Hensten-Pettersen, A. (1979) in *Saliva and Dental Caries, Proceedings of a Workshop* (Kleinberg, I., Ellison, S. A., & Mandel, I. D., Eds.) Special Suppl. to Microbiology Abstracts, pp 81-87, IRL, Oxford.
Isemura, S., Saitoh, E., & Sanada, K. (1982) *J. Biochem. (Tokyo) 91*, 2067-2075.
Kallner, A. (1975) *Clin. Chim. Acta. 59*, 35-39.
Kauffman, D. L., & Keller, P. J. (1979) *Arch. Oral Biol. 24*, 249-256.
Kauffman, D. L., & Keller, P. J. (1983) *Arch. Oral Biol. 28*, 61-67.
Kauffman, D., Wong, R., Bennick, A., & Keller, P. (1982) *Biochemistry 2*, 6558-6562.
Kauffman, D. L., Bennick, A., & Keller, P. J. (1986a) *J. Dent. Res. 65*, Special Issue, Abstract 86.
Kauffman, D., Hoffman, T., Bennick, A., & Keller, P. (1986b) *Biochemistry 25*, 2387-2392.
Levine, M., & Keller, P. J. (1977) *Arch. Oral Biol. 22*, 37-41.
Lowry, O. H., & Lopez, J. A. (1946) *J. Biol. Chem. 162*, 421-428.
Lyons, K. M., Stein, J. H., & Smithies, O. (1988) *Genetics 120*, 267-278.
Madapallimattan, G., & Bennick, A. (1990) *Biochem. J. 270*, 297-304.
Maeda, N. (1985) *Biochem. Genet. 23*, 455-464.
Maeda, N., Kim, H.-S., Azen, E. A., & Smithies, O. (1985) *J. Biol. Chem. 260*, 11123-11130.
Martinsen, T. M. (1984) *Methods Enzymol. 107*, 3-22.
Menshansho, H., Butler, L. G., & Carlson, D. M. (1987) *Annu. Rev. Nutr. 7*, 423-440.
Milne, E. J. & Dawes, C. (1973) *Vox Sang. 25*, 298-307.
Minaguchi, K., & Bennick, A. (1989) *J. Dent. Res. 68*, 2-15.
Moreno, E. L., Varaghese, K., & Hay, D. I. (1979) *Calcif. Tissue Int. 28*, 7-16.
Moser, P. W., & Rickli, E. E. (1979) *J. Chromatogr. 176*, 451-455.
Robinson, R., Kauffman, D. L., Waye, M. M. Y., Blum, M., Bennick, A., & Keller, P. J. (1989) *Biochem. J. 263*, 497-503.
Saitoh, E., Isemura, S., & Sanada, K. (1983a) *J. Biochem. (Tokyo) 93*, 883-888.
Saitoh, E., Isemura, S., & Sanada, K. (1983b) *J. Biochem. (Tokyo) 93*, 495-502.
Saitoh, E., Isemura, S., & Sanada, K. (1983c) *J. Biochem. (Tokyo) 94*, 1991-1999.
Schwartz, T. W. (1986) *FEBS Lett. 200*, 1-10.
Warner, T. F., & Azen, E. A. (1984) *Am. Rev. Respir. Dis. 130*, 115-118.
Wong, R. S. C., & Bennick, A. (1980) *J. Biol. Chem. 255*, 5943-5948.

^{1}H Assignments and Secondary Structure Determination of the Soybean Trypsin/Chymotrypsin Bowman-Birk Inhibitor†

Milton H. Werner‡ and David E. Wemmer*,§

Chemical Biodynamics Division, Lawrence Berkeley Laboratory, and Department of Chemistry, University of California, Berkeley, Berkeley, California 94720

Received September 28, 1990; Revised Manuscript Received January 2, 1991

ABSTRACT: The ^{1}H resonance assignments and secondary structure of the trypsin/chymotrypsin Bowman-Birk inhibitor from soybeans were determined by nuclear magnetic resonance spectroscopy (NMR) at 600 MHz in an 18% acetonitrile-d_3/aqueous cosolvent. Resonances from 69 of 71 amino acids were assigned sequence specifically. Residues Q11-T15 form an antiparallel β-sheet with residues Q21-S25 in the tryptic inhibitory domain and an analogous region of antiparallel sheet forms between residues S38-A42 and Q48-V52 in the chymotryptic inhibitory domain. The inhibitory sites of each fragment (K16-S17 for trypsin, L43-S44 for chymotrypsin) are each part of a type VI like turn at one end of their respective region of the antiparallel β-sheet. These structural elements are compared to those found in other Bowman-Birk inhibitors.

Bowman-Birk inhibitor (BBI) is a small serine protease inhibitor containing seven disulfide bonds (Birk, 1985). Isoforms of the Bowman-Birk type are found in a wide variety of leguminous plants with up to 87% sequence homology among different species (Morhy & Ventura, 1987), thus forming a distinct family of serine protease inhibitors in plants (Steiner & Frattali, 1969). BBI is an intriguing protein among the serine protease inhibitors since it is the only inhibitor capable of binding 2 equiv of serine protease/equiv of inhibitor with kinetically independent binding sites (Harry & Steiner, 1970; Odani & Ikenaka, 1972).

Classical BBI from soybeans simultaneously inhibits trypsin and chymotrypsin, although several isoinhibitors have been isolated which are capable of inhibiting either two trypsin molecules or trypsin at one site and either chymotrypsin or elastase at a second site (Odani & Ikenaka, 1977). Each is characterized by 7 disulfide bonds, an isoelectric point near 4.2, and between 60 and 76 amino acids. The trypsin/chymotrypsin inhibitor (hereafter referred to as BBI-I) may be

† This work was supported by the Office of Energy Research, Office of Health and Environmental Research, Health Effects Research Division of the U.S. Department of Energy, under contract DE-AC03-76 SF 00098 and through instrumentation grants from the U.S. Department of Energy (DE FG05-86 ER 75281) and the National Science Foundation (DMB 86-09035 and BBS 87-20134).

* To whom correspondence should be addressed.
‡ Chemical Biodynamics Division.
§ Department of Chemistry.

0006-2960/91/0430-3356$02.50/0

characterized as a relatively tight binding inhibitor with $K_I = 5 \times 10^{-9}$ (pH 8.2) for trypsin (Turner et al., 1975) and $K_I = 5.2 \times 10^{-8}$ (pH 7.98) (Harry & Steiner, 1970) for chymotrypsin. Near-UV circular dichroism does not indicate the presence of substantial regular secondary structure (α-helix or β-sheet) (Birk et al., 1980), and despite successful crystallization of BBI-I and the ternary complex of BBI-I, trypsin, and chymotrypsin, the structure of BBI-I has not yet been determined by X-ray crystallography (Wei et al., 1979; Gaier et al., 1981).

Because of the lack of structural information able to explain the unusual and potent inhibitory properties of BBI-I, we have undertaken a study to understand the nature of its interaction with serine proteases. We report here the sequential ^{1}H NMR assignments and secondary structure determination of BBI-I determined at 600 MHz. The tendency of BBI-I to self-associate (Millar et al., 1969; Harry & Steiner, 1969) required a search for solvent conditions which gave predominantly monomer at millimolar concentration. The best conditions were found to be an 18% acetonitrile/aqueous cosolvent. As it is common to assay the inhibitory properties of BBI-I in 5–10% acetone or acetonitrile (Frattali, 1969; Walsh, 1970; Wilcox, 1970), it was assumed at the outset that the presence of a cosolvent would not disrupt the native protein structure significantly. Subsequent evaluation of the NMR spectra in pure water and in cosolvent during the course of the assignment process has borne this assumption out.

Materials and Methods

Isolation and Purification of BBI-I. Commercially defatted soybean meal (Nutrasoy 7B flakes, a generous gift of the Archer Daniels Midland Co., 1.3 kg) was extracted in a 6:1 ratio (v/w) of 60% ethanol at 35 °C with stirring for 24–30 h. The solution was filtered to clarity under vacuum or centrifuged at 4000*g* and the filtrate mixed with a 2:1 ratio (v/v) of −20 °C acetone. The measured pH was lowered to 4 with concentrated HCl and allowed to stand at −20 °C for 36 h. The resulting suspension and precipitate were centrifuged at 8000*g*; the sticky pellets were resuspended in water and dialyzed against 50 mM NH_4HCO_3, pH 8, 4 °C, in a 2000 or 3500 molecular weight cutoff membrane (Spectrapor 7, Fisher Scientific) until the dialysate was clear and only slightly yellow in color (4–7 days). The dialysate was filtered, lyophilized, and stored at −20 °C prior to purification. The acetone precipitate was considered stable indefinitely prior to purification.

Purification was achieved through successive steps of ion-exchange chromatography. Initial separation was achieved by using carboxymethyl (CM)-cellulose (2.6 × 100 cm) equilibrated with 5 mM sodium acetate, pH 4, and eluted with a 0–0.5 M NaCl linear gradient (2 L each). The trypsin/chymotrypsin inhibitory fraction was further purified on diethylaminoethyl (DEAE)-cellulose (2.6 × 70 cm) equilibrated with 50 mM ammonium acetate, pH 6.5, and eluted with a linear gradient in which 0.3 M ammonium acetate, pH 5.0, was mixed into 50 mM ammonium acetate, pH 6.5 (0.35 L each). The main trypsin/chymotrypsin inhibitory fraction was further purified on Sephadex G50F (2.6 × 100 cm) equilibrated at 1 mM HCl and eluted in an ascending mode. The predominant chymotrypsin inhibitory fraction was finally purified to homogeneity on a high-performance DEAE column (0.75 × 7.5 cm) equilibrated with 10 mM sodium phosphate and eluted with a 0–0.35 M NaCl gradient in which a 50:50 mixture was achieved after 30 min of elution and a 0:100 mixture was achieved at 60 min of elution. Purity and homogeneity were determined by 15% SDS–PAGE (Laemmli, 1970), native 5% PAGE in 0.089 M Tris–borate–EDTA buffer, pH 8.4 (Frattali & Steiner, 1968), analytical ion-exchange high-performance liquid chromatography, and N-terminal sequence analysis. Assays of inhibitory activity were done essentially according to Frattali (1969).

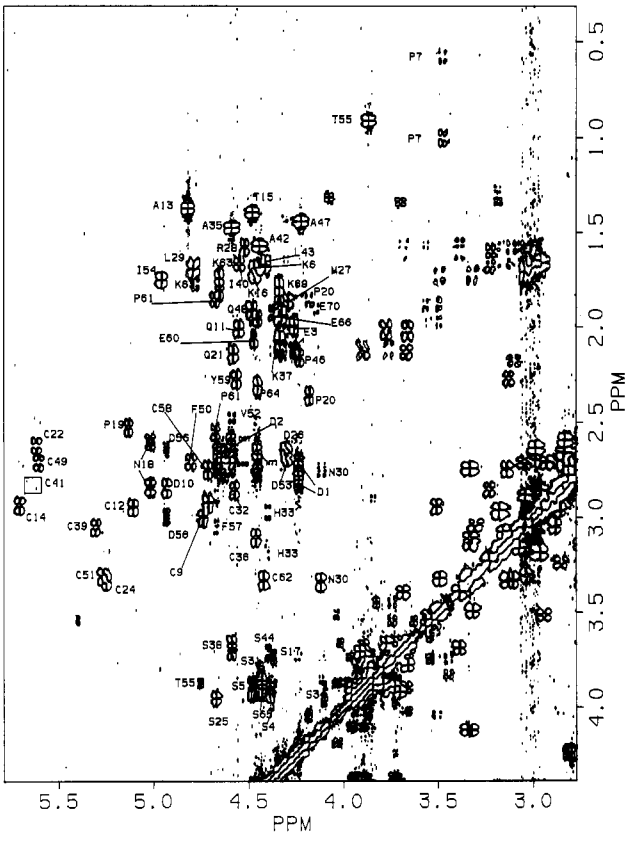

FIGURE 1: $\alpha\beta$ cross-peak region from DQF-COSY, 35 °C, pH 5.7, in D_2O showing assignments for all 71 spin systems in BBI-I.

NMR Spectroscopy. All spectra were recorded with a Bruker AMX-600 spectrometer equipped with digital-phase shifters, a ^{1}H probe, and a UX32 computer. All two-dimensional spectra were collected in the pure-phase absorption mode according to the time-proportional phase incrementation (TPPI) technique (Redfield & Kunz, 1975; Drobny et al., 1978; Bodenhausen et al., 1980) as described by Marion and Wüthrich (1983). Spectra were collected at pH 5.7 (uncorrected for the presence of acetonitrile) in 2 mM protein, 9 mM sodium phosphate buffer, and 200 μM sodium azide at 15 and 35 °C. All spectra were processed by using the program FTNMR (Hare Research, Inc.).

NOESY (Jeener et al., 1979; Macura et al., 1981) spectra were recorded in 18% acetonitrile-d_3/D_2O (D_2O) at 100 and 125 ms and at 80, 125, 150, and 200 ms in 18% acetonitrile-d_3/H_2O (H_2O). For NOESY spectra recorded in D_2O, the residual water resonance was suppressed by selective irradiation during the relaxation delay and mixing time. For NOESY spectra recorded in H_2O, the water resonance was suppressed by employing a "jump and return" sequence ($90°-\tau-90°$) (Plateau & Guéron, 1982) with phase cycling and processing according to Driscoll et al. (1989). All NOESY spectra were recorded at 15 and 35 °C except for one additional NOESY in H_2O at 200 ms which was recorded at 45 °C. Typically, 512 increments of 1K complex data were collected for spectra recorded in D_2O, and 512 increments of 2K complex data were collected for spectra recorded in H_2O. The time-domain data were multiplied by a sine-bell window with a 50° phase shift in F_2 and F_1, and all spectra were base-line-corrected with a either cubic-spline or a quadratic

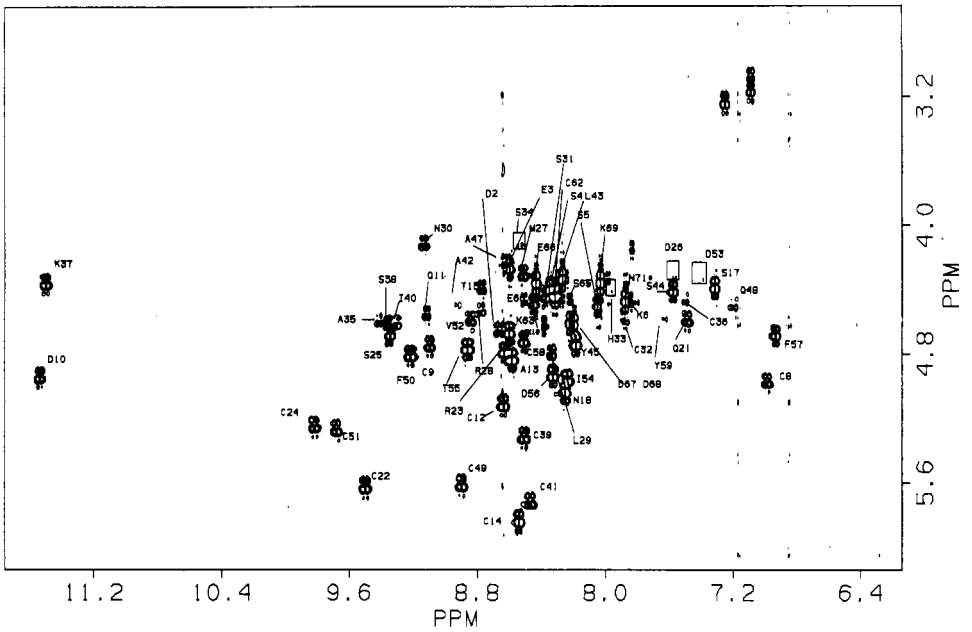

FIGURE 2: NHα cross-peak region from the DQF-COSY spectrum, 35 °C, pH 5.7, in H_2O showing the assignment of backbone amide protons in BBI-I. N71 is assigned on the basis of its spin system only.

polynomial base-line routine prior to Fourier transformation in F_1. Residual noise in F_1 was suppressed according to Klevit (1985) and Otting and Wüthrich (1986).

DQF-COSY (Rance et al., 1983) spectra were recorded at 15 and 35 °C in H_2O and at 35 °C in D_2O. Suppression of the solvent resonance was achieved by selective irradiation during the relaxation delay. A total of 512 increments of 2K complex data points were recorded for spectra collected in H_2O and 1024 increments for spectra collected in D_2O. Time-domain data were multiplied by a sine-bell window with a 30–50° phase shift in both F_2 and F_1.

To assist in the assignment of the spin systems, HOHAHA (Braunschweiler & Ernst, 1983; Bax & Davis, 1985) spectra were recorded in H_2O at 35 °C at 35- and 50-ms spin-locking period using an MLEV-17$_y$ sequence with 1.5-ms trim pulses. HOHAHA were also recorded in D_2O at 35 °C, 50 ms, in a similar manner. Spectra were recorded and processed as in DQF-COSY experiments.

Slowly exchanging amide protons were determined from 80-ms NOESY spectra in D_2O (99.9%) at pH 4.6, 20 °C. Four NOESY spectra were collected over a 24-h period with a protein sample freshly dissolved in D_2O. Slowly exchanging amide protons were identified from the NOESY spectrum collected from 18 to 24 h of exchange in D_2O.

RESULTS

Assignment of Spin Systems. The methodology for identification of amino acid residue spin systems in NMR spectra is now well established (Wagner & Wüthrich, 1982; Clore & Gronenborn, 1987; Torchia et al., 1989). Spin system assignments were made by using combinations of DQF-COSY and HOHAHA spectra in D_2O to distinguish the nonexchangable protons of three-spin, four-spin, five-spin, and greater spin systems from one another (Figure 1). HOHAHA and NOESY spectra in H_2O were then used to connect the nonexchangeable protons to their corresponding amide proton (Figure 2). During this process, all 71 spin systems could be distinguished. The collection of redundant data at 15 and 35 °C permitted the discrimination of nearly degenerate amide, α-, and β-protons that made complete assignment of the proton spin systems possible (Table I).

Since 17% of the spin systems were acidic, these were identified on the basis of the relative sensitivity of their α-proton chemical shifts to pH titration below a pH of about 3.5 in one-dimensional spectra. The multiplicity of the $\alpha\beta$ cross-peaks in DQF-COSY spectra in D_2O was used to discriminate between Asx (AMX) and Glx [AM(PT)X] spin systems. AM(PT)X spin system assignments were confirmed by assignment of the $\alpha\gamma$ relay cross-peaks in HOHAHA spectra in D_2O.

Aliphatic amino acids were assigned by using a combination of DQF-COSY and NOESY spectra collected in D_2O. Valine and isoleucine have unique spin system patterns and were clearly distinguished from one another (Wüthrich, 1986), although V52 had an unusual β-chemical shift roughly 0.7 ppm downfield from that commonly seen in proteins (Gross & Kalbitzer, 1988). L29 was found to have unusual β'- and γ-chemical shifts with the β'-proton upfield of the γ-proton (Table I). $\alpha\beta$, $\alpha\beta'$ cross-peaks and the relay cross-peak $\alpha\gamma$ in the HOHAHA spectrum permitted initial discrimination of this spin system to greater than three protons. In both DQF-COSY and HOHAHA spectra, the $\beta'\gamma$ cross-peak appeared to be upfield of the $\beta'\beta$ cross-peak (Figure 3). This was resolved by observation of the relay cross-peaks $\beta'\delta$ and $\beta'\delta'$ in the HOHAHA spectrum, confirming the identity of the $\beta'\beta'$, $\beta'\gamma$ cross-peaks, and therefore making the assignment of the spin system to leucine clear. The assignment was completed by observation of the $\alpha\delta$ and $\alpha\delta'$ cross-peaks in the 100-ms NOESY spectrum in D_2O. For L43, the β-, β'-, and γ-protons appear to be completely degenerate (Figure 3). The basis for this spin system assignment was the observation of α to methyl cross-peaks in the NOESY spectrum in D_2O, which distinguished this spin system as either valine or leucine. The final assignment to that of L43 could only be accomplished by sequential assignment.

T55 was assigned by observation of the $\alpha\beta$ and $\beta\gamma$ cross-peaks in DQF-COSY and the corresponding $\alpha\gamma$ NOE in NOESY spectra in D_2O. However, T15 was observed to have its α-proton chemical shift upfield of its β-proton. A putative $\alpha\beta$ cross-peak was observed in both DQF-COSY and HOHAHA spectra in D_2O, but final confirmation of this assignment again came through the sequential connectivities.

Phenylalanine and tyrosine were assigned on the basis of

Table I: ^{1}H Assignments for BBI-I (ppm)[a]

