Spiess, M., Schwartz, A. L., & Lodish, H. F. (1985) *J. Biol. Chem. 260*, 1979–1982.
Steer, C. J., Kempner, E. S., & Ashwell, G. (1981) *J. Biol. Chem. 256*, 5851–5856.
Stoorvogel, W., Geuze, H. J., Griffith, J. M., Schwartz, A. L., & Strous, G. J. (1989) *J. Cell Biol. 108*, 2137–2148.
Südhof, T. C., Goldstein, J. L., Brown, M. S., & Russell, