2Tyr). The mutagenic primers used were 5′ TAC-CCT-AAG-T*G*T-GC*C-TGC-ACG-AT 3′ for Asp → Ala and Pro → Cys at the P_4 and P_5 sites, respectively, and 5′ AAC-TTC-TGC-T*G*T-GCA-GTC-GT 3′ for Asn → Cys at the P'_{21} site. The expression plasmid for the P14C/N39C gene was constructed by ligation of the *Pst*I−*EcoRI* fragment of the mutated OV3 gene, the *Eco*RI−*Sph*I fragment (2.8 kb) of plasmid pMK2, and the *Sph*I−*Pst*I fragment (550 bp) of plasmid pMKAPOV (*18*) and named pOVC14/C39. *E. coli* JM105 was separately transformed with pOVWT and pOVC14/C39.

For expressions of the P12A and the P14A genes, the pET system was applied. The *Pst*I−*Eco*RI fragment (170 bp) was amplified by PCR with pOVWT as a template. The mutagenic forward primers used were 5′ C-TGC-AGT-GAG-TAC-G*CT-AAG-CCT 3′ for Pro → Ala, the P_7 site in the P12A gene, or 5′ C-TGC-AGT-GAG-TAC-CCT-AAG-G*CT-GCC-TG 3′ for Pro → Ala, the P_5 site in the P14A gene. A reverse primer was 5′ GAATTCTCTCAGCTCT-GATA 3′ containing an *Eco*RI site. On the other hand, double-stranded DNA fragments encoding the N-terminal region of OMSVP3 were chemically synthesized. A plus strand was 5′ CC-GCT-GTG-AGT-GTT-GAC-TGC-A 3′ and a minus strand was 5′ GTC-AAC-ACT-CAC-AGC-GG 3′. The annealed fragments were ligated with the mutated *Pst*I−*Eco*RI fragment (170 bp) and then inserted into the *Bal*I−*Eco*RI backbone fragment of pET22b vector (Novagen, Madison, WI) to obtain the expression vectors pETOV12A and pETOV14A. *E. coli* BL21(DE3)*p*LysS was then transformed with pETOV12A or pETOV14A, respectively. The base substitutions were confirmed by dideoxy sequencing of the plasmids. The recombinant inhibitors were designed to start with Ala^3 of the natural ovomucoid third domain after cleavage with signal peptidase, as previously described (*18*).

Expression and Purification of Recombinant Inhibitors. Cultivation and induction with IPTG were performed as previously described (*18*). Each secreted inhibitor in the supernatant of *E. coli* cells was precipitated with ammonium sulfate. The precipitate was dialyzed against 25 mM sodium acetate, pH 5.3. The dialyzate was subjected to SP-Sephadex C-50 chromatography using the acetate buffer and a linear gradient of NaCl. Inhibitor fractions then were subjected to anhydrochymotypsin−Sepharose 4B chromatography using 50 mM Tris-HCl, pH 8.0, containing 20 mM $CaCl_2$ (*20*). Elution was carried out by addition of 5 mM HCl. Inhibitor fractions were finally subjected to SP-5PW (Tosoh, Tokyo) chromatography using the acetate buffer and a linear gradient of NaCl. The purification procedure was monitored by measuring the inhibitory activity toward chymotrypsin (*18*).

Protein Characterization. Performic acid oxidation and reduction and *S*-pyridyl ethylation were carried out according to the published methods (*21*, *22*). Amino acid composition analysis was performed as previously described (*23*). The N-terminal sequence of peptides was determined on an Applied Biosystems Model 492 protein sequencer. Polyacrylamide gel electrophoresis under nondenaturing conditions (native PAGE) with 15% gel was carried out at 4 °C using Laemmli's buffer system (*24*) in the absence of SDS. The proteins were stained with Coomassie Brilliant Blue R-250.

Location of Disulfide Bridges. Native P14C/N39C protein (50 nmol) was digested at 37 °C with thermolysin (Seikagaku Kogyo, Tokyo) (E/S = 1/70, mol/mol) in 150 mM ammonium formate, pH 6.5, containing 2 mM $CaCl_2$ and 0.02% sodium azide for 3 days. The digest was subjected to reversed-phase HPLC as previously described (*20*) to separate cystine-containing peptides, i.e., T1−T3. T3 was found to be still heterogeneous by peptide analysis. Thus T3 (30 nmol) was further digested at 37 °C with *Staphylococcus aureus* V8 protease (Sigma Chemical Co.) (E/S = 1/10, mol/mol) in 50 mM sodium phosphate, pH 7.8, for 36 h. T3V2 containing three chains was applied to the protein sequencer to determine the disulfide partners for each of the directly linked two half-cystine residues (Cys^{38}−Cys^{39}).

Enzyme Kinetics. S. griseus proteases A and B (SGPA and SGPB) were prepared as described by Narahashi (*25*). Porcine pancreatic elastase (PPE) was kindly donated by Eisai Co., Tokyo. The following materials were purchased from the indicated commercial sources: bovine α-chymotrypsin (CHT), human leukocyte elastase (HLE), Suc-Ala-Ala-Ala-*p*NA, and Suc-Ala-Ala-Pro-Phe-*p*NA (Sigma Chemical Co.); Suc-Ala-Ala-Pro-Phe-MCA and Suc-Ala-Ala-Ala-MCA (Peptide Institute, Osaka). All assays were carried out at 25 °C in 50 mM Tris-HCl, pH 7.8, containing 20 mM $CaCl_2$ and 0.005% Triton X-100. Molar concentrations of inhibitors were determined by amino acid composition analysis. The active site concentration of CHT was first determined, as described by Kézdy and Kaiser (*26*). The wild-type OMSVP3 solution was then titrated with CHT of known concentration. Concentrations of SGPA, SGPB, and PPE were then determined by titrating with wild type of known concentration. Suc-Ala-Ala-Pro-Phe-MCA was used as a substrate for assays of CHT, SGPA, and SGPB. For elastase assay Suc-Ala-Ala-Ala-MCA and Suc-Ala-Ala-Ala-*p*NA were used. Substrate concentration at least three times lower than its K_m was used for the study of inhibitor−enzyme

reactions. All of the K_m values were determined under assay conditions. The equilibrium dissociation constants (K_is) were determined essentially as previously described (*27*). For the measurement of association rate constants (k_{on}s), the enzyme solution was added to a mixture of inhibitor and substrate, and the progress curve corresponding to the pre steady state was obtained according to Bieth (*28*). The inhibitor concentration was chosen so that $[I]_0 \geq 5[E]_0$. A substrate concentration less than $0.3K_m$ was used. Data fitting were carried out using nonlinear regression analysis (KaleidaGraph for Windows, Synergy Software).

NMR Experiment. The OMSVP3 sample was dissolved in 250 μL of 100% D_2O or a 90% H_2O/10% D_2O solution to give a final concentration of approximately 4 mM. The pH of the solution was adjusted to 4.0 with 1 M HCl. All NMR spectra were obtained on Bruker AMX500 and DRX600 spectrometers with quadrature detection in the phase-sensitive mode by TPPI (*29*) and States−TPPI (*30*). The following spectra were recorded at 25, 30, and 35 °C with 12 ppm spectral widths in the t_1 and t_2 dimensions: two-dimensional (2D) double-quantum-filtered correlated spectroscopy (DQF-COSY) (*31*), recorded with 512 and 2048 complex points in the t_1 and t_2 dimensions; 2D homonuclear total correlated spectroscopy (TOCSY) (*32*) with a MLEV-17 mixing sequence, recorded with mixing times of 40 and 80 ms and 512 and 2048 complex points in the t_1 and t_2 dimensions; 2D nuclear Overhauser effect spectroscopy (NOESY) (*33*), recorded with mixing times of 60, 80, 100, and 200 ms and 512 and 2048 complex points in the t_1 and t_2 dimensions. A high digital resolution DQF-COSY was recorded using 512 and 4096 complex points in the t_1 and t_2 dimensions. Water suppression was performed using the WATERGATE sequence (*34*, *35*). Slowly exchanging amide protons were determined by lyophilizing the protein from H_2O solution, dissolving the protein in D_2O, and collecting sequential 2 h 2D TOCSY spectra. All NMR spectra were processed using XWINNMR (Bruker Instruments) and the NMR Pipe program (*36*). Before Fourier transformation the shifted sine-bell window function was applied to the t_1 and t_2 dimensions. Chemical shifts were referenced to internal 2,2-dimethyl-2-silapentane-5-sulfonate (DSS) at 25 °C.

Structure Calculations. NOE-derived distance restraints were classified into three ranges, 1.8−2.5 Å, 1.8−3.5 Å, and 1.8−5.0 Å, according to the relative NOE intensities. Upper distance limits for NOEs involving methyl protons and nonstereospecifically assigned methylene protons were corrected appropriately for center averaging (*37*). In addition, a distance of 0.5 Å was added to the upper distance limits for only NOEs involving methyl protons (*38*) after correction for center averaging. Torsion angle constraints on backbone ϕ and ψ angles were derived from $^3J_{HNH\alpha}$ coupling constants estimated from the high digital resolution 2D DQF-COSY spectra and sequential and short-range NOEs. The 29 and 36 ϕ angle restraints were obtained for wild-type OMSVP3 and P14C/N39C, respectively. Backbone ϕ angles were restrained to $-60° \pm 30°$ for $^3J_{NH\alpha} < 6$ Hz, $-120° \pm 50°$ for $^3J_{NH\alpha} = 8-8.5$ Hz, $-120° \pm 40°$ for $^3J_{NH\alpha} = 8.5-9$ Hz, and $-120° \pm 30°$ for $^3J_{HN\alpha} > 9$ Hz. The 11 and 13 ψ angles in α-helical regions of wild type and P14C/N39C, respectively, were restrained to $-40° \pm 30°$. Side-chain χ_1 angles were determined by $^3J_{H\alpha H\beta}$ coupling constants qualitatively estimated from short-mixing TOCSY connectivities in combination with NH−Hβ and Hα−Hβ NOEs (*39*). The 20 and 19 χ_1 angle restraints, respectively, for wild type and P14C/N39C were obtained. The χ_1 angle restraints were normally restricted to a ±60° range from staggered conformations, g^+ (+60°), t (180°) or g^- (−60°). Hydrogen−deuterium exchange experiments identified 25 and 26 hydrogen bond donors for wild type and P14C/N39C, respectively. Corresponding hydrogen bond acceptors were determined on the basis of NOE patterns observed for regular secondary structural regions and preliminary calculated structures without restraints regarding hydrogen bonds. Hydrogen bond constraints were applied to N−H and C═O groups: 1.7−2.3 Å for the H−O distance and 2.7−3.3 Å for the N−O distance.

Structure calculations were performed using the hybrid distance geometry-simulated annealing method (*40*) using X-PLOR 3.1 (*41*). A set of 581 NOE-derived distance restraints (including 111 intraresidue, 174 sequential, 117 medium range, and 179 long range), 50 hydrogen bond restraints, and 60 dihedral angle restraints were obtained for wild-type OMSVP3. In the case of P14C/N39C, a set of 562 NOE-derived distance restraints (including 114 intraresidue, 180 sequential, 101 medium range, and 167 long range), 52 hydrogen bond restraints, and 68 dihedral angle restraints were obtained. The structure calculation proceeded in two stages by using the standard X-PLOR protocol. In the first stage, a low-resolution structure was preliminarily determined using NOE-derived distance restraints and dihedral angle restraints except for ψ angle restraints. In the second stage of the calculation, the same protocol was applied by adding hydrogen bond restraints and ψ angle restraints. The force constants for the distance restraints were set to 50 kcal mol^{-1} Å^{-2} throughout all the calculations, and dihedral angle restraints were initially set to 5 kcal mol^{-1} rad^{-2} during the high-temperature dynamics and gradually increased to 200 kcal mol^{-1} rad^{-2} during the annealing stage. The final round of calculations began with 100 initial structures, 38 for wild type and 23 for P14C/N39C of which, after refinement, had no distance and dihedral angle violations greater than 0.5 Å and 5°, respectively. Among the 38 and 23 final structures, the 15 structures of wild type and P14C/N39C with low total energy and low deviation from mean structure were used for further analyses. The structures have good covalent geometry and stereochemistry, as evidenced by the low rmsd values for bonds, angles, and improper torsion angles from idealized geometry. The Ramachandran plot confirmed the high quality of the 15 structures of wild type or P14C/N39C, respectively, which showed that over 90% of ϕ and ψ angles in both proteins are found in the most favored and additionally allowed regions and less than 1% of ϕ and ψ angles are found within the disallowed region. The average coordinates of the ensembles of the 15 low-energy structures were subjected to 500 cycles of Powell restrained minimization to improve stereochemistry and nonbonded contacts. All subsequent numerical analyses were performed using X-PLOR, Insight II (MSI), PROCHECK-NMR (*42*), and MOLMOL (*43*). Structure figures were generated using Molscript (*44*), Insight II, and MOLMOL.

RESULTS AND DISCUSSION

Preparation of OMSVP3 Variants. P14C/N39C, P12A, P14A, and wild-type OMSVP3 were efficiently expressed

Table 1: Amino Acid Compositions and Disulfide Positions of the Cystine-Containing Peptides Derived from Native P14C/N39C

	T1	T2	T3V2
	Amino Acid Compositions (Residues per Molecule)[a]		
Cya[b]	1.81 (2)	2.16 (2)	3.73 (4)
Asp	1.29 (1)	2.39 (2)	2.15 (2)
Glu			1.09 (1)
Ser		1.20 (1)	0.67 (1)
Gly	1.12 (1)	2.60 (2)	
Thr	0.96 (1)	1.10 (1)	
Ala	*1* (1)		0.80 (1)
Pro			1.13 (1)
Tyr	0.92 (1)		1.05 (1)
Val			1.20 (1)
Leu		1.07 (1)	
Phe		*1* (1)	*1* (1)
Lys	0.96 (1)	1.89 (2)	1.12 (1)
total	(8)	(12)	(14)
	Disulfide Positions		
fragments	Ala^{15}-*Cys*^{16}-Thr^{17} plus Tyr^{31}-Gly^{32}-Asn^{33}-Lys^{34}-*Cys*^{35}	Leu^{23}-*Cys*^{24}-Gly^{25}-Ser^{26}-Asp^{27}-Asn^{28}-Lys^{29}-Thr^{30} plus Phe^{53}-Gly^{54}-Lys^{55}-*Cys*^{56}	Val^{6}-Asp^{7}-*Cys*^{8}-Ser^{9}-Glu^{10} plus Tyr^{11}-Pro^{12}-Lys^{13}-*Cys*^{14} plus Asn^{36}-Phe^{37}-*Cys*^{38}-*Cys*^{39}-Ala^{40}
S−S bridge	Cys^{16}−Cys^{35}	Cys^{24}−Cys^{56}	Cys^{8}−Cys^{38}, Cys^{14}−Cys^{39} or Cys^{8}−Cys^{39}, Cys^{14}−Cys^{38}

[a] Analytical values are indicated as molar ratios with respect to the amino acid with an italicized value in each column. Values in parentheses are those deduced from the sequence data. [b] Cysteic acid.

Table 2: Kinetic Parameters for the OMSVP3−Enzyme Interaction

	CHT		PPE		SGPA		SGPB	
	K_i (M)	k_{on} (M^{-1} s^{-1})	K_i (M)	k_{on} (M^{-1} s^{-1})	K_i (M)	k_{on} (M^{-1} s^{-1})	K_i (M)	k_{on} (M^{-1} s^{-1})
natural[a]	5.6×10^{-12}	1.3×10^{7}	8.3×10^{-11}	1.7×10^{6}				
wild type	6.5×10^{-12}	1.7×10^{7}	4.3×10^{-11}	3.0×10^{6}	1.3×10^{-11}	2.1×10^{6}	2.9×10^{-11}	2.5×10^{6}
P14C/N39C	6.6×10^{-11}	1.5×10^{7}	4.8×10^{-6}	$<5.0 \times 10^{2}$	5.4×10^{-11}	1.8×10^{6}	5.9×10^{-10}	1.6×10^{6}
P14A	1.4×10^{-12}	1.1×10^{7}	5.1×10^{-11}	2.9×10^{6}	2.1×10^{-11}	1.7×10^{6}	2.9×10^{-11}	2.1×10^{6}
P12A	2.0×10^{-11}	1.2×10^{7}	1.6×10^{-10}	1.3×10^{6}	3.8×10^{-11}	1.8×10^{6}	4.9×10^{-11}	1.3×10^{6}

[a] Data from ref *10*.

by *E. coli* and purified by three steps of chromatography, including affinity chromatography on anhydrochymotrypsin−Sepharose. The yields from 1 L of the culture are 3 mg for wild type and the two Ala variants and 0.6 mg for P14C/N39C. The overall yield in the purification step was about 60%. The purity was checked on native PAGE analysis. Amino acid compositions of these proteins were all in excellent agreement with expected ones. They had the same N-terminal sequence starting with Ala^{3}-Val^{4}-Ser^{5}-, as designed.

Locations of the Disulfide Bridges of P14C/N39C. Prior to this experiment, native P14C/N39C protein was treated with 4-vinylpyridine in the absence and presence of dithioerythritol. The amino acid analysis showed that no *S*-(pyridylethyl)cysteine was detectable unless the protein had been pyridylethylated in the presence of dithioerythritol. This indicates that the eight cysteines are all involved in the formation of disulfide bonds. From the thermolysin digest of native P14C/N39C protein three cystine-containing peptides, i.e., T1, T2, and T3, were obtained on reversed-phase HPLC (data not shown). The amino acid compositions of T1 and T2 correspond to residues 15−17 plus 31−35 and residues 23−30 plus 53−56, respectively (Table 1). The presence of Cys^{16}−Cys^{35} and Cys^{24}−Cys^{56} was thus established. Peptide analysis indicated that T3 was a 1:3 mixture of two peptides consisting of residues 6−14 plus 37−40 and residues 6−14 plus 36−40, respectively. Thus this fraction was further treated with V8 protease to cleave the Glu^{10}−Tyr^{11} bond, and the digest was again subjected to reversed-phase HPLC (data not shown). T3V2 thus obtained as a major component is a three-chain peptide, corresponding to residues 6−10 (Val^{6}-Asp^{7}-Cys^{8}-Ser^{9}-Glu^{10}) plus 11−14 (Tyr^{11}-Pro^{12}-Lys^{13}-Cys^{14}) plus 36−40 (Asn^{36}-Phe^{37}-Cys^{38}-Cys^{39}-Ala^{40}) (see also Table 1). To determine the disulfide partners for each of the directly linked two half-cystine residues in segment 36−40, this fraction was applied to protein sequencer. PTH-cystine was detected after both the third and fourth cycles of the Edman degradation (data not shown), supporting the combinations of Cys^{8}−Cys^{38} and Cys^{14}−Cys^{39}. Thus, the recombinant P14C/N39C protein actually has a nonnatural Cys^{14}−Cys^{39} bond, together with three natural disulfide bonds, as expected.

Inhibitory Specificity Change due to Introduction of a Nonnatural Cys^{14}−Cys^{39} Bond into OMSVP3. The K_i values of P14C/N39C, P12A, P14A, and wild-type OMSVP3 toward four different serine proteases were determined (Table 2). Recombinant wild type, despite the deletion of the Leu^{1}-Ala^{2} sequence, has essentially the same binding affinity as natural OMSVP3 (*10*). None of the variants has much more potent inhibitory activity than wild type. The most striking features of the data are that P14C/N39C shows a dramatic decrease in inhibition toward PPE by a factor of 10^{5} in comparison with wild type, whereas the variant retains potent inhibitory activities toward CHT, SGPA, and SGPB. A preliminarily experiment also shows that P14C/N39C is a 100 times weaker inhibitor for HLE. These findings indicate

that P14C/N39C loses most the inhibitory activity only for elastase. Next, the association rate constants (k_{on}s) were determined in order to investigate the alteration of the inhibitory activity for proteases in more detail. Wild-type OMSVP3 gives high k_{on} values ($>3 \times 10^5$ M^{-1} s^{-1}) for CHT, SGPA, SGPB, and PPE. P14C/N39C also gives high values for CHT, SGPA, and SGPB but a very low value ($<5 \times 10^2$ M^{-1} s^{-1}) for PPE. The dramatic decrease of inhibition toward PPE is mainly caused by the decrease of the k_{on} value, suggesting that P14C/N39C cannot practically associate with PPE. Then, we examined the susceptibility of P14C/N39C to elastase digestion by native PAGE analysis. The results indicated that P14C/N39C was hardly degraded even after 3 day incubation with PPE (E/S = 1/10, mol/mol) at pH 7.8 and 25 °C. Thus, the dramatic decrease in inhibitory activity toward PPE was due to the lack of tight binding to the enzyme.

It is generally accepted that the P_1 site residue of serine protease inhibitors is a predominant determinant for the inhibitory specificity (*45*−*50*). Comparative studies using more than 100 species of the third domains demonstrated that changes in residues other than the P_1 site residue can often exert large differential effects toward the different enzymes (*2*). Among them, the two ovomucoid third domains from chestnut-belied scaled quail (OMSCQ3N) and blue scaled quail (OMSCQ3S) are notable from a viewpoint of their inhibitory specificity. In both cases, substitution of Ala by Asp in the P_4 position of OMTKY3 results in a dramatic decrease in affinity for elastase and subtilisin, but it has no effect for CHT. Unfortunately, the tertiary structures of these domains have not yet been analyzed. In the case of P14C/N39C, it is the highly conserved Pro residue in the P_5 position that is replaced by Cys to form the non-native Cys^{14}−Cys^{39} bond. Our preliminary experiment shows that introduction of the Cys^{14}−Cys^{39} bond results in some unfavorable effect for subtilisin BPN′, as are the cases for OMSCQ3N and OMSCQ3S. Furthermore, P12A and P14A, in contrast with P14C/N39C, exhibit essentially the same inhibitory activity toward the corresponding proteases as wild type. Thus, these results suggest that the alteration in inhibitory specificity of P14C/N39C may be due to the conformational changes introduced by the nonnatural disulfide bond between residues 14 and 39 but not by a mere amino acid substitution at the P_5 site. On the other hand, the Tyr^{11}−Ala^{12} bond in P12A is expected to adopt the trans configuration, differing from the cis configuration of the respective Tyr^{11}−Pro^{12} bond in natural OMSVP3, although the configuration of this peptide bond is not yet analyzed. Our kinetic data suggest that the cis configuration in the Tyr^{11}−Pro^{12} bond in OMSVP3 may not be essential for the inhibitory activity for serine proteases. Then we tried to determine the solution structure of P14C/N39C, as well as that of wild-type OMSVP3, by two-dimensional nuclear magnetic resonance (2D NMR) methods and compare their structures to elucidate the structural basis of the inhibitory specificity change.

Resonance Assignments and Secondary Structure. Essentially complete ^{1}H resonance assignments were obtained for wild-type OMSVP3 and P14C/N39C using the spin system identification and sequential assignment procedure (*51*). The spin systems of all amino acid residues were identified via through-bond connectivities observed in DQF-COSY and TOCSY spectra. Any ambiguous peaks due to overlapping with the water resonance or degeneration in chemical shift were resolved by comparison of NOESY and TOCSY spectra at three different temperatures, namely, 25, 30, and 35 °C. Sequential assignments were obtained using through-space connectivities NH(i)−NH(i + 1) (d_{NN}), Hα-(i)−NH(i + 1) ($d_{\alpha N}$), and Hβ(i)−NH(i + 1) ($d_{\beta N}$) observed in the NOESY spectra. In these assignments, Hα(i)−Hδ(i + 1: Pro) ($d_{\alpha\delta}$) or Hα(i)−Hα(i + 1: Pro) ($d_{\alpha\alpha}$) NOEs instead of $d_{\alpha N}$ were used for the Pro residues. Wild-type OMSVP3 has three proline residues (Pro^{12}, Pro^{14}, and Pro^{22}), while P14C/N39C has two (Pro^{12} and Pro^{22}). In wild type, Pro^{12} gave a strong $d_{\alpha\alpha}$ NOE, while Pro^{14} and Pro^{22} gave $d_{\alpha\delta}$ NOEs, indicating that the first Pro residue adopts the cis configuration, while the latter two adopt trans configurations. In P14C/N39C, Pro^{12} and Pro^{22} both adopt the trans configuration, as evidenced by the observation of $d_{\alpha\delta}$ NOEs. These findings clearly show that introduction of the non-natural Cys^{14}−Cys^{39} bond into the OMSVP3 molecule results in conversion of the Tyr^{11}−Pro^{12} *cis*-peptide linkage to the trans configuration. The resonance assignments were extended by determining stereospecific assignments of many methylene protons and methyl protons of Val in order to obtain high-precision NMR structures. Stereospecific assignments of β-methylene protons were obtained for 18 out of 36 residues of wild-type OMSVP3 and for 17 out of 37 residues of P14C/N39C. The γ-methyl proton signals of two Val residues out of four in both proteins were stereospecifically assigned from the qualitative $^3J_{H\alpha H\beta}$ coupling constant and NOEs between α−β and α−γ protons. ^{1}H resonance assignments for wild-type OMSVP3 and P14C/N39C are provided in Tables S1 and S2, respectively.

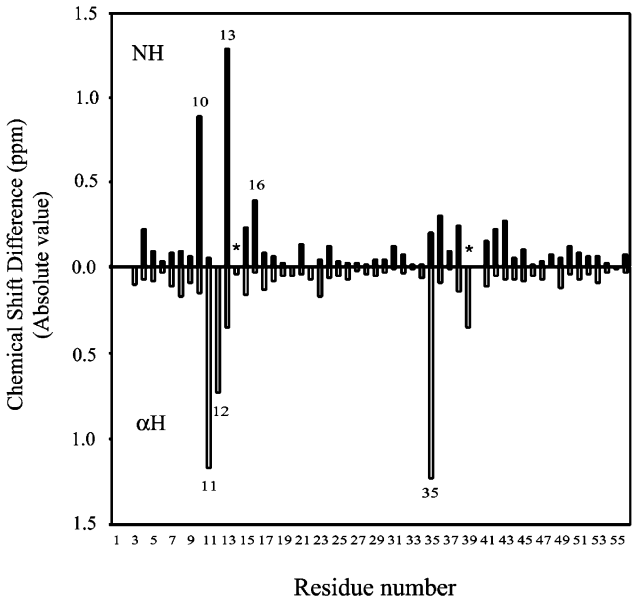

FIGURE 2: ^{1}H chemical shift difference of CαH and NH between wild-type OMSVP3 and P14C/N39C. Asterisks indicate substituted amino acid residues from Pro or Asn to Cys.

To investigate structural differences between wild-type OMSVP3 and P14C/N39C, the ^{1}H chemical shifts of CαH and NH were compared (Figure 2). Significant differences (>0.35 ppm) are detected for the localized region from Glu^{10} to Cys^{16}, and also at Cys^{35}, indicating that the two proteins assume different conformations for the N-terminal region and almost the same structure for the C-terminal region. This idea is supported by the observation that both proteins contain

Table 3: Structural Statistics[a]

	OMSVP3 wild type		P14C/N39C	
	⟨SA⟩	⟨SA⟩r	⟨SA⟩	⟨SA⟩r
total distance restraints	631		614	
intraresidue	111		114	
sequential	174		180	
medium ($1 < \|i - j\| < 5$)	117		101	
long ($\|i - j\| \geq 5$)	179		167	
hydrogen bond (two per bond)	50		52	
total dihedral angle restraints	60		68	
ϕ	29		36	
ψ	11		13	
χ_1	20		19	
rmsd from experimental restraints				
distance restraints (Å)	0.021 ± 0.003	0.017	0.020 ± 0.005	0.014
dihedral angles (deg)	0.169 ± 0.085	0.079	0.291 ± 0.162	0.197
rmsd from idealized geometry				
bonds (Å)	0.003 ± 0.000	0.002	0.002 ± 0.000	0.002
angles (deg)	0.638 ± 0.019	0.627	0.555 ± 0.029	0.530
impropers (deg)	0.373 ± 0.015	0.372	0.357 ± 0.043	0.312
energies (kcal/mol)				
total	127.5 ± 12.6	113.3	100.5 ± 17.1	82.5
bond	5.86 ± 0.95	4.71	4.71 ± 0.93	3.78
angle	88.9 ± 5.6	85.6	66.5 ± 7.2	60.5
impropers	8.77 ± 0.72	8.71	8.02 ± 2.24	6.03
van der Waals[b]	8.62 ± 3.52	4.59	7.97 ± 2.80	5.07
dihedral[c]	0.13 ± 0.13	0.02	0.46 ± 0.49	0.16
NOE[c]	15.1 ± 5.19	9.73	12.8 ± 6.78	6.83
Ramachandran plot analysis[d]				
% most favored region	65.9	68.9	55.5	56.5
% additionally allowed	31.3	26.7	41.7	39.1
% generously allowed	1.9	4.4	2.3	4.3
% disallowed	0.9	0.0	0.4	0.0
average rms difference to mean structure (Å)				
all (1−54) for backbone	1.04 ± 0.22		1.00 ± 0.16	
all (1−54) for all heavy atoms	1.28 ± 0.15		1.45 ± 0.16	
secondary structure region[e] for backbone	0.33 ± 0.10		0.44 ± 0.14	
secondary structure region[e] for all heavy atoms	0.70 ± 0.06		0.92 ± 0.19	

[a] ⟨SA⟩ represents the 15 and 15 individual structures for OMSVP3 and P14C/N39C calculated with the X-PLOR program (*41*). ⟨SA⟩r is the refined structure obtained by energy minimization of the mean structure obtained by simple averaging of the coordinates of the SA structures. [b] van der Waals energy was calculated using a final value of 4 kcal mol^{-1} Å^{-2} with the van der Waals hard sphere radii set to 0.75 times the standard values used in the CHARMM empirical energy function. [c] Dihedral and NOE energies were calculated using force constants of 200 kcal mol^{-1} rad^{-2} and 50 kcal mol^{-1} Å^{-2}, respectively. [d] The PROCHECK-NMR program (*42*) was used to assess the stereochemical parameters of the family of conformers. [e] The secondary structure region was residues 22−26, 27−31, 33−44, and 52−54.

the same secondary structure elements, an α-helix (Asn^{33}−Ser^{44}) and a three-stranded antiparallel β-sheet (Pro^{22}−Ser^{26}, Asp^{27}−Tyr^{31}, and His^{52}−Gly^{54}), deduced from NOE patterns, $^3J_{HN-H\alpha}$ coupling constants, and amide−proton exchange data (see Figure 1S). All slowly exchanging amide protons observed for wild type were also observed for P14C/N39C with the exception of Lys^{13} and Cys^{35}. The large chemical shift changes observed for the region of Glu^{10}−Cys^{16} may reflect a conformational change including the conversion of the Tyr^{11}−Pro^{12} peptide bond configuration from cis to trans. The loss of the hydrogen bond between the Lys^{13} amide proton and the side chain of Asn^{39} contributes to the chemical shift change in P14C/N39C. Our data suggest that the amide proton of Cys^{35} may form a hydrogen bond with the side chain of Asn^{33} in wild-type OMSVP3. Contrary to this, the respective hydrogen bond might be broken in P14C/N39C probably by some conformational change due to disulfide bond formation between Cys^{14} and Cys^{39}. The chemical shift of Hα in Cys^{35} for P14C/N39C moves upfield by 1.23 ppm from that for wild type and by 1.30 ppm from the random coil value (BioMagResBank, University of Wisconsin). It is worthwhile mentioning that a similar extraordinary upfield shift tendency has been observed for Hα of the respective Cys^{37} of ATI, which has the same cystine pattern as P14C/N39C (*15*). Thus, both of these extraordinary upfield shifts of Hα may be due to the newly introduced disulfide bond near them.

Tertiary Structure and Structural Comparison. Three-dimensional structures of wild-type OMSVP3 and P14C/N39C were determined by a hybrid distance geometry−dynamical simulated annealing approach based upon 691 and 682 experimental restraints, respectively, derived from NMR spectroscopy. Structural statistics for the final 15 structures and the restrained energy-minimized average structure are given in Table 3. Figures 3A and 4A show the best-fit superpositions of backbone atoms of the 15 structures for wild-type OMSVP3 and P14C/N39C, respectively. Ribbon diagrams of the corresponding restrained energy-minimized average structures are shown in Figures 3B and 4B. The solution structures of these proteins are comprised of a three-stranded antiparallel β-sheet (strand 1, Pro^{22} to Ser^{26}; strand 2, Asp^{27} to Tyr^{31}; strand 3, His^{52} to Gly^{54}) and a central α-helix (Asn^{34} to Ser^{44}) like the crystal structure of natural OMSVP3 (*7*). The rmsd for all backbone atoms relative to the mean structure is 1.0 Å for both proteins. The rmsd value decreases to 0.33 Å for wild-type OMSVP3 and to 0.44 Å

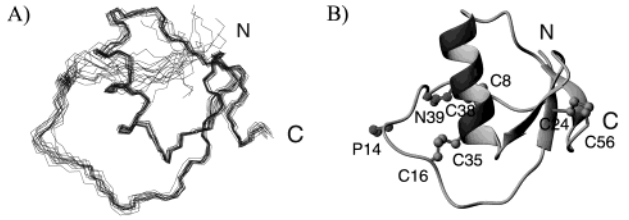

FIGURE 3: Superimposition of the best 15 structures (A) and ribbon diagram of the minimized average structure of wild-type OMSVP3 (B). Three disulfide bridges (Cys^{8}−Cys^{38}, Cys^{16}−Cys^{35}, and Cys^{24}−Cys^{56}) and the side chains of Asn^{39} and Pro^{14} are ball-and-stick representations. This figure is generated using MOLMOL (*43*).

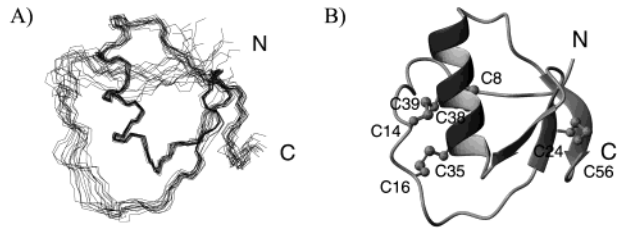

FIGURE 4: Superimposition of the best 15 structures (A) and ribbon diagram of the minimized average structure of P14C/N39C (B). Four disulfide bridges (Cys^{8}−Cys^{38}, Cys^{14}−Cys^{39}, Cys^{16}−Cys^{35}, and Cys^{24}−Cys^{56}) are ball-and-stick representations. This figure is generated using MOLMOL (*43*).

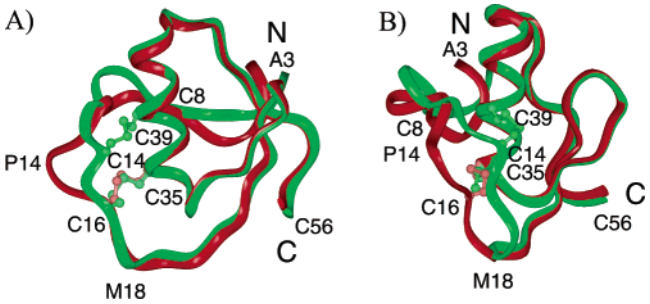

FIGURE 5: Schematic representation of main chain folding of wild-type OMSVP3 (red) overlaid with that of P14C/N39C (green) (A). Part B has been rotated 90° relative to part A. Disulfide bonds (Cys^{14}−Cys^{39}, Cys^{16}−Cys^{35}) are ball-and-stick representations. Some residues described in the text are also represented. These figures are generated by Insight II (MSI, San Diego, CA).

for P14C/N39C when the backbone atoms within the secondary structure regions are used to align the structures.

The overall structure of P14C/N39C is quite similar to that of wild-type OMSVP3. In fact, the backbones of the well-defined C-terminal portions from 19 to 56 of the two proteins are almost superimposable (Figure 5). However, the same figure shows significant structural difference in the N-terminal portion (residues 3−17). Residual backbone rmsd values are less than 1 Å for the C-terminal portion while larger than 2 Å for the N-terminal portion (see Figure S2 in detail). The most significant deviation is observed within the flexible loop from Cys^{8} to Thr^{17}, which includes the half-cystine residues of the newly introduced Cys^{14}−Cys^{39} disulfide bond as well as the preexisting Cys^{8}−Cys^{38} and Cys^{16}−Cys^{35} ones (see also Figure S2). These results are in good agreement with the ^{1}H chemical shift differences of CαH and NH between wild-type OMSVP3 and P14C/N39C. This loop is drawn approximately 4 Å closer to the central helix in P14C/N39C as compared with the same loop in wild type. A disulfide bond can generally be formed when the distance between the two α-carbons pairing cysteine residues is within the range from 4.6 to 6.6 Å (*52*). The Cα−Cα distance between Pro^{14} and Asn^{39} of wild-type OMSVP3 is calculated to be 9.2 Å in solution, which is in good agreement with the corresponding value, 8.6 Å, observed for natural OMSVP3 in the crystal (*7*). Furthermore, the side chain of Pro^{14} faces to the solvent, although that of Asn^{39} shows a good geometry for disulfide bond formation. The disulfide bond formation between positions 14 and 39 led to the large conformational changes in the loop, accompanied by side-chain rotation, main-chain movement, and the configuration change of the Tyr^{11}−Pro^{12} peptide bond from cis to trans. It is, however, worthwhile emphasizing that introduction of the disulfide bond has little effect on the C-terminal well-defined structural portion because it is designed to afford a "cystine-stabilized α-helical (CSH) motif (*17*)" to the OMSVP3 molecule by cooperating with the preexisting Cys^{16}−Cys^{35} disulfide bond.

Structural Basis of the Inhibitory Specificity Change due to Introduction of the Cys^{14}−Cys^{39} Bond. Comparison of the solution structures of wild-type OMSVP3 and P14C/N39C revealed that different structures were observed for the loop region from Cys^{8} to Thr^{17}, which is located near the reactive site of Met^{18}-Glu^{19}. The loop moves by ca. 4 Å to the central helix due to the disulfide formation. Concomitantly, the Tyr^{11}−Pro^{12} *cis*-peptide linkage converts to the trans form in P14C/N39C. This is the first example on the ovomucoid third domain containing the Tyr^{11}−Pro^{12} *trans*-peptide linkage. In the kinetic study we demonstrated that P12A, in contrast with P14C/N39C, exhibits essentially the same inhibitory activity toward the corresponding proteases as wild type. The configuration in the respective peptide linkage of this variant is expected to adopt the trans form by the replacement of Pro to Ala at position 12. The configuration of the Tyr^{11}-Ala^{12} peptide linkage in P12A has not yet been analyzed, but it will offer further valuable information on the structure−function relationship of the ovomucoid third domain. Considering our structural data together with the inhibitory activities of P12A, P14A, and P14C/N39C toward the corresponding proteases, we could conclude that the change in inhibitory specificity of P14C/N39C with the engineered disulfide bridge is not attributed to the cis-to-trans conversion in the Tyr^{11}−Pro^{12} bond.

Next we discuss the inhibitory change of P14C/N39C based on the comparison with the complex structures of OMTKY3 bound to serine proteases. So far, the complexes of OMTKY3 with CHT, SGPB, and HLE have been analyzed by X-ray crystallography (*4*, *5*, *53*), but those of the ovomucoid third domain including OMSVP3 with PPE and SGPA have not been analyzed. However, almost the same binding modes elucidated using OMTKY3 would be appropriate for those of OMSVP3, since their primary and tertiary structures are almost the same (*7*). The general binding mode is that residues P_6 through P'_3 of the inhibitor interact with subsites of the enzyme in an antiparallel β-sheet fashion with the intermolecular hydrogen bonds. In a CHT−OMTKY3 complex (*4*), there are four hydrogen bonds between them: Ser^{218} NH···CO Pro^{14i} (where i denotes an inhibitor residue); Gly^{216} NH···CO Cys^{16i}; Gly^{216} CO···NH Cys^{16i}; Ser^{214} NH···CO Leu^{18i}. In this regard, the interaction mode between the HLE and OMTKY3 complex (*53*) resembles the complex. Contrary to this, in the SGPB−OMTKY3 complex (*5*), a hydrogen bond corresponding to

Ser[218] NH···CO Pro[14i] does not occur, because this enzyme has a different conformation in the corresponding region. The respective structure of SGPA is known to be similar to that of SGPB (*54*). Thus, the hydrogen bond pattern between SGPA and OMSVP3 in the complex is likely to be the same as that between SGPB and OMTKY3.

The Pro[14]- (or Cys[14]-) Ala[15]-Cys[16] segment (or P_5-P_4-P_3) in OMSVP3 should interact with the Gly[216]-Ser[217]-Ser[218] segment of CHT and the Val[216]-Arg[217]-Ser[217A]-Gly[218] segment of PPE in an antiparallel β-sheet fashion, respectively. As described above, the Cys[8]−Cys[16] loop of P14C/N39C is drawn by approximately 4 Å toward the central helix. Therefore, it is easily assumed that the Cys[14]-Ala[15]-Cys[16] segment would no longer interact with the subsites of PPE or CHT in the same fashion as in the case for natural OMSVP3. Essentially no interaction near the P_3 site of inhibitor can be demonstrated. In P14C/N39C−CHT complex, the above segment of P14C/N39C would form only one hydrogen bond of Ser[214] NH···CO Leu[18i]. A similar assumption could be made for the case of the PPE (or SGPB)−P14C/N39C interaction. This means that P14C/N39C strongly inhibits CHT and SGPB despite the loss of several hydrogen bonds, except for Ser[214] NH···CO Leu[18i]. Thus the loss of the intermolecular hydrogen bonds between the P_3 site of the inhibitor and residue 216 of the proteases may not be the main cause of the inhibitory specificity change.

To examine the complementarity between enzymes and P14C/N39C, we performed a simple model-building study where the putative complex between P14C/N39C and CHT or HLE is constructed by superimposing the reactive site regions between P14C/N39C and OMTKY3 in the CHT−OMTKY3 complex (PDB id: 1CHO) or in the HLE−OMTKY3 complex (PDB id: 1PPF) and by exchanging the two inhibitors using the method of Read and James (*12*). In P14C/N39C−CHT complex, the above segment of P14C/N39C would form only one hydrogen bond of Ser[214] NH···CO Leu[18i]. A similar assumption could be made for the case of the elastase (or SGPB)−P14C/N39C interaction. This means that P14C/N39C strongly inhibits CHT and SGPB despite the loss of several hydrogen bonds, except for Ser[214] NH···CO Leu[18i]. Thus the loss of the intermolecular hydrogen bonds between the P_3 site of the inhibitor and residue 216 of the proteases may not be the main cause of the inhibitory specificity change. On the other hand, we found that the orientation of the whole structure of P14C/N39C relative to the reactive site region in the hypothetical complex changed compared to that of wild type. Such a change causes steric conflict between Ala[44] of HLE and Tyr[11] of P14C/N39C, while there is no bad contact between CHT and P14C/N39C. The Cys[14]−Cys[39] bond formation in P14C/N39C probably makes it impossible to adjust the conformational change in the inhibitor as required to form a tight complex with elastase. These results provide a possible explanation why P14C/N39C loses most of the inhibitory activity for elastase. To estimate the flexibility of OMSVP3 and its variant, we calculated the backbone rmsd values per residue. However, no substantial difference was found between them (Figure S2). Probably, the introduced disulfide bridge does not strongly attenuate thermal fluctuations but reduces the conformational plasticity required to fit with a enzyme. The quantitative analyses of dynamics by NMR relaxation study and molecular dynamics simulation on these complexes will give additional information about the flexibility and are in progress.