residue	NH	α	β	γ	δ	ε	ζ
D1		4.24	2.82, 2.74				
D2	8.66	4.67	2.75, 2.65				
E3	8.60	4.26	2.12, 2.00	2.33			
S4	8.35	4.39	3.96				
S5	8.05	4.50	3.91, 3.87				
K6	7.85	4.46	1.72, 1.67	1.38	2.96		
P7		3.48	1.01, 0.59	1.63, 1.56	3.69, 3.40		
C8	6.99	4.98	3.19, 2.95				
C9	9.09	4.74	3.02, 2.99				
D10	11.54	4.94	2.84, 2.67				
Q11	9.11	4.46	2.17, 2.01	2.28		6.66, 7.47	
C12	8.63	5.11	3.52, 2.95				
A13	8.58	4.82	1.37				
C14	8.54	5.83	2.96, 2.86				
T15	8.77	4.44	4.49	1.40			
K16	8.45	4.49	2.00, 1.70	1.52, 1.43	3.02		
S17	7.32	4.37	3.90, 3.74				
N18	8.25	5.04	2.84, 2.60		6.79, 7.48		
P19		5.14	2.53, 2.11	1.99, 1.89	3.57, 3.55		
P20		4.19	2.37, 1.86	2.12, 2.09	3.91, 3.76		
Q21	7.48	4.59	2.16, 1.78	2.24		6.67, 7.42	
C22	9.50	5.63	2.62, 2.47				
R23	8.63	4.78	1.78, 1.69	1.60			
C24	9.81	5.25	3.35, 3.15				
S25	9.34	4.67	3.95				
D26	7.58	4.30	2.67				
M27	8.51	4.30	1.99, 1.88				
R28	8.77	4.52	1.84, 1.58				
L29	8.34	4.79	1.69, 1.21	1.61	0.84, 0.72		
N30	9.13	4.12	3.35, 2.74		6.66, 7.56		
S31	8.38	4.45	3.91, 3.80				
C32	7.88	4.59	3.24, 2.85				
H33	7.98	4.40	3.19, 2.98		7.54	7.22	
S34	8.54	4.10	3.96, 3.88				
A35	9.42	4.60	1.48				
C36	7.49	4.19	3.11				
K37	11.50	4.34	2.08, 1.88	1.62, 1.50	3.07		
S38	9.36	4.59	3.72, 3.67				
C39	8.51	5.31	3.05, 2.72				
I40	9.30	4.56	1.67	1.48, 1.18 0.84 CH_3	0.80		
C41	8.47	5.72	2.94, 2.72				
A42	8.92	4.45	1.57				
L43	8.34	4.42	1.67	1.67	0.96, 0.89		
S44	7.58	4.40	3.89, 3.70				
Y45	8.18	4.71	2.90		7.15	6.81	
P46		4.30	2.16, 1.81	1.80, 1.74	3.50, 3.33		
A47	8.63	4.23	1.45				
Q48	7.19	4.48	2.19, 1.91	2.34			
C49	8.89	5.61	2.88, 2.71				
F50	9.22	4.81	2.99, 2.70		7.07	7.18	7.00
C51	9.68	5.27	3.31, 3.04				
V52	8.84	4.59	2.48	0.90, 0.82			
D53	7.40	4.35	2.65				
I54	8.24	4.96	1.75	1.72, 1.09 0.90 CH_3	0.84		
T55	8.86	4.75	3.87	0.92			
D56	8.33	4.94	3.00, 2.64				
F57	6.94	4.67	3.05, 2.89		7.00	6.90	7.15
C58	8.52	4.72	3.14, 2.75				
Y59	7.63	4.58	3.13, 2.28		7.11	6.79	
E60	8.51	4.48	2.16, 2.06	2.44			
P61		4.68	2.55, 1.85	2.14	3.90, 3.67		
C62	8.32	4.43	3.34, 3.11				
K63	8.60	4.65	1.83, 1.74	1.46	2.98		
P64		4.46	2.31, 1.95	2.03	3.78, 3.66		
S65	8.23	4.44	3.90, 3.85				
E66	8.44	4.34	2.13, 1.95	2.31			
D67	8.23	4.60	2.73, 2.60				
D68	8.19	4.60	2.73				
K69	8.03	4.34	1.89, 1.80	1.45	3.02		
E70[b]		4.15	1.91				
N71[b]	7.88	4.46	2.76, 2.65				

[a] pH 5.7, 35 °C, referenced to TSP. [b] Assignment based on spin system only.

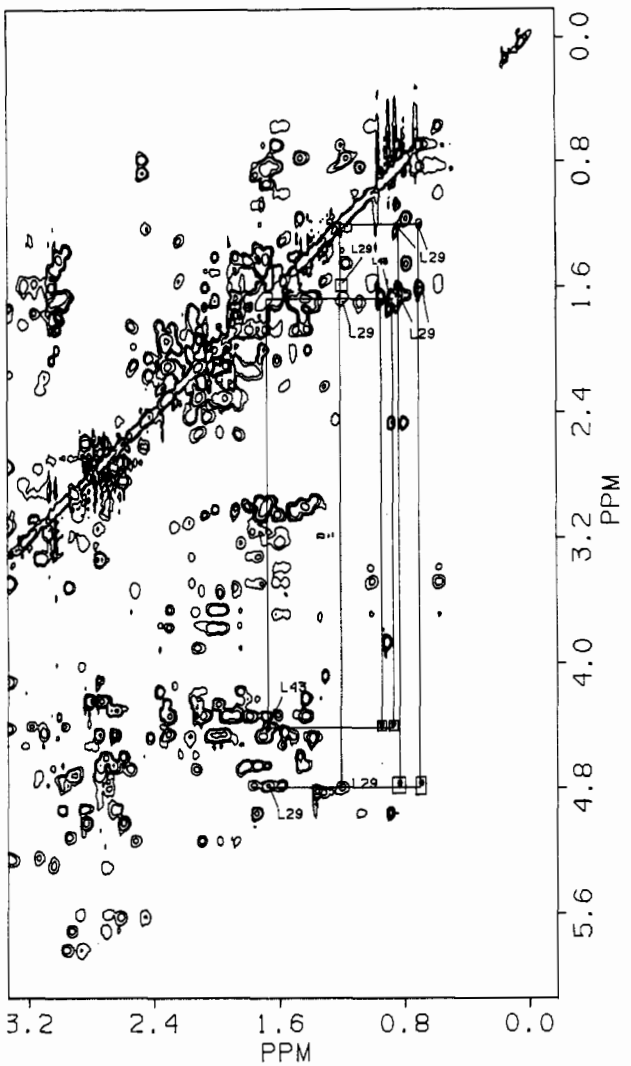

FIGURE 3: Assignment of L29 and L43 from the 50-ms HOHAHA spectrum, 35 °C, pH 5.7, in D_2O. Stars in boxes indicate NOEs observed at 100 ms in D_2O. Unstarred box indicates the $\beta'\gamma$ cross-peak of L29 identified in the DQF-COSY spectrum under the same conditions.

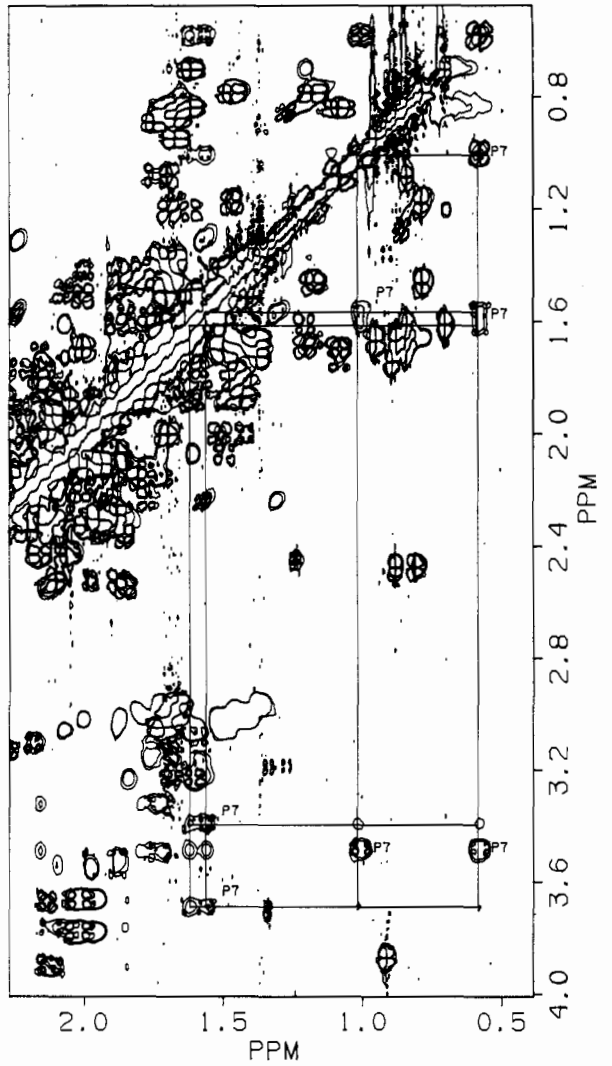

FIGURE 4: Overlay of DQF-COSY and 50-ms HOHAHA spectra showing the assignment of the P7 spin system at 35 °C pH 5.7, in D_2O.

their $\alpha\beta$ cross-peak multiplicity (AMX) in DQF-COSY spectra and the observation of corresponding $\alpha\delta$ and $\beta\delta$ cross-peaks in the 100-ms NOESY spectrum in D_2O. The corresponding NHα, NHβ and NHβ' cross-peaks observed in the H_2O NOESY spectrum completed the assignment for these residues. The assignment of the single histidine residue was completed in a similar manner. The NHα cross-peak of H33 was very weak in the DQF-COSY spectrum but was observed as a weak peak in the HOHAHA spectrum at 35 °C in H_2O. This assignment was confirmed by the observation of the corresponding NHβ and NHβ' cross-peaks in the NOESY spectrum at the same temperature.

The $\alpha\beta$ cross-peaks in the DQF-COSY spectrum having a β-proton chemical shift above 3.7 ppm (that were not previously assigned to threonine) were considered to be putative serine spin systems. The multiplicity permitted identification of these spin systems as AMX type. Their subsequent assignment as serine was established through sequence-specific connectivities. Data obtained at both 15 and 35 °C were used to resolve degeneracies with other α-proton chemical shifts.

Arginine and lysine spin systems were discriminated from all others primarily on the basis of relay cross-peaks $\alpha\gamma$, $\alpha\delta$, $\beta\delta$, and $\beta'\delta$ in HOHAHA spectra and $\alpha\gamma$ cross-peaks in the 100-ms NOESY spectrum in D_2O.

The assignment of the six proline residues is demonstrated by example with the assignment of P7 (Figure 4). DQF-COSY in D_2O permitted identification of the $\alpha\beta$, $\alpha\beta'$, and $\beta\beta'$ cross-peaks. The corresponding γ-, γ'-, δ-, and δ'-protons were then identified by relay cross-peaks in HOHAHA spectra in D_2O. For P7, $\beta\delta$, $\beta\delta'$, $\beta'\delta$, and $\beta'\delta'$ cross-peaks were all observed and corresponded with the $\gamma\delta$, $\gamma\delta'$, $\gamma'\delta$, and $\gamma'\delta'$ cross-peaks observed in the DQF-COSY spectrum. Finally, the relay cross-peaks $\alpha\gamma$ and $\alpha\gamma'$ observed in the HOHAHA spectrum in D_2O completed the spin system assignment. Prolines-19, -20, -46, -61, and -64 were all assigned in the same manner. For P61 and P64, the γ-protons were found to be essentially degenerate.

The spin system corresponding to C8 was only identified on the basis of the presence of its NHα cross-peak in the DQF-COSY spectrum and NHβ and NHβ' cross-peaks observed in 125-ms (35 °C) and 200-ms (45 °C) NOESY spectra in H_2O. Neither $\alpha\beta$ nor $\alpha\beta'$ cross-peaks were observed in DQF-COSY or HOHAHA spectra in D_2O, but putative $\alpha\beta$ and $\alpha\beta'$ NOEs were observed in D_2O, consistent with the apparent NHβ and NHβ' NOEs in H_2O. It was observed that the β- and β'-protons were nearly degenerate with those of H33, but these could be distinguished in the 200-ms NOESY spectrum at 45 °C and the 150-ms NOESY spectrum at 15 °C. In addition, a clear sequential NOE $d_{\alpha N(i,i+1)}$ between P7 and C8 was observed which confirmed that the observed NHα

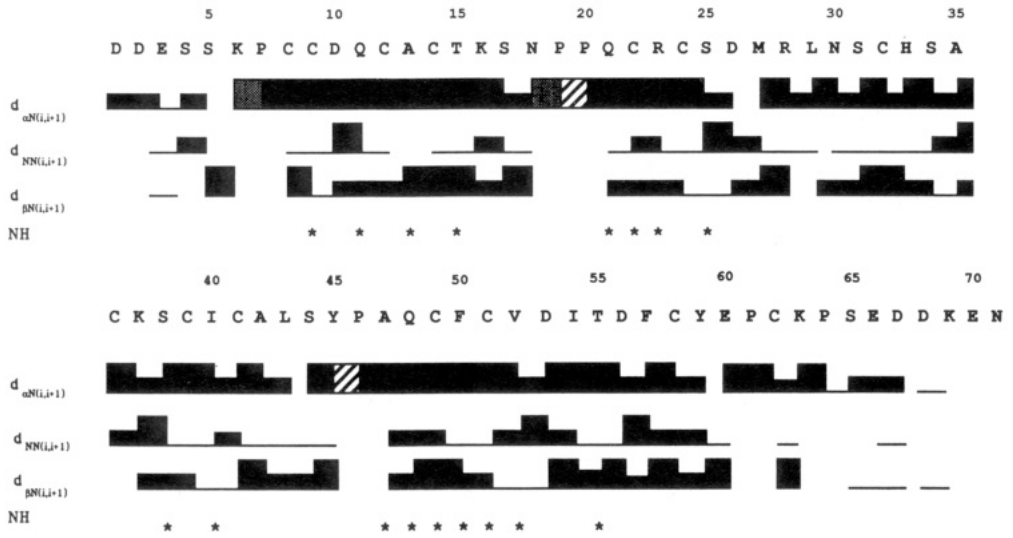

FIGURE 5: Sequential connectivities observed for BBI-I at 35 °C, pH 5.7, including geometry about proline. Shaded connectivities indicate *trans*-X–Pro bond; hatched connectivities indicate *cis*-X–Pro bond. Asterisks indicate slowly exchanging amide protons as discussed in the text. Thick bars indicate strong NOEs, medium thickness indicates medium intensity NOEs, and thin bars indicate weak NOEs.

cross-peak in the DQF-COSY spectrum belonged to a specific amino acid residue.

Sequence-Specific Assignments and Description of the Secondary Structure

Figure 5 summarizes the observed sequential connectivities for BBI-I. For the N- and C-terminal residues (D1–P7 and K63–N71, respectively), assignment was possible only at 15 °C as amide proton degeneracies precluded the observation of unambiguous $d_{\alpha N(i,i+1)}$ and $d_{NN(i,i+1)}$ connectivities for many of these residues at higher temperatures. No unambiguous sequential assignment could be made for E70 or N71.

Proline Geometry. The backbone geometry about the proline may be determined by the observation of sequential $d_{\alpha\delta(i,i+1)}$ and $d_{\alpha\delta'(i,i+1)}$ (trans) or $d_{\alpha\alpha(i,i+1)}$ (cis) connectivities in NOESY spectra at short mixing times for the X–Pro bond (Wüthrich et al., 1984). K6–P7, P19–P20, and K63–P64 were all shown to have trans geometry based on the observation of sequential $d_{\alpha\delta(i,i+1)}$ and $d_{\alpha\delta'(i,i+1)}$ connectivities in the NOSEY spectrum with a 100-ms mixing time in D_2O. N18 was observed to be connected by a $d_{\alpha\alpha}$ connectivity to P19 and by $d_{\alpha\delta}$ and $d_{\alpha\delta'}$ connectivities to P20. Subsequent analysis of the medium- and long-range contacts permitted assignment of the N18–P19 geometry to cis and the $d_{\alpha\delta}$ and $d_{\alpha\delta'}$ connectivities to P20 as $d_{(i,i+2)}$ connectivities of a type VI like tight turn (see below). The E60–P61 geometry has been assigned to be trans; however, the defining $d_{\alpha\delta(i,i+1)}$ connectivity of E60–P61 is believed to be overlapped with the $\alpha\beta$ and $\alpha\beta'$ cross-peaks of S5, and the $d_{\alpha\delta'(i,i+1)}$ cross-peak of E60–P61 is nearly degenerate with the $d_{\alpha\delta(i,i+1)}$ cross-peak of K6–P7. No $d_{\alpha\alpha(i,i+1)}$ cross-peak was observed in any of the NOESY spectra collected in D_2O which, if seen, would suggest a cis geometry about the E60–P61 bond. For all six proline residues, a sequential $d_{\alpha N(i,i+1)}$ connectivity was observed for the Pro–X bond which permitted assignment through the prolines in the primary sequence. As mentioned earlier, the observation of an unambiguous $d_{\alpha N(i,i+1)}$ NOE for the P7–C8 bond aided in confirming the assignment for the spin system of C8, even in the absence of the corresponding $\alpha\beta$ or $\alpha\beta'$ cross-peaks for this residue in DQF-COSY or HOHAHA spectra.

Elements of Secondary Structure. Although circular dichroism apparently shows no evidence of extensive regions of regular secondary structure in BBI-I (Birk et al., 1980), it is observed that two short regions of antiparallel β-sheet are present in BBI-I.

Residues Q11–T15 and residues Q21–S25 form a short region of antiparallel β-sheet in the tryptic inhibitory region of the molecule. This was determined on the basis of strong, nonsequential $d_{\alpha\alpha(i,j)}$ NOEs between residues C14 and C22 and between residues C12 and C24. Nonsequential $d_{NN(i,j)}$ connectivities were observed between residues Q11–S25 and T15–Q21; however, that between residues A13 and R23 was not observed because the amide protons of R23 and A13 were nearly degenerate at 15, 35, and 45 °C. Nonsequential NOEs $d_{\alpha N(i,j)}$ were observed between residues C14α–R23NH and C22α–T15NH and between residues C24α–A13NH and C12α–S25NH. Finally, slowly exchanging amide protons were observed for the residues Q11, A13, T15, Q21, R23, and S25, indicative of hydrogen bonding consistent with this region of β-sheet.

An analogous region of antiparallel β-sheet was also observed in the chymotryptic inhibitory domain between residues S38–A42 and residues Q48–V52. Again, nonsequential $d_{\alpha\alpha(i,j)}$ NOEs were observed between C41–C49 and C39–C51. $d_{\alpha N(i,j)}$ NOEs occur between C39α–V52NH and between residues C41α–F50NH and C49α–A42NH. $d_{NN(i,j)}$ NOEs occur between I40–F50 and A42–Q48. The amide protons of residues S38, I40, Q48, F50, and V52 exchange slowly, indicative of hydrogen bonding in this region of β-sheet. The pattern of NOEs which defined both regions of β-sheet is reviewed in Figure 6.

Two tight turns have been identified, one each near the inhibitory sites for trypsin and chymotrypsin. Strong NOEs $d_{\alpha N(i,i+3)}$ are seen between P20α and S17NH, with an additional contact $d_{\alpha\delta(i,i+2)}$ between N18α and P20δ. An analogous turn is observed in the chymotryptic inhibitory domain between residues L43 and A47. A $d_{\alpha N(i,i+3)}$ NOE occurs between A47α and S44NH, with an additional contact $d_{\alpha\alpha(i,i+2)}$ between Y45 and A47 (Figure 6). This pattern of NOEs is consistent with that expected from a type VI like turn (Richardson, 1981). The combination of a tight turn and antiparallel β-sheet form a β-hairpin in each of the inhibitory domains.

Disulfide Cross-Links. The seven disulfide bonds of BBI-I have previously been identified by chemical analysis by Odani and Ikenaka (1973). The identification of disulfide cross-links can be confirmed spectroscopically by observation of NOEs between amino acid side chain protons located distantly in the

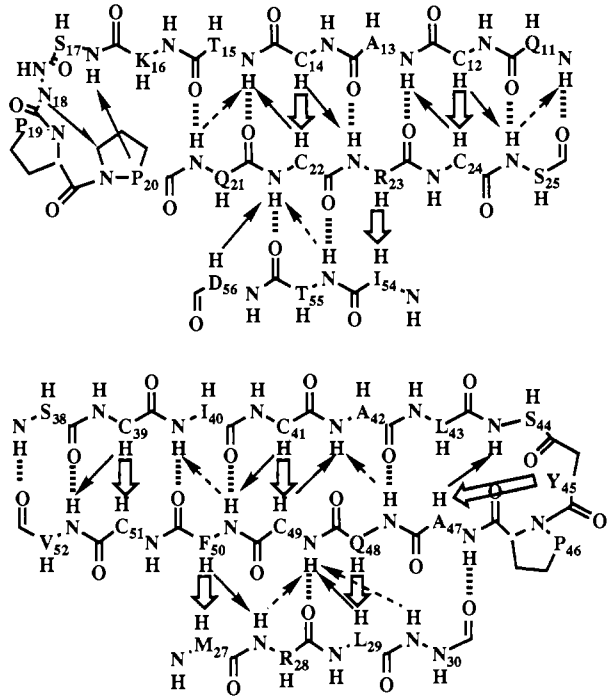

FIGURE 6: Secondary structure contacts and hydrogen bonding for BBI-I determined from NOESY spectra in H_2O at 35 °C, pH 5.7.

primary sequence. The disulfide cross-link C14–C22 was identified by strong, nonsequential NOEs 14α–22β and the reciprocal pairs 22α–14β and 22α–14β' at 100 ms in D_2O. Similarly, C41–C49 was identified from the pairs 41α–49β' and 49α–41β.

Much weaker $\alpha\beta$ NOEs were observed at 100 ms for the cross-links C12α–C58β and C32β–C39α. Cross-links C9–C24 and C36–C51 could not be identified on the basis of nonsequential $\alpha\beta$ cross-peaks in the NOESY spectrum. These were shown to exist only by chemical methods (Odani & Ikenaka, 1973). Finally, as mentioned earlier, the β-protons of C8 are nearly degenerate with those of H33. The cross-link C8–C62 was identified at 15 °C on the basis of a weak 62α–8β NOE where the β-proton degeneracy of C8 and H33 can be resolved.

Medium- and Long-Range Contacts. The overall fold of the molecule can begin to be addressed by the intermediate- and long-range contacts observed between distant residues in the primary sequence. Table II lists the medium- and long-range NOEs observed in BBI-I that have been unambiguously identified at this point in the spectral analysis.

Each inhibitory domain is found to be in contact with a short extended intervening sequence. The tryptic inhibitory domain is in contact in a sheetlike structure with residues I54–D56. A $d_{\alpha\alpha(i,j)}$ NOE is observed between residues R23 and I54. The amide proton of C22 is in contact with D56α and with the amide proton of T55; the amides of C22 and T55 are both slowly exchanging. This pattern of NOEs and the relative rates of amide exchange are suggestive of an extended or sheetlike structure; however, the NOEs are weaker than those which were previously identified in the β-hairpin of the inhibitory domain.

A similar pattern of NOEs is observed between the chymotryptic inhibitory domain and residues M27–N30. $d_{\alpha\alpha(i,j)}$ NOEs are observed between residues L29 and Q48 and residues M27 and F50. $d_{\alpha N(i,j)}$ NOEs are observed between F50α and R28NH and between L29α and C49NH. The amide proton of C49 is found to be in contact with the amide proton of R28 and of N30, and both A47 and C49 are slowly exchanging amide protons. Again, although this pattern suggests

Table II: Medium- and Long-Range NOEs[a]

C9α–Q11NH	N18α–P20δ,δ'
C9NH–E60NH	C24α–D26NH
C9NH–C62NH	C36α–S38NH
D10α–P61α	S38NH–D53NH
D10α–C62NH	I40NH–V52NH
Q11NH–D26NH	A42NH–F50NH
A13NH–S25NH	S44NH–A47α
T15NH–R23NH	Y45α–A47α
S17NH–P20α	

[a] pH 5.7, 35 °C, 125 ms, 15 °C, 150 ms, excluding sheet contacts.

an extended conformation between residues M27 and N30 in the chymotryptic inhibitory domain, the NOEs connecting them are weaker than those identified in the β-hairpin of the domain itself.

A number of turns in the primary sequence have been identified, primarily based on the observation of $d_{\alpha N(i,i+2)}$ NOEs listed in Table II. The inhibitory site at one end of each inhibitory domain appears to be part of a type VI like tight turn, as mentioned previously. In addition, the opposing end of each region of the β-sheet is connected to the rest of the molecule by half-turns to the adjoining intervening sequence.

Discussion

Effect of Cosolvent. The presence of 18% acetonitrile raises the question of whether a native or nearly native structure of BBI-I was preserved in the mixed solvent. The inhibitory activity of BBI-I is routinely measured in the presence of 5–10% methanol or acetonitrile to improve the solubility of specific ester substrates. Thus, it was believed that the presence of 18% acetonitrile would not substantially alter the assayed inhibitory activity. A more sensitive assessment of the effect of the cosolvent on the native structure is provided by comparison of 200-ms NOESY spectra in H_2O in the presence and absence of acetonitrile. Spectra were compared at pH 7.2, 35 °C, without acetonitrile to spectra collected at pH 5.7, 35 °C, in the presence of 18% acetonitrile-d_3. The choice of pH in the absence of acetonitrile was made to maximize solubility in the absence of cosolvent. Two things were apparent in this comparison. First, the pattern of intraresidue and sequential contacts which was observed at pH 7.2 and at pH 5.7 was the same, although a greater number of $d_{\beta N(i,i+1)}$ NOEs could be identified at pH 5.7. This is believed to be partly due to the suppression of amide exchange at the lower pH and partly due to the greater solubility of the BBI-I monomer in the cosolvent which permitted identification of the sequential contacts. The same observation was made for the $d_{\alpha N(i,i+1)}$ NOEs in comparing the two data sets.

Second, the pattern of medium-range NOEs defining the β-sheet was the same in the presence and absence of acetonitrile; however, again a larger number of these contacts were resolved at pH 5.7 in the cosolvent. In addition, the relative NOE intensities were comparable for these important medium-range contacts. Thus, the secondary structure contacts were preserved in the cosolvent, and their relative intensities were consistent with that expected for such contacts as they occur in β-sheet.

Finally, the patterns all other NOEs between the backbone NH and the side chains were comparable at both pH 7.2 and pH 5.7 indicating that the global fold of the molecule is probably very similar in the presence and absence of acetonitrile. The only differences between the spectra were that a larger number of contacts were observed in the cosolvent spectra which we believe is largely due to the greater solubility of the BBI-I monomer.

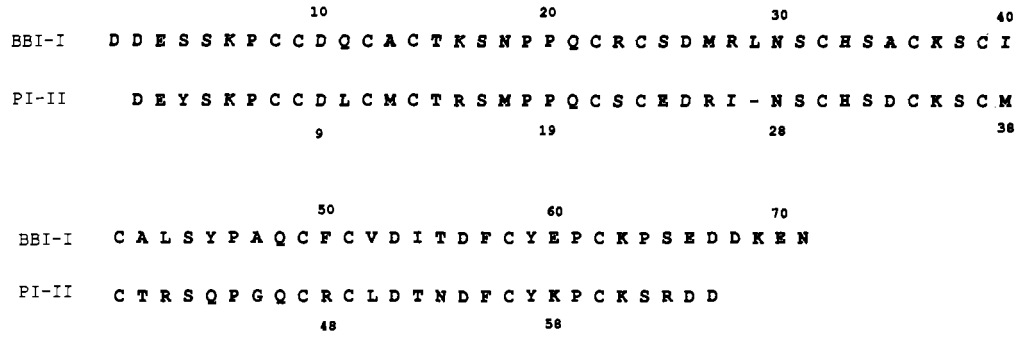

FIGURE 7: Comparison of the amino acid sequences of soybean BBI-I and soybean PI-II.

^1H Assignments. Of 71 residues in the primary sequence of BBI-I, 69 were assigned sequence-specifically. The nonexchangeable protons of residues E70 and N71 have been tentatively assigned, but it was not possible to connect the observed spin systems to their location in the primary sequence; the assignment of N71 shown in Figure 2 is based on the observation of an AMQX spin system in H_2O that was not assigned to any other amino acid residue in the primary sequence. The spin system of L43 could not be distinguished from that of valine, and its assignment as leucine was based solely on its position in the primary sequence. Comparison of the primary sequences of BBI-I with other members of the Bowman-Birk family (Morhy & Ventura, 1987) does not indicate the presence of valine at the chymotryptic reactive site (position 43), and N-terminal sequence analysis of our preparation of BBI-I through residue N30 did not indicate any differences between BBI-I isolated in our laboratory and the published sequence (Odani & Ikenaka, 1972). We therefore feel that the assignment to L43 is correct and that the β- and γ-protons are degenerate for this spin system.

Secondary Structure. Recently, a number of reports have appeared regarding the structure of trypsin/trypsin Bowman-Birk inhibitors determined in part or in whole by X-ray crystallography. Tsunogae et al. (1986) reported that the tryptic inhibitory domain of AB-I from azuki beans showed no evidence of β-sheet when bound to trypsin. The same observation was reported by Suzuki et al. (1987) for both trypsin inhibitory domains in the uncomplexed trypsin/trypsin inhibitor A-II isolated from peanuts. This is in clear contrast to the secondary structure of BBI-I from soybeans as determined here by NMR spectroscopy as well as to that of the trypsin/trypsin inhibitor PI-II crystallographic structure recently reported by Chen (1989). Perhaps a higher degree of refinement by Tsunogae and co-workers can help elucidate the differences between the BBI-type inhibitors they have studied and that reported here.

Tryptic Inhibitory Domain. Chen (1989) has recently reported the crystallographic determination of the trypsin/trypsin PI-II inhibitor from Tracy soybeans. This inhibitor shares 65% sequence homology with BBI-I (Figure 7) and has disulfide cross-links analogous to those of BBI-I. Thus, PI-II might be expected to have a structure similar to that of BBI-I. However, comparison of the crystallographic secondary structure determination to that of BBI-I determined by NMR spectroscopy shows some significant differences.

PI-II is reported to have a short region of antiparallel β-sheet in one tryptic inhibitory domain between residues M12–T14 and residues Q20–S22. This contrasts to that observed in BBI-I where the sheet extends from Q11 to T15 and from Q21 to S25. The observation of the NOE $d_{NN(i,j)}$ Q11–S25 and the slowly exchanging amide protons of Q11 and S25 includes these residues as part of the β-sheet in this domain. This is supported by observation of the NOE $d_{NN(i,j)}$ Q11–D26 as well as the disulfide cross-link C9–C24, which places Q11 across from S25 at one end of the sheet in the tryptic inhibitory domain of BBI-I.

Chymotryptic Inhibitory Domain. In PI-II, the second antitryptic domain contains a region of antiparallel β-sheet between residues S36–T40 and residues Q46–L50. This is analogous to that of BBI-I where the sheet occurs between residues S38–A42 and residues Q48–V52.

Intervening Region. In addition to distinct regions of β-sheet in each inhibitory domain, two short sheetlike regions have been identified in BBI-I linking each inhibitory domain to the rest of the molecule.

