Solution Structure of Ascidian Trypsin Inhibitor Determined by Nuclear Magnetic Resonance Spectroscopy†

Hikaru Hemmi,[‡,§] Takuya Yoshida,[||,⊥] Takashi Kumazaki,[#,△] Nobuaki Nemoto,[||] Jun Hasegawa,[||,▽] Fujio Nishioka,[§] Yoshimasa Kyogoku,[||] Hideyoshi Yokosawa,[#] and Yuji Kobayashi*[,||,⊥]

National Food Research Institute, 2-1-12 Kannondai, Tsukuba, Ibaraki 305-8642, Japan, National Research Institute of Fisheries Science, 2-12-4 Fukuura, Kanazawa-ku, Yokohama, Kanagawa 236-0004, Japan, Institute for Protein Research, Osaka University, 3-2 Yamadaoka, Suita, Osaka 565-0871, Japan, Graduate School of Pharmaceutical Sciences, Osaka University, 1-6 Yamadaoka, Suita, Osaka 565-0871, Japan, and Graduate School of Pharmaceutical Sciences, Hokkaido University, Kita-ku, Sapporo 060-0812, Japan

Received April 26, 2002; Revised Manuscript Received June 26, 2002

ABSTRACT: The three-dimensional solution structure of ascidian trypsin inhibitor (ATI), a 55 amino acid residue protein with four disulfide bridges, was determined by means of two-dimensional nuclear magnetic resonance (2D NMR) spectroscopy. The resulting structure of ATI was characterized by an α-helical conformation in residues 35−42 and a three-stranded antiparallel *β*-sheet in residues 22−26, 29−32, and 48−50. The presence of an α-helical conformation was predicted from the consensus sequences of the cystine-stabilized α-helical (CSH) motif, which is characterized by an α-helix structure in the Cys-X_1-X_2-X_3-Cys portion (corresponding to residues 37−41), linking to the Cys-X-Cys portion (corresponding to residues 12−14) folded in an extended structure. The secondary structure and the overall folding of the main chain of ATI were very similar to those of the Kazal-type inhibitors, such as Japanese quail ovomucoid third domain (OMJPQ3) and leech-derived tryptase inhibitor form C (LDTI-C), although ATI does not show extensive sequence homology to these inhibitors except for a few amino acid residues and six of eight half-cystines. On the basis of these findings, we realign the amino acid sequences of representative Kazal-type inhibitors including ATI and discuss the unique structure of ATI with four disulfide bridges.

Ascidian trypsin inhibitor (ATI)[1] was first isolated from the hemolymph of the solitary ascidian *Halocynthia roretzi* (*1*). ATI strongly inhibits mammalian trypsin and plasmin. It also inhibits acrosin purified from sperm of the ascidian. However, this inhibitor has been thought to take a role in the defense mechanism through functioning in the inhibition of the activity of trypsin-like proteases which are released from hemocytes on phagocytosis (*1*). ATI has four disulfide bridges in a molecule composed of 55 amino acid residues with a sequence showing no extensive homology to other protease inhibitors (*2*, *3*). The reactive site for trypsin is the Lys^{16}−Met^{17} bond (*4*). The most striking feature of ATI is that this inhibitor has two kinds of cystine motifs in a small-sized molecule. The first motif is a cystine framework of the chelonianin family of serine protease inhibitors that contain four disulfide bridges in a domain. This is called WAP (whey acidic protein) motif (*5*, *6*). This family of protease inhibitors includes secretory leukocyte protease inhibitor (SLPI) (*7*), antileukoprotease (*8*), and elafin (*9*). Among them tertiary structures of SLPI and elafin have been elucidated (*10*−*12*). Figure 1 shows the topological relation between the disulfide bridge structure and the reactive site position of ATI and several selected serine protease inhibitors. The cystine framework of ATI is the same as that of elafin, but the reactive site position is not. It should be noticed here that three of four disulfide bridges except for the Cys^{12}−Cys^{41} bond in ATI display a topology, including the reactive site position, common to that of all three disulfide bridges in the Kazal-type inhibitors such as OMJPQ3 and bdellin B-3 from *Hirudo medicinalis* (*13*). On the basis of the findings ATI has been tentatively classified into the Kazal-type inhibitor family (*3*). The Kazal-type inhibitor family is now divided into two subgroups, i.e., the "classical" and "nonclassical" inhibitors. The former includes pancreatic secretory trypsin inhibitor (PSTI) and ovomucoids, and the

† The atomic coordinates for the 20 best conformers described in this paper have been deposited with the Protein Data Bank (accession number 1IW4).

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

‡ National Food Research Institute.

§ National Research Institute of Fisheries Science.

|| Institute for Protein Research, Osaka University.

⊥ Graduate School of Pharmaceutical Sciences, Osaka University.

Graduate School of Pharmaceutical Sciences, Hokkaido University.

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

▽ Present address: Daiichi Pharmaceutical Co., Ltd., Tokyo R&D Center, 1-16-13 Kita-Kasai, Edogawa-ku, Tokyo 134-0081, Japan.

[1] Abbreviations: ATI, ascidian trypsin inhibitor; CSH motif, cystine-stabilized α-helical motif; DQF-COSY, double-quantum-filtered chemical shift correlation spectroscopy; LDTI-C, leech-derived tryptase inhibitor form C; NMR, nuclear magnetic resonance; NOESY, nuclear Overhauser effect spectroscopy; OMJPQ3, Japanese quail ovomucoid third domain; OMTKY3, turkey ovomucoid third domain; OMSVP3, silver pheasant ovomucoid third domain; PSTI, porcine pancreatic secretory trypsin inhibitor; rmsd, root-mean-square deviation; TOCSY, total correlation spectroscopy; E-COSY, exclusive two-dimensional scalar correlation spectroscopy; 2D, two dimensional; WAP, whey acidic protein.

10.1021/bi026035o CCC: $22.00
Published on Web 07/31/2002

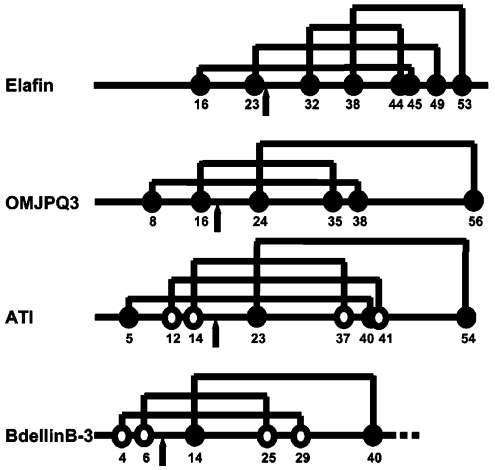

FIGURE 1: Topological relation between the disulfide bridge structure and the reactive site position of elafin, OMJPQ3, ATI, and bdellin B-3. Arrows indicate reactive sites. Circles indicate half-cystines. Among them, the open circle in ATI and bdellin B-3 especially represents half-cystines constituting the CSH motif.

latter includes bdellin B-3, the elastase-specific inhibitor from *Anemonia sulcata* (*14*), rhodniin (a Kazal-type thrombin inhibitor from *Rhodnius prolixus*) (*15*), and leech-derived tryptase inhibitor form C (LDTI-C) (*16*). The common features of nonclassical Kazal-type inhibitors are that the positioning of half-cystine residues forming disulfides is significantly different from that of the classical inhibitors, even though their sequences and the cystine pattern are homologous. Next, the second cystine motif found in ATI is the cystine-stabilized α-helical (CSH) motif, which is composed of an α-helical segment spanning the Cys-X_1-X_2-X_3-Cys sequence portion that is cross-linked by two disulfide bridges to the sequence portion Cys-X-Cys, folded in an extended β-strand-type structure. Tamaoki et al. (*17*) has pointed out that ATI, as well as most of the nonclassical Kazal-type inhibitors such as bdellin B-3, LDTI-C, and rhodoniin, has the CSH motif in the sequence (also see Figure 1). The CSH motif has been widely found in bioactive peptides, such as endothelin-1, honeybee toxins (apamin), and scorpion toxins (charybdotoxin) (*18*−*20*). The solution structures of these peptides have been determined by using NMR methods, and each Cys-X_1-X_2-X_3-Cys segment was proven to form an α-helical structure (*18*, *21*−*23*). In the case of ATI, Cys^{37}-Ala^{38}-Leu^{39}-Cys^{40}-Cys^{41} and Cys^{12}-Arg^{13}-Cys^{14} correspond to the respective segments of the CSH motif, where the disulfide bridges linking Cys^{12} to Cys^{41} and Cys^{14} to Cys^{37} are formed. ATI has the unique and crowded disulfide bridges in a small-sized molecule, but its tertiary structure has not yet been analyzed. It is worthwhile, therefore, for definite classification of this inhibitor to determine the secondary and tertiary structure of ATI and to compare it with those of protease inhibitors from various other sources. It is interesting to clarify if Cys^{40} interferes with α-helix formation in the CSH motif.

In the present study, we have determined the three-dimensional solution structure of ATI by two-dimensional (2D) NMR spectroscopy and simulated annealing calculations. The results indicated that the Cys-X_1-X_2-X_3-Cys segment in question actually forms an α-helical structure and that the secondary structure and overall folding of main chain of ATI are very similar to those of the Kazal-type inhibitor family but not to those of the chelonianin inhibitor family. On the basis of the results obtained, we will discuss the unique structure of ATI with four disulfide bridges.

MATERIALS AND METHODS

Materials. ATI was isolated from the hemolymph of *H. roretzi* by affinity chromatography on immobilized bovine trypsin, and it was further separated into ATI-I (intact form) and ATI-II (reactive site cleaved form) by ion-exchange chromatography as previously reported (*1*). ATI-I was further separated by reversed-phase high-performance liquid chromatography into ATI-Ia (minor component) and ATI-Ib (major component) (*4*). ATI-Ib thus isolated was used in this study and designated herein as ATI in brief. The purity of the final product was verified by sodium dodecyl sulfate gel electrophoresis and amino acid composition analysis after acid hydrolysis.

Measurement of NMR Spectra. A desalted and lyophilized sample was dissolved in either D_2O or a 90% H_2O/10% D_2O solution to give a final concentration of 1.5−2.0 mM. The pH of the solution was adjusted to 3.7 with 1 M HCl. All ^{1}H NMR spectra were obtained on Bruker AVANCE-500, DRX-600, and DMX-750 and JEOL α-600 spectrometers with quadrature detection in the phase-sensitive mode by TPPI (*24*) and States−TPPI (*25*). The following spectra were collected at 25, 30, and 35 °C with 15 ppm spectral widths in the t_1 and t_2 dimensions: 2D double-quantum-filtered chemical shift correlation spectroscopy (DQF-COSY) (*26*), recorded with 512 and 2048 complex points in the t_1 and t_2 dimensions; 2D nuclear Overhauser effect spectroscopy (NOESY) (*27*), 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; 2D total correlation spectroscopy (TOCSY) (*28*), recorded with mixing times of 35 and 70 ms and 512 and 2048 complex points in the t_1 and t_2 dimensions. The high digital resolution DQF-COSY and 2D exclusive scalar correlation spectroscopy (E-COSY) (*29*) spectra, respectively, were recorded using 800 and 4096 complex points in the t_1 and t_2 dimensions. Water suppression was performed using the WATERGATE sequence (*30*, *31*). For identification of the slowly exchanging amide protons the lyophilized sample was initially dissolved in 250 μL of D_2O, and then sequential 2 h 2D-TOCSY spectra were recorded at 30 °C. All NMR spectra were processed using the NMR Pipe Program (*32*). 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 4,4-dimethyl-4-silapentane-1-sulfonate (DSS).

Assignments. Proton cross-peaks were assigned on the basis of the NOESY spectrum in combination with COSY and TOCSY data by following the well-established sequence-specific methodology (*33*). Apart from the data being used directly for structure determination, a set of NOESY spectra was collected under various temperature conditions to assign the overlapping peaks. The χ_1 angles of stereospecific assignments of the β-methylene protons were obtained using standard procedures (*34*).

Structure Determination. The structures of ATI were calculated by simulated annealing using torsion angle dynamics with the program CNS (*35*). The force constants used were 75 kcal mol^{-1} Å^{-2} and 400 kcal mol^{-1} rad^{-2} for

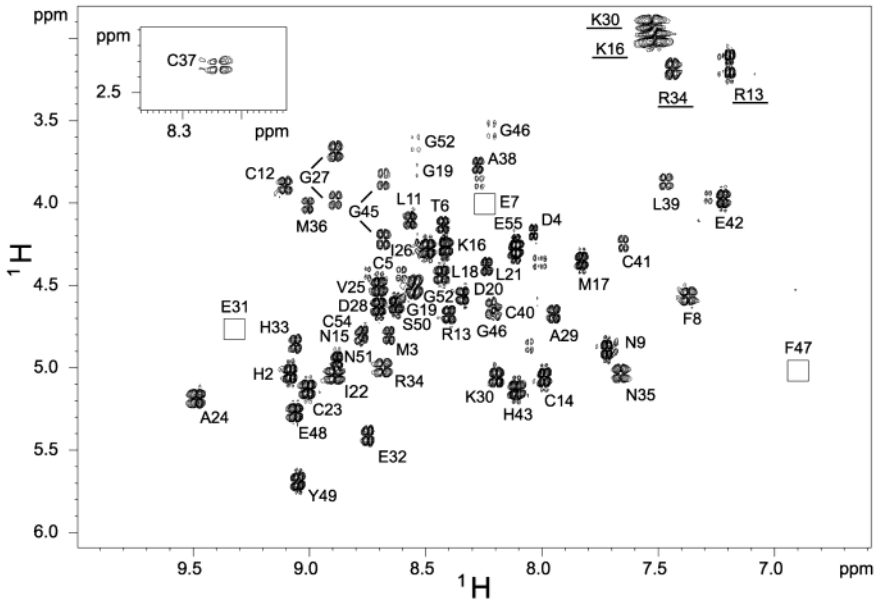

FIGURE 2: Fingerprint region of the DQF-COSY spectrum of ATI in a 90% H_2O/10% D_2O solution at pH 3.7 and 30 °C collected at 600 MHz. Boxes indicate cross-peaks that are not observed in the H_2O solution and that have been located through other evidence. Underlined assignments are for side-chain resonances. The inset box shows a peak for Cys[37] with Hα chemical shift upfield.

distance and torsion angle restraints, respectively. The structure calculation proceeded in three stages. In the first stage, a low-resolution structure was preliminarily determined using only NOE-derived distance restraints. In the second stage, the same protocol was applied by adding hydrogen bond restraints and dihedral angle restraints except for ψ angle restraints. Additional NOE constraints were added in each round of calculations, and restraints that were consistently violated were removed. In the third stage, the additional NOE constraints and ψ angle restraints were added and employed into the final structure calculation. The averaged minimized structure of ATI was obtained by restrained minimization of the mean structure of the 20 final structures with low energy.

Proton−proton distance restraints were determined from NOEs recorded in NOESY spectra with mixing times of 80, 100, and 200 ms. The NOE cross-peaks were translated into upper distance limit restraints according to their intensity by using the following qualitative criterion: strong intensity NOEs were set to a distance lower than 2.5 Å, medium intensity NOEs to lower than 3.5 Å, and weak intensity NOEs to lower than 5.0 Å. For distances involving methyl groups, nonstereospecifically assigned methylene protons, and aromatic protons, a pseudo-atom correction was added to the distance limit according to Wüthrich (*33*). Torsion angle restraints on the backbone ϕ angle were derived from $^3J_{HNH\alpha}$ coupling constants from the high digital resolution 2D DQF-COSY spectra and intraresidue and sequential NOEs. The 39 ϕ angle restraints were obtained. Backbone ϕ angles were restrained to $-60° \pm 30°$ for $^3J_{HNH\alpha} < 6$ Hz, $-120° \pm 50°$ for $^3J_{HNH\alpha} = 8-9$ Hz, and $-120° \pm 40°$ for $^3J_{HNH\alpha} > 9$ Hz. An additional ϕ angle restraint of $-100° \pm 80°$ was applied for residues for which the intraresidue Hα−HN NOE was clearly weaker than the NOE between HN and the Hα of the preceding residue (*36*). A ψ angle restraint was used for residues in α-helix and β-strand structures, as predicted from the Hα chemical shift index (CSI) (*37*), NOE patterns characteristic of secondary structure and the preliminary structures. The 15 ψ angle restraints were obtained. The ψ angle was restrained to $-40° \pm 30°$ in the α-helix regions and $120° \pm 60°$ in the β-sheet regions. Side-chain χ_1 angles were determined by $^3J_{H\alpha H\beta}$ coupling constants from E-COSY and short-mixing TOCSY connectivities in combination with NH−Hβ and Hα−Hβ NOEs (*34*). The 8 χ_1 angle restraints 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 exchanging experiments identified 14 hydrogen bond donors. 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 restraints were applied to N−H and C═O groups: 1.7−2.4 Å for the H−O distance and 2.7−3.4 Å for the N−O distance (*33*).

The final structures were analyzed using MOLMOL (*38*) and PROCHECKNMR (*39*). The molecular structures were drawn graphically using the programs MOLSCRIPT (*40*) and MOLMOL. The coordinate data for OMJPQ3 (*41*), elafin (*12*), and LDTI-C (*42*) were obtained from the Protein Data Bank.

RESULTS

NMR Analysis and Secondary Structural Elements. The sequence-specific assignments of the proton resonance from each residue in ATI were made using standard procedures (*33*) from 2D NMR spectra recorded at 25, 30, and 35 °C. The fingerprint region in the DQF-COSY spectrum at 30 °C is shown in Figure 2. The proton peak assignments of ATI in the 2D NMR spectra were completed. Signals for Pro were assigned from $d_{\alpha\delta(i,i+1)}$ connectivities, and they were identified to be all in a trans conformation. The detailed chemical shift data at pH 3.7 and 30 °C are given in the Supporting Information (Table S1). Stereospecific assignments for eight residues were then carried out. As the result,

FIGURE 3: Summary of the sequential and medium-range NOE connectivities observed for ATI and chemical shift index (CSI) derived from Hα chemical shifts of ATI. Bars, the size of which indicates the NOE intensity (strong, medium, and weak), represent sequential NOEs. Slowly exchanging amide protons are also represented as closed circles. $^3J_{HN-H\alpha}$ are three-bond coupling constants between HN and Hα, where the symbols represent <6.0 Hz (▲) and >8.0 Hz (■). The secondary structure of ATI is schematically represented.

χ_1 angles were restrained to $-60° \pm 60°$ for Asp[20], Asn[35], and Asn[51], $60° \pm 60°$ for Leu[18], Cys[41], and Tyr[49], and $180° \pm 60°$ for Asn[15] and Leu[39]. The NOE connectivities along the peptide backbone in the NOESY spectrum of ATI, $^3J_{HNH\alpha}$ coupling constants, and slowly exchanging amide protons provide its secondary structure profile as shown in Figure 3, where d_{NN}, $d_{\alpha N}$, and $d_{\beta N}$ symbolize the connectivities between NH(i)−NH(i+1) and Hβ(i)−NH(i+1) and so on. Particular $d_{\alpha N(i,i+3)}$, $d_{\alpha N(i,i+4)}$, and $d_{\alpha\beta(i,i+3)}$ connectivities observed in the sequence region 35−42 strongly suggest an α-helical structure in this region. Successive d_{NN} connectivities, small $^3J_{HNH\alpha}$ coupling constants (<6 Hz), and the slowly exchanging amide protons in the same region further support the presence of an α-helical structure. On the other hand, judging from strong $d_{\alpha N(i,i+1)}$ and $d_{\beta N(i,i+1)}$, three peptide segments, i.e., residues 22−26, 29−32, and 48−50, are considered to be in the extended form, and long-range NOEs between CαH protons among these three segments suggest that they take a part of an antiparallel β-sheet. Large $^3J_{HNH\alpha}$ coupling constants (>8 Hz) and the slowly exchanging amide protons in the same region further support the presence of an antiparallel β-sheet structure. Additional support for helix and β-sheet conformation is provided by analysis of the Hα chemical shift index (also see Figure 3). The Hα resonances of most residues in the sequence 35−42 gave particular upfield shift, while those in the sequences 22−26, 29−32, and 48−50 gave particular downfield shift from their random coil values. It should be noted that the chemical shift of Hα in Cys[37] remarkably moved to upfield from the random coil value. We cannot explain the cause of this high upfield shift at present. Recently, we prepared an OMSVP3 variant with a non-native disulfide bridge between positions 14 and 39, corresponding to a disulfide bond between positions 12 and 41 of ATI, by site-directed mutagenesis. Such a high upfield shift was also observed in Hα in Cys[35] of the variant, corresponding to that of Cys[37] of ATI. It is obvious from the above results that the two segments of Cys[37]-Ala[38]-Leu[39]-Cys[40]-Cys[41] and Cys[12]-Arg[13]-Cys[14] take an expected structure

FIGURE 4: Comparison of the secondary structural elements of ATI (a), OMJPQ3 (b), and PSTI (c). Box and arrowheads represent α-helix and β-strands, respectively. Vertical arrowheads indicate the reactive site.

satisfying the CSH motif. Present results revealed that Cys[40] at the X_3 position in the CSH motif, forming a disulfide bridge with Cys[5], does not interfere with the α-helical formation of the corresponding sequence portion in ATI. There have been two other examples reported so far that a Cys residue occurs in the amino acid sequence as X_1-X_2-X_3 between the two Cys residues. Insectotoxin I_5A (*43*) and chlorotoxin (*44*) contain an additional disulfide bond that bridges the Cys-X-X-Cys-Cys sequence portion in the α-helix to the additional N-terminal β-strand. We can identify a β-turn at residues 26−29 on the basis of characteristic NOE patterns. The first and second β-strands are thought to be connected to each other by a type I or I′ β-hairpin loop consisting of two residues, Gly[27]-Asp[28], since the $d_{\alpha N(i,i+2)}$, $d_{NN(i,i+2)}$, and $d_{\alpha N(i,i+3)}$ NOEs in the sequence region, 26 to 29, are observed. The secondary structure of ATI is schematically represented in Figure 4, showing that ATI exhibits a remarkable pattern similar to those of typical or classical Kazal-type inhibitors such as OMJPQ3 and PSTI. The α-helix of ATI is somewhat shorter than that of typical Kazal-type inhibitors, probably due to Pro[44] as an α-helix breaker. On the other hand, the secondary structure of elafin with a WAP motif contains a one-turn 3_{10}-helix followed

Table 1: Statistics for Best 20 NMR Structures of Ascidian Trypsin Inhibitor

total distance restraints	536
intraresidue	77
sequential	176
medium ($1 < \|i - j\| < 5$)	89
long ($\|i - j\| \geq 5$)	166
hydrogen bond (two per bond)	28
total dihedral angle restraints	62
ϕ	39
ψ	15
χ_1	8
rmsd from experimental restraints	
NOE distance restraints (Å)	0.0353 ± 0.0019
dihedral angle restraints (deg)	0.4269 ± 0.1094
rmsd from ideal covalent geometry	
bonds (Å)	0.0038 ± 0.00023
angles (deg)	0.5438 ± 0.0163
impropers	0.3616 ± 0.0349
ϕ and ψ in core and allowed regions (%)[a]	96.2
rmsd relative to mean structure (Å)	
whole molecule (residues 1−55)	
backbone (N, Cα, and C′ atoms)	1.163 ± 0.231
all non-H	1.873 ± 0.276
core region (residues 21−50)[b]	
backbone (N, Cα, and C′ atoms)	0.682 ± 0.105
all non-H	1.283 ± 0.113

[a] The program PROCHECK-NMR (*39*) was used for Ramachandran plot analysis. [b] Core region involves all secondary structures, one α-helix (residues 35−42), and three β-strands (residues 23−26, 29−32, and 48−50).

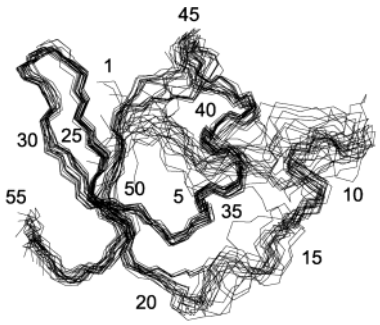

FIGURE 5: Superposition of 20 converged structures of ATI determined by ^{1}H NMR and simulated annealing calculations.

by a two-stranded antiparallel β-sheet. Thus, the secondary structural topology of ATI, $\beta\beta\alpha\beta$, is clearly different from that of elafin, $3_{10}\beta\beta$ (*5*, *9*).

Tertiary Structure. The three-dimensional structure of ATI was determined by simulated annealing calculations using 508 NOE-derived distance restraints (including 77 intraresidue, 176 sequential residue, 89 medium range, and 166 long range), 28 hydrogen bond restraints, and 62 dihedral angle restraints. Twenty conformations that give low conformation energy and give no distance and dihedral angle violations greater than 0.5 Å and 5°, respectively, were chosen among 100 resulting structures. Statistical data for the 20 structures of ATI are given in Table 1. The structures thus obtained had good covalent geometry and stereochemistry, as evidenced by the low rmsd values for bond, angle, and improper from idealized geometry. The Ramachandran plot confirmed the high quality of these structures, which showed that 88% of ϕ and ψ angles are found within the most favored and additionally allowed regions and that less than 4% of ϕ and ψ angles are found within the disallowed region. Figure 5 represents the resulting solution structures of ATI, where these structures are superimposed to give the best fit in space. The rmsd value from the mean structure is 1.16 Å for all backbone atoms in the whole molecule, while the corresponding value is 0.68 Å for all backbone atoms in the core region of residues 21−51, involving all secondary structure elements of ATI. These data indicate that the core region converges very well in the calculated structures. Figure 6 gives the schematic drawing of the mean structure of the peptide backbone of ATI, together with those of OMJPQ3, LDTI-C, and elafin. The mean structure of ATI is well characterized by an α-helix in the sequence of Asn^{35}−Glu^{42} and a three-stranded antiparallel β-sheet consisting of the peptide segments of Ile^{22}−Ile^{26}, Ala^{29}−Glu^{32}, and Glu^{48}−Ser^{50}. The gross structure of ATI is very similar to those of OMJPQ3 and LDTI-C but not to elafin, which is characterized by a central twisted β-hairpin accompanied by two external segments linked by the protease-binding loop. The results thus suggest that ATI could be classified into a Kazal-type inhibitor family on the basis of tertiary structure.

DISCUSSION

So far, serine protease inhibitors have been classified not only by the homology in amino acid sequences but also by the similarity in topological relation between the disulfide bridge pattern and the reactive site position (*45*). Structural analysis of these inhibitors using X-ray crystallography and NMR methods revealed that the inhibitor molecules thus classified have a common characteristic tertiary structure (*46*, *47*). It would be apparent that ATI examined in this study could not be easily classified on the basis of the data of sequence homology and cystine framework. For definite classification, it was necessary to determine the tertiary structure and to compare it with those of other familiar inhibitors. In this study we have determined the tertiary structure of ATI by 2D NMR methods. The resulting structure of ATI was well characterized by an α-helix in the sequence of Asn^{35}−Glu^{42} and a three-stranded antiparallel β-sheet consisting of the peptide segments of Ile^{22}−Ile^{26}, Ala^{29}−Glu^{32}, and Glu^{48}−Ser^{50}. This structural profile is quite similar to those of OMJPQ3 and PSTI, suggesting that ATI can be classified as a member of Kazal-type inhibitors from the viewpoint of tertiary structure. This study further confirmed that ATI is clearly different from a member of the chelonianin inhibitor family, even though both inhibitors have a common cystine framework consisting of four disulfide bridges.

The Kazal-type inhibitor family is now divided into two subgroups, the classical and nonclassical inhibitors. The former includes OMJPQ3 and PSTI, and so on, and the latter includes bdellin B-3 (*13*), *Anemonia* elastase inhibitor (*14*), rhodniin (*15*), and LDTI-C (*16*). Their sequences have been so far aligned with those of the classical inhibitors by placing the half-cystines in corresponding positions introducing deletion arbitrarily to improve the alignment (*13*, *14*). Fink et al. have already pointed out that the number of amino acids between the first and second half-cystine, as well as the fourth and fifth, is apparently variable between bdellin B-3 and various classical Kazal-type inhibitors. Tamaoki et al. (*17*) have found the CSH motif in the sequences of bdellin B-3, LDTI-C, and rhodniin domain 1, as well as ATI. The solution and crystal 3D structures of LDTI-C (*11*, *42*) and rhodniin (*48*) have been determined. In LDTI-C the Cys-

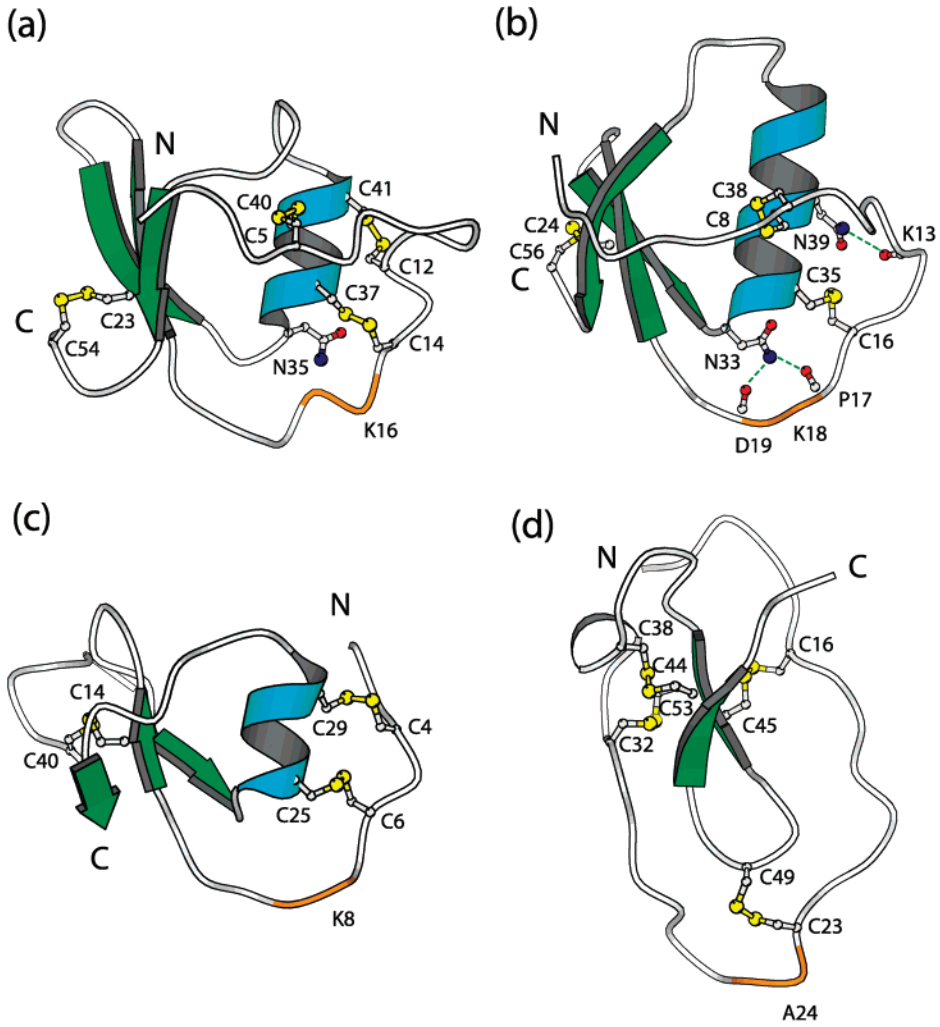

FIGURE 6: Schematic representation of the peptide backbone structures of ATI (a), OMJPQ3 (b), LDTI-C (c), and elafin (d) drawn with MOLSCRIPT (*40*). The disulfide bridges and the side chains of three asparagine residues (N35 in ATI, N33 and N39 in OMJPQ3) are shown by ball-and-stick representation. The main chains (C′, O, and N) of other residues are also shown by ball-and-stick representation. Hydrogen bonds are shown by dotted lines. Note that $Cys^{12}-Cys^{41}$ plus $Cys^{14}-Cys^{37}$ in ATI (a) and $Cys^{4}-Cys^{29}$ plus $Cys^{6}-Cys^{25}$ in LDTI-C (c) constitute the CSH motif. The reactive site residues are indicated by K16, K18, K8, and A24, respectively.

X_1-X_2-X_3-Cys portion of this motif is folded into a short 3_{10}-helix conformation in solution and of α-helix in crystals of the LDTI−trypsin complex. Rhodniin possesses two Kazal-type inhibitor domains in the molecule, and the corresponding portion of rhodniin domain 1 is also folded into an α-helix conformation. Although structural analysis of bdellin B-3 has not been carried out, its tertiary structure has been modeled by using the coordinates of the homologous OMSVP3. In this modeled structure the Cys-X_1-X_2-X_3-Cys portion is folded into an α-helical conformation (*13*). On the basis of the cystine framework including the CSH motif and the information of their tertiary structures, we tried to realign the amino acid sequences of representative Kazal-type inhibitors and ATI. The results are shown in Figure 7. In brief, the classical Kazal-type inhibitors such as PSTI, OMJPQ3, and OMSVP3 have the cystine framework of I−V, II−IV, and III−VI. ATI has an extra disulfide bridge of α−β between residues Cys^{12} and Cys^{41} in addition to these three disulfide bridges. On the other hand, bdellin B-3, LDTI-C, and rhodoniin, which are classified into the nonclassical inhibitors with the CSH motif, do not have the disulfide bond corresponding to I−V in the classical inhibitor. The disulfide bridge between the first and fifth half-cystine in these nonclassical inhibitors would correspond to an extra disulfide bridge of α−β in ATI. Thus our alignment could dissolve discrepancy in the numbers of amino acids among the Kazal-type inhibitors described above.

Finally, we would like to discuss the reactive site loop of ATI. In the case of OMJPQ3, two hydrogen bonds provided by Asn^{33} and Asn^{39} on the central helix, together with $Cys^{8}-Cys^{38}$ (I−V) and $Cys^{16}-Cys^{35}$ (II−IV), are thought to play a role in stabilizing the reactive site conformation (*49*). From our alignment shown in Figure 7, it is clear that the respective residues in ATI are Asn^{35}, Cys^{41}, $Cys^{5}-Cys^{40}$ (I−V), and $Cys^{14}-Cys^{37}$ (II−IV). From comparison of the tertiary structure of ATI with that of OMJPQ3, the spatial locations of these residues in both proteins are very similar each other (Figure 6). Thus, disulfide bonds I−V and II−IV in ATI could be expected to contribute to stabilization of the reactive site loop of this inhibitor. On the other hand, the side-chain amide proton of Asn^{33} in OMJPQ3 and PSTI forms hydrogen bonds with the main-chain carbonyl oxygen of Pro^{17} and Asp^{19} for OMJPQ3 or Pro^{17} and Ile^{19} for PSTI (*41*, *50*). Thus, Asn^{33} of the Kazal-type inhibitors acts as a spacer between the reactive site loop and the central helix (*46*). NMR studies showed that the Hδ2 of Asn^{33} in OMTKY3 and PSTI gives

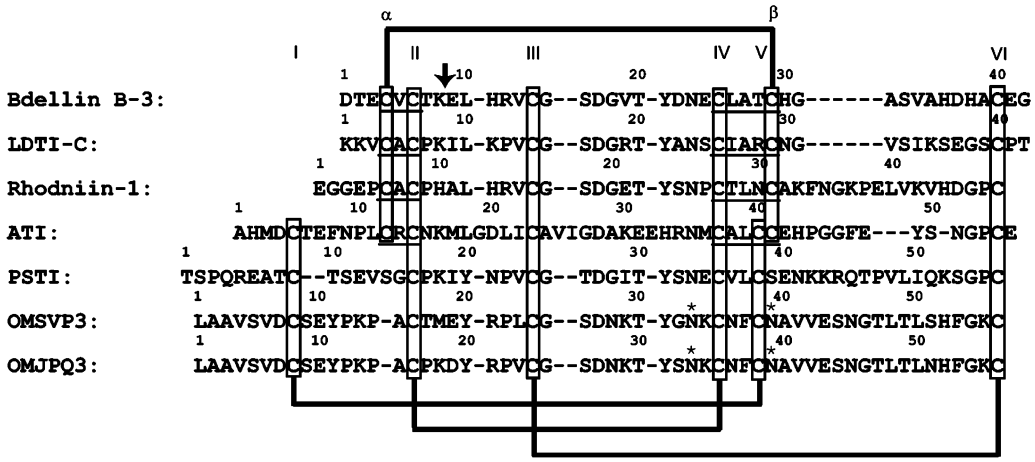

FIGURE 7: Alignment of amino acid sequences of ATI and some selected classical and nonclassical Kazal-type inhibitors. The sequence of ATI was aligned with those of Kazal-type inhibitors referring to the data obtained by CLUSTAL-W through Genome Net at the Institute for Chemical Research, Kyoto University. The arrow indicates the reactive site. Disulfide pairs are I−V, α−*β*, II−IV, and III−VI. The sequence regions assigned to the CSH motif are shown in underlines. Asterisks represent residues that play a role in stabilizing the reactive site conformation of OMSVP3 and OMJPQ3. References: bdellin B-3 (*13*), LDTI-C (*16*), rhodniin domain 1 (*15*), ATI (*3*), PSTI (*50*), and OMSVP3 and OMJPQ3 (*52*).

a considerable low-field chemical shift, favoring its participation in the mentioned hydrogen bond (*49*, *51*). In the case of Asn[35] in ATI, however, such a low-field chemical shift of the Hδ2 was not observed (see Table S1). This finding suggests that a similar contribution in stabilization of the reactive site loop by forming the hydrogen bonds may not be expected for Asn[35] in ATI. On the other hand, Asn[39] in OMJPQ3 replaces Cys[41] in ATI (also see Figure 7). This is the case for the nonclassical Kazal-type inhibitors with the CSH motif. In OMJPQ3, the side-chain amide proton of Asn[39] on the central helix forms a hydrogen bond with the main-chain carbonyl oxygen of Lys[13] on the reactive site loop (*41*). Instead of the hydrogen bond, ATI, together with the nonclassical Kazal-type inhibitors described above, has a disulfide bridge of α−*β* between Cys[41] and Cys[12] for stabilization of the reactive site conformation. It should be noted that all of the inhibitors with the CSH motif are classified into the family of trypsin inhibitors and that rhodniin and LDTI-C, the members of the nonclassical Kazal-type inhibitors, specifically and strongly inhibit thrombin and mast cell tryptase, respectively. The physiological target protease of ATI has not been identified yet, but it is expected that the target may be some specialized trypsin-like protease. Recently, we succeeded in preparing an OMSVP3 variant with a non-native disulfide bridge between positions 14 and 39, corresponding to the extra disulfide bond of α−*β* in ATI, by site-directed mutagenesis. These findings would provide further justification for our alignment. Furthermore, we obtained the preliminary results that this variant scarcely inhibits elastase, even though it exhibits almost the same potent inhibitory activities toward α-chymotrypsin and *Streptomyces griseus* proteases A and B as wild-type OMSVP3. These results suggest that increase of the rigidity near the reactive site may produce an inhibitor with a narrower specificity. These results will be published elsewhere.

ACKNOWLEDGMENT

We thank Dr. T. Yamazaki, National Institute of Agrobiological Sciences, for NMR measurement on a Bruker DMX-750.

SUPPORTING INFORMATION AVAILABLE

One table containing proton resonance assignments for ATI. This material is available free charge via the Internet at http://pubs.acs.org.

REFERENCES

1. Yokosawa, H., Odajima, R., and Ishii, S. (1985) *J. Biochem. 97*, 1621−1630.
2. Kumazaki, T., Hoshiba, N., Yokosawa, H., and Ishii, S. (1990) *J. Biochem. 107*, 409−413.
3. Kumazaki, T., and Ishii, S. (1990) *J. Biochem. 107*, 414−419.
4. Kumazaki, T., Ishii, S., and Yokosawa, H. (1994) *J. Biochem. 116*, 787−793.
5. Dandekar, A. M., Robinson, E. A., Appella, E., and Qasba, P. K. (1982) *Proc. Natl. Acad. Sci. U.S.A. 79*, 3987−3991.
6. Hennighausen, L. G., and Sippel, A. E. (1982) *Nucleic Acids Res. 10*, 2677−2684.
7. Thompson, R. C., and Ohlsson, K. (1986) *Proc. Natl. Acad. Sci. U.S.A. 83*, 6692−6696.
8. Seemüller, U., Arnhold, M., Fritz, H., Wiedenmann, K., Machleidt, W., Heinzel, R., Appelhans, H., Gassen, H. G., and Lottspeich, F. (1986) *FEBS Lett. 199*, 43−48.
9. Tsunemi, M., Kato, H., Nishiuchi, Y., Kumagaye, S.-I., and Sakakibara, S. (1992) *Biochem. Biophys. Res. Commun. 185*, 967−973.
10. Grütter, M. G., Fendrich, G., Huber, R., and Bode, W. (1988) *EMBO J. 7*, 345−351.
11. Mühlhahn, P., Czisch, M., Morenweiser, R., Habermann, B, Engh, R. A., Sommerhoff, C. P., Auerswald, E. A., and Holak, T. A. (1994) *FEBS Lett. 355*, 290−296.
12. Tsunemi, M., Matsuura, Y., Sakakibara, S., and Katsube, Y (1996) *Biochemistry 35*, 11570−11576.
13. Fink, E., Rehm, H., Gippner, C., Bode, W., Eulitz, M., Machleidt, W., and Fritz, H. (1986) *Biol. Chem. Hoppe-Seyler 367*, 1235−1242.
14. Tschesche, H., Kolkenbrock, H., and Bode, W. (1987) *Biol. Chem. Hoppe-Seyler 368*, 1297−1304.
15. 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.
16. Sommerhoff, C. P., Sollner, C., Mentele, R., Piechottka, G. P., Auerswald, E. A., and Fritz, H. (1994) *Biol. Chem. Hoppe-Seyler 375*, 685−694.
17. Tamaoki, H., Miura, R., Kusunoki, M., Kyogoku, Y., Kobayashi, Y., and Moroder, L. (1998) *Protein Eng. 11*, 649−659.
18. Tamaoki, H., Kobayashi, Y., Nishimura, S., Ohkubo, T., Kyogoku, Y., Nakajima, K., Kumagaye, S., Kimura, T., and Sakakibara, S. (1991) *Protein Eng. 4*, 509−518.

19. Kobayashi, Y., Sato, A., Takashima, H., Tamaoki, H., Nishimura, S., Kyogoku, Y., Ikenaka, K., Kondo, T., Mikoshiba, K., Hojo, H., Aimoto, S., and Moroder, L. (1991) *Neurochem. Int. 18*, 525−534.
20. Kobayashi, Y., Takashima, H., Tamaoki, H., Kyogoku, Y., Lambert, P., Kuroda, H., Chino, N., Watanabe, T. X., Kimura, T., Sakakibara, S., and Moroder, L. (1991) *Biopolymers 31*, 1213−1220.
21. Pease, J. H., and Wemmer, D. E. (1988) *Biochemistry 27*, 8491−8498.
22. Takashima, H., Kobayashi, Y., Tamaoki, H., Kyogoku, Y., Lambert, P., Kuroda, H., Chino, N., Watanabe, T. X., Kimura, T., and Sakakibara, S. (1991) in *Peptides 1990* (Giralt, E., and Andreu, D., Eds.) pp 557−559, ESCOM, Leiden, The Netherlands.
23. Bruix, M., Jimenez, M. A., Santoro, J., Gonzalez, C., Colilla, F. J., Mendez, E., and Rico, M. (1993) *Biochemistry 32*, 715−724.
24. Marion, D., and Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 113*, 967−974.
25. Marion, D., Ikura, M., Tschundin, R., and Bax, A. (1989) *J. Magn. Reson. 85*, 393−399.
26. Rance, M., Sorensen, O. W., Bodenhausen, G., Wagner, G., Ernst, R. R., and Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 117*, 479−485.
27. Kumar, A., Ernst, R. R., and Wüthrich, K. (1980) *Biochem. Biophys. Res. Commun. 95*, 1−6.
28. Davis, D. G., and Bax, A. (1985) *J. Am. Chem. Soc. 107*, 2820−2821.
29. Griesinger, C., Sorensen, O. W., and Ernst, R. R. (1987) *J. Magn. Reson. 75*, 474−492.
30. Piotto, M., Saudek, V., and Sklenar, V. (1992) *J. Biomol. NMR 2*, 661−665.
31. Sklenar, V., Piotto, M., Leppik, R., and Saudek, V. (1993) *J. Magn. Reson. 102*, 241−245.
32. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) *J. Biomol. NMR 6*, 277−293.
33. Wüthrich, K. (1986) in *NMR of Proteins and Nucleic Acids*, John Wiley & Sons, New York.
34. Wagner, G., Braun, W., Havel, T. F., Schaumann, T., Go, N., and Wüthrich, K. (1987) *J. Mol. Biol. 196*, 611−639.
35. Brünger, A. T., Adams, P. D., Clore, G. M., DeLano, W. L., Gros, P., Grosse-Kunstleve, R. W., Jiang, J.-S., Kuszewski, J., Nilges, M., Pannu, N. S., Read, R. J., Rice, L. M., Simonson, T., and Warren, G. L. (1998) *Acta Crystallogr., Sect. D 54*, 905−921.
36. Clubb, R. T., Ferguson, S. B., Walsh, C. T., and Wagner, G. (1994) *Biochemistry 33*, 2761−2772.
37. Wishart, D. S., and Sykes, B. D. (1994) *Methods Enzymol. 239*, 363−392.
38. Koradi, R., Billeter, M., and Wüthrich, K. (1996) *J. Mol. Graphics 14*, 51−55.
39. Laskowski, R. A., MacArthur, M. W., Moss, D. S., and Thronton, J. M. (1993) *J. Appl. Crystallogr. 26*, 283−291.
40. Kraulis, P. J. (1991) *J. Appl. Crystallogr. 24*, 946−950.
41. Papamokos, E., Weber, E., Bode, W., Huber, R., Empie, M. W., Kato, I., and Laskowski, M., Jr. (1982) *J. Mol. Biol. 158*, 515−537.
42. Stubbs, M. T., Morenweiser, R., Stürzenbecher, J., Bauer, M., Bode, W., Huber, R., Piechottka, G. P., Matschiner, G., Sommerhoff, C. P., Fritz, H., and Auerswald, E. A. (1997) *J. Biol. Chem. 272*, 19931−19937.
43. Lomize, A. L., Mairov, V. N., and Arseniev, A. S. (1991) *Sov. J. Bioorg. Chem. 17*, 921−940 (translated from *Bioorg. Khim. 17*, 1613−1632).
44. Lippens, G., Najib, J., Wodak, S. J., and Tartar, A. (1995) *Biochemistry 34*, 13-21.
45. Laskowski, M., Jr., and Kato, I. (1980) *Annu. Rev. Biochem. 49*, 593−626.
46. Read, R. J., and James, M. N. G. (1986) in *Proteinase Inhibitors* (Barrett, A. J., and Salvesen, G., Eds.) pp 193−208, Elsevier, Amsterdam.
47. Bode, W., and Huber, R. (1992) *Eur. J. Biochem. 166*, 673−692.
48. van de Locht, A., Lamba, D., Bauer, M., Huber, R., Friedrich, T., Kröger, B., Höffken, W., and Bode, W. (1995) *EMBO J. 14*, 5149−5157.
49. Robertson, A. D., Westler, M. W., and Markley, J. L. (1988) *Biochemistry 27*, 2519−2529.
50. Bolognesi, M., Gatti, G., Menegatti, E., Guarneri, M., Marquart, M., Papamokos, E., and Huber, R. (1982) *J. Mol. Biol. 162*, 839−868.
51. Klaus, W., and Schomburg, D. (1993) *J. Mol. Biol. 229*, 695−706.
52. 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., Whatley, H. E., Wieczorek, A., and Wieczorek, M. (1987) *Biochemistry 26*, 202−221.

BI026035O