The elastases, PPE and HLE, have a general chain fold of a chymotrypsin-like serine protease, including CHT and SGPB (*53*, *55*). Most of the structural differences between the elastases and other serine proteases are located in the surface loop (*56*). Many structural determinants controlling substrate specificity reside on surface loops, and this allows for the possibility of rapid and varied evolutionary divergence with conservation of the overall tertiary fold (*55*). Thus, the structural difference of surface loops between CHT and PPE also may be related with the inhibitory specificity alteration of P14C/N39C. The surface loop structures in CHT may be able to allow the conformational change of P14C/N39C but those in PPE may not.

Finally, it was suggested in this study that the introduction of an engineered disulfide bond forming the CSH motif near the reactive site might produce an inhibitor with a narrower specificity by increasing the rigidity of the reactive site loop. It should be noted that all of the inhibitors with the CSH motif, which have been found, are classified into the Kazal-type trypsin inhibitor family. They include leech-derived tryptase inhibitor (LDTI) (*57*), rhodniin (a highly specific inhibitor of thrombin from *Rhodnius prolixus*) (*58*), budellin B-3 from *Hirudo medicinalis* (*59*), and ascidian trypsin inhibitor (*60*, *61*). Thus, these inhibitors may enhance the inhibitory specificity not only by a arginine or lysine residue as a primary determinant at the P_1 site but also by the introduction of a disulfide bond near the reactive site. Removal of an additional disulfide bridge constituting the CSH motif in these Kazal-type inhibitors, including ATI, would constitute a meaningful test of how the rigidity in the reactive site region is related to the restricted inhibitory specificity.

CONCLUSIONS

To address the structural basis of the broad specificity of ovomucoid, we prepared the disulfide variant of OMSVP3, P14C/N39C, introducing an engineered disulfide bond between positions 14 and 39 by site-directed mutagenesis. The mutation sites were chosen on the basis of the structural homology between OMSVP3 and ascidian trypsin inhibitor (ATI) with the cystine-stabilized α-helical motif in the sequence. The Cys[14]−Cys[39] bond thus designed introduces this motif into the OMSVP3 molecule by cooperating with the preexisting Cys[16]−Cys[35] disulfide bond near the reactive site. Introduction of the disulfide Cys[14]−Cys[39] bond gave the unexpected results that the disulfide variant loses almost all inhibitory activity toward elastase, although it retains the potent inhibitory activities toward CHT, SGPA, and SGPB. To elucidate the structural basis of the inhibitory specificity change, the solution structure of P14C/N39C, as well as of wild-type OMSVP3, was analyzed by 2D NMR methods. The results demonstrated that both proteins are comprised of a three-stranded antiparallel β-sheet (strand 1, Pro[22] to Ser[26]; strand 2, Asp[27] to Tyr[31]; strand 3, His[52] to Gly[54]) and a central α-helix (Asn[34] to Ser[44]) like the crystal structure of natural OMSVP3, indicating that the preexisting α-helix structure has not been broken due to introduction of the nonnatural disulfide bond. Successful introduction of the disulfide bridge into a small-sized and compact inhibitor such

as OMSVP3 is credited to the precise design of the mutation sites based on the structural homology between OMSVP3 and ascidian trypsin inhibitor with the CSH motif. Structural studies further revealed that the loop region from Cys^{8} to Thr^{17} of P14C/N39C moves by 4 Å toward the central helix and that the cis configuration in the $Tyr^{11}-Pro^{12}$ bond in wild-type OMSVP3 converted to the trans configuration in P14C/N39C. Our structural data, together with the inhibitory activities of P12A, P14A, and P14C/N39C toward the corresponding proteases, suggest that the main cause of the inhibitory specificity change would be the conformational change of the reactive site loop due to the disulfide formation, but not the configuration change in the $Tyr^{11}-Pro^{12}$ bond. The combination of the site-specific introduction of the engineered disulfide bond and spectroscopic approaches developed here should prove useful in further investigations of the relationship between the inhibitory specificity and the conformational flexibility of the ovomucoid inhibitor.

ACKNOWLEDGMENT

We are grateful to Mrs. H. Matsumoto, Center for Instrumental Analysis, Hokkaido University, for determination of the amino acid compositions and sequences of proteins.

SUPPORTING INFORMATION AVAILABLE

Two tables containing proton chemical shifts for wild-type OMSVP3 and P14C/N39C and two figures (1) summarizing the sequential NOE connectivities and (2) showing three plots of backbone rmsd per residue in 15 structures of wild-type OMSVP3 and P14C/N39C and the backbone pairwise rmsd per residue between each minimized average structure of the two proteins. This material is available free of charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Laskowski, M., Jr., and Kato, I. (1980) *Annu. Rev. Biochem. 49*, 593−626.
2. Laskowski, M., Jr., Kato, I., Ardelt, W., Cook, J., Denton, A., Empie, M. W., Kohr, W. J., Park, S. J., Parks, K., Schatzley, B. L., Schoenberger, O. L., Tashiro, M., Vichot, G., Whatkey, H. F., Wieczorek, A., and Wieczorek, M. (1987) *Biochemistry 26*, 202−221.
3. Bode, W., and Huber, R. (1992) *Eur. J. Biochem. 204*, 433−451.
4. Read, R. J., Fujinaga, M., Sielecki, A. R., and James, M. N. G. (1983) *Biochemistry 22*, 4420−4433.
5. Fujinaga, M., Sielecki, A. R., Read, R. J., Ardelt, W., Laskowski, M., Jr., and James, M. N. G. (1987) *J. Mol. Biol. 195*, 397−418.
6. Papamokos, E., Weber, E., Bode, W., Huber, R., Empie, M. W., Kato, I., and Laskowski, M., Jr. (1982) *J. Mol. Biol. 158*, 515−537.
7. Bode, W., Epp, O., Huber, R., and Laskowski, M., Jr. (1985) *Eur. J. Biochem. 147*, 387−395.
8. Krezel, A. M., Darba, P., Robertson, A. D., Fejzo, J., Macura, S., and Markley, J. L. (1994) *J. Mol. Biol. 242*, 203−214.
9. Schechter, I., and Berger, A. (1967) *Biochem. Biophys. Res. Commun. 27*, 157−162.
10. Empie, M. W., and Laskowski, M., Jr. (1982) *Biochemistry 21*, 2274.
11. Ardelt, W., and Laskowski, M., Jr. (1985) *Biochemistry 24*, 5313−5320.
12. Read, R. J., and James, M. N. G. (1986) in *Protease inhibitors* (Barrett, A. J., and Salvesen, G., Eds.) pp 301−336, Elsevier, Amsterdam.
13. Hopkins, P. C. R., Chang, W.-S. W., Wardell, M. R., and Stone, S. R. (1997) *J. Biol. Chem. 272*, 3905−3909.
14. Kojima, S., Kumagai, I., and Miura, K. (1993) *J. Mol. Biol. 230*, 395−399.
15. Hemmi, H., Yoshida, T., Kumazaki, T., Nemoto, N., Hasegawa, J., Nishioka, F., Kyogoku, Y., Yokosawa, H., and Kobayashi, Y. (2002) *Biochemistry 41*, 10657−10664.
16. Yokosawa, H., Odajima, R., and Ishii, S. (1985) *J. Biochem. 97*, 1621−1630.
17. Tamaoki, H., Miura, R., Kusunoki, M., Kyogoku, Y., Kobayashi, Y., and Moroder, L. (1998) *Protein Eng. 8*, 649−659.
18. Kojima, S., Fushimi, N., Ikeda, A., Kumagai, I., and Miura, K. (1994) *Gene 143*, 239−243.
19. Kunkel, T. A. (1985) *Proc. Natl. Acad. Sci. U.S.A. 82*, 488−492.
20. Kumazaki, T., Fujitani, A., Terasaka, K., Shimura, K., Kasai, K., and Ishii, S. (1988) *J. Biochem (Tokyo) 103*, 297−301.
21. Hirs, C. H. W. (1967) *Methods Enzymol. 11*, 197−199.
22. Hermodson, M. A., Ericsson, L. H., Neurath, H., and Walsh, K. A. (1973) *Biochemistry 12*, 3146−3153.
23. Bidlingmeyer, B. A., Cohen, S. A., and Travis, T. L. (1984) *J. Chromatogr. 336*, 93−104.
24. Laemmli, U. K. (1970) *Nature 227*, 680−685.
25. Narahashi, Y. (1970) *Methods Enzymol. 19*, 3−20.
26. Kezdt, F. J., and Kaiser, E. T. (1970) *Methods Enzymol. 19*, 3−20.
27. Kumazaki, T., Ishii, S., and Yokosawa, H. (1994) *J. Biochem. (Tokyo) 116*, 787−793.
28. Bieth, J. G. (1995) *Methods Enzymol. 248*, 59−84.
29. Marion, D., and Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 113*, 967−974.
30. Marion, D., Ikura, M., Tschundin, R., and Bax, A. (1989) *J. Magn. Reson. 85*, 393−399.
31. Rance, M., Sorensen, O. W., Bodenhausen, G., Wagner, G., Ernst, R. R., and Wüthrich, K. (1983) *Biochim. Biophys. Acta 117*, 479−485.
32. Davis, D. G., and Bax, A. (1985) *J. Am. Chem. Soc. 107*, 2820−2821.
33. Kumar, A., Ernst, R. R., and Wüthrich, K. (1980) *Biochem. Biophys. Res. Commun. 95*, 1−6.
34. Piotto, M., Saudek, V., and Sklenar, V. (1992) *J. Biomol. NMR 2*, 661−665.
35. Sklenar, V., Piotto, M., Leppik, R., and Saudek, V. (1993) *J. Magn. Reson. 102*, 241−245.
36. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) *J. Biomol. NMR 6*, 277−293.
37. Wüthrich, K., Billetter, M., and Braun, W. (1983) *J. Mol. Biol. 169*, 949−961.
38. Clore, G. M., Gronenborn, A. M., Nilges, M., and Ryan, C. A. (1987) *Biochemistry 26*, 8012−8013.
39. Wagner, G., Braun, W., Havel, T. F., Schaumman, T., Go, N., and Wüthrich, K. (1987) *J. Mol. Biol. 196*, 611−639.
40. Nilges, M., Clore, G. M., and Gronenborn, A. M. (1988) *FEBS Lett. 229*, 317−324.
41. Brünger, A. T. (1992) *X-PLOR. A system for X-ray crystallography and NMR*, Yale University Press, New Haven, CT.
42. Laskowski, R. A., Rullmann, J. A. C., MacArthur, M. W., Kaptein, R., and Thornton, J. M. (1990) *J. Biomol. NMR 8*, 477−486.
43. Koradi, R., Billeter, M., and Wüthrich, K. (1996) *J. Mol. Graphics 14*, 51−55.
44. Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946−950.
45. Kojima, S., Obata, S., Kumagai, I., and Miura, K. (1990) *Biotechnology 8*, 449−452.
46. Kowalski, D., Leary, T. R., McKee, R. E., Sealock, R. W., Wang, D., and Laskowski, M., Jr. (1974) in *Proceeding of Bayer Symposium, V. Protease Inhibitors* (Fritz, H., Tschesche, H., Greene, L. J., and Truscheit, E., Eds.) pp 311−324, Springer, Berlin.
47. Jallat, S., Carvallo, D., Tessier, L. H., Roecklin, D., Roitsch, C., Ogushi, F., Crystal, R. G., and Courtney, M. (1986) *Protein Eng. 1*, 29−35.
48. von Wilcken-Bergmann, B., Tilis, D., Sartorius, J., Auerwald, E. A., Schroder, W., and Müller-Hill, B. (1986) *EMBO J. 5*, 3219−3225.
49. Longstaff, C., Campbell, A. F., and Fersht, A. R. (1990) *Biochemistry 29*, 7339−7347.
50. Heinz, D. W., Hyberts, S. G., Peng, J. W., Priestle, J. P., Wagner, G., and Grutter, M. G. (1992) *Biochemistry 31*, 8755−8766.
51. Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, John Wiley & Sons, New York.
52. Thornton, J. M. (1981) *J. Mol. Biol. 151*, 261−287.
53. Bode, W., Wei, A.-Z., Huber, R., Meyer, E., Travis, J., and Neumann, S. (1986) *EMBO J. 5*, 2453−2458.
54. James, M. N. G., Sielecki, A. R., Brayer, G. D., and Delbaere, L. T. J. (1980) *J. Mol. Biol. 144*, 43−88.

55. Perona, J. J., and Craik, C. S. (1997) *J. Biol. Chem. 272*, 29987−29990.
56. Bode, W., Meyer, E., Jr., and Powers, J. C. (1989) *Biochemistry 28*, 1951−1963.
57. Sommerhoff, C. P., Sollner, C., Mentele, R., Piechottka, G. P., Auerswald, E. A., and Fritz, H. (1994) *Biol. Chem. Hoppe-Seyler 375*, 685−694.
58. Friedrich, T., Kroger, B., Bialojan, S., Lemaire, H. G., Hoffken, H. W., Reuscenbach, P., Otte, M., and Dodt, J. (1993) *J. Biol. Chem. 268*, 16216−16222.
59. Fink, E., Rehm, H., Gippner, C., Bode, W., Eulitz, M., Machleidt, W., and Fritz, H. (1986) *Biol. Chem. Hoppe-Seyler 367*, 1235−1242.
60. Kumazaki, T., Hoshiba, N., Yokosawa, H., and Ishii, S. (1990) *J. Biochem. 107*, 409−413.
61. Kumazaki, T., and Ishii, S. (1990) *J. Biochem. 107*, 414−419.

BI026727C