I54–D56 forms a sheetlike region with Q21–R23 in the tryptic inhibitory domain. In the chymotryptic inhibitory domain, residues A47–F50 are connected to residues M27–N30. In both these cases, the pattern of NOEs and the observation of slowly exchanging amide protons as discussed above are consistent with a β-sheet-like structure. However, qualitative evaluation of the NOE intensities suggests that these intervening sequence contacts are much weaker than those observed in the β-hairpin of each of the inhibitory domains. Thus, it is unclear whether BBI-I forms a continuous four-stranded β-sheet across both inhibitory domains as suggested by Chen for PI-II. A detailed evaluation of the tertiary structure should resolve these structural issues.

Disulfide cross-links between C8 and C62 and between C12 and C58 bring the two terminii into close contact. In addition to those contacts described above between the tryptic inhibitory domain and I54–D56, contacts exist between D10α and P61α and between C9NH and C62NH and between C9NH and E60NH; C9NH is slowly exchanging while C62NH is moderately slow exchanging. This pattern of relative exchange rates and NOEs suggests some type of structure exists in these residues, but it is not well-defined as yet.

No contacts were observed between D1–P7 and K63–N71, and few intraresidue contacts are observed in NOESY spectra in H_2O even at 8 °C for these terminal amino acids. It is unusual that the α- and β-chemical shifts of P7 are so far upfield from that commonly seen in proteins (Table I). Medium-range NOEs are observed between P7α and H33δ and H33ϵ, suggesting that the unusual chemical shifts in P7 may be due to a ring-current shift from H33. The lack of substantial NOE contacts between the N- or C-terminus and other regions distant in the primary sequence suggests that both terminii are irregularly structured. It will remain for the tertiary structure analysis in progress to better characterize the structure present in these regions of the protein.

The pattern of disulfide cross-links and the secondary structure described imply that the structure of BBI-I is fairly rigid. Each inhibitory domain is held in place by the disulfide cross-links within a domain and by cross-links of a domain to

an intervening sequence. C32–C39 and C36–C51 in the anti-chymotryptic domain link it to the intervening sequence residues S31–K37. A similar pattern of contacts and cross-links exists in the tryptic inhibitory domain. Further, the disulfide cross-link within each domain's β-sheet (e.g., C41–C49 in the anti-chymotryptic domain) maintains a rigid, extended combining site for the protease; this is consistent with the observation of such cross-links in other serine protease inhibitors (Laskowski & Kato, 1980). Overall, as implied in Figure 6, the preliminary evaluation of the structure of BBI-I indicates that the inhibitory sites are extended on opposite ends of the molecule. This would be consistent with the observation that the inhibition of protease at one site is kinetically independent from that of the second site.

In conclusion, we have described the ^{1}H resonance assignments and secondary structure determination of the trypsin/chymotrypsin Bowman–Birk inhibitor from soybeans using ^{1}H NMR. Evaluation of the tertiary structure using NOE intensities as an input to a distance geometry algorithm and studies of the protein binary and ternary complexes shall help elucidate the interesting inhibitory properties of the Bowman–Birk family of protease inhibitors.

References

Bax, A., & Davis, D. G. (1985) *J. Magn. Reson. 65*, 355–360.

Birk, Y. (1985) *Int. J. Pept. Protein Res. 25*, 113–131.

Birk, Y., Jibson, M. D., & Bewley, T. A. (1980) *Int. J. Pept. Protein Res. 15*, 193–199.

Bodenhausen, G., Vold, R. L., & Vold, R. R. (1980) *J. Magn. Reson. 37*, 93–106.

Braunschweiler, L., & Ernst, R. R. (1983) *J. Magn. Reson. 53*, 521–528.

Chen, P. (1989) Ph.D. Thesis, University of Pittsburgh.

Clore, G. M., & Gronenborn, A. M. (1987) *Protein Eng. 1*, 275–288.

Driscoll, P. C., Clore, G. M., Beress, L., & Gronenborn, A. M. (1989) *Biochemistry 28*, 2178–2187.

Drobny, G., Pines, A., Sinton, S., Weitekamp, D., & Wemmer, D. E. (1978) *Faraday Symp. Chem. Soc. 13*, 49–55.

Frattali, V. (1969) *J. Biol. Chem. 244*, 274–280.

Frattali, V., & Steiner, R. F. (1968) *Biochemistry 7*, 521–530.

Gaier, J. R., Tulinsky, A., & Liener, I. E. (1981) *J. Biol. Chem. 256*, 11417–11419.

Gross, K.-H., & Kalbitzer, R. (1988) *J. Magn. Reson. 76*, 87–99.

Harry, J. B., & Steiner, R. F. (1969) *Biochemistry 8*, 5060–5064.

Harry, J. B., & Steiner, R. F. (1970) *Eur. J. Biochem. 16*, 174–179.

Jeener, J., Meier, B. H., Bachmann, P., & Ernst, R. R. (1979) *J. Chem. Phys. 71*, 4546–4553.

Laemmli, U. K. (1970) *Nature 227*, 680–685.

Laskowski, M., Jr., & Kato, I. (1980) *Annu. Rev. Biochem. 49*, 593–626.

Macura, S., & Ernst, R. R. (1980) *Mol. Phys. 41*, 95–117.

Marion, D., & Wüthrich, K. (1983) *Biochem. Biophys, Res. Commun. 113*, 967–974.

Millar, D. B. S., Willick, G. E., Steiner, R. F., & Frattali, V. (1969) *J. Biol. Chem. 244*, 281–284.

Morhy, L., & Ventura, M. M. (1987) *An. Acad. Brasil. Cien 59*, 71–81.

Odani, S., & Ikenaka, T. (1972) *J. Biochem. (Tokyo) 71*, 839–848.

Odani, S., & Ikenaka, T. (1973) *J. Biochem. (Tokyo) 74*, 697–715.

Odani, S., & Ikenaka, T. (1977) *J. Biochem. (Tokyo) 82*, 1513–1522.

Plateau, P., & Guêron, M. (1982) *J. Am. Chem. Soc. 104*, 7310–7311.

Rance, M., Sorensen, G., Bodenhausen, G., Wagner, G., Ernst, R. R., & Wüthrich, K. (1983) *Biochem. Biophys. Res. Commun. 117*, 479–485.

Redfield, A. G., & Kunz, S. D. (1975) *J. Magn. Reson. 19*, 250–254.

Richardson, J. S. (1981) *Adv. Protein Chem. 34*, 167–339.

Steiner, R. F., & Frattali, V. (1969) *J. Agric. Food Chem. 17*, 513–518.

Suzuki, A., Tsunogae, Y., Tanaka, I., Yamane, T., Ashida, T., Norioka, S., Hara, S., & Ikenaka, T. (1987) *J. Biochem. (Tokyo) 101*, 267–274.

Torchia, D. A., Sparks, S. W., & Bax, A. (1989) *Biochemistry 28*, 5509–5524.

Tsunogae, Y., Tanaka, I., Yamane, T., Kikkawa, J., Ashida, T., Ishikawa, C., Watanabe, K., Nakamura, S., & Takahashi, K. (1986) *J. Biochem. (Tokyo) 100*, 1637–1646.

Turner, R., Liener, I. E., & Lovrien, R. E. (1975) *Biochemistry, 14*, 275–282.

Wagner, G., & Wüthrich, K. (1982) *J. Magn. Reson. 155*, 347–366.

Walsh, K. A. (1970) *Methods Enzymol. 19*, 41–63.

Wei, C. H., Sankar, P. B., & Einstein, J. R. (1979) *J. Biol. Chem. 254*, 4892–4894.

Wilcox, P. E. (1970) *Methods Enzymol. 19*, 64–108.

Wüthrich, K. (1986) *NMR of Proteins and Nucleic Acids*, Wiley, New York.

Wüthrich, K., Billeter, M., & Braun, W. (1984) *J. Mol. Biol. 180*, 715–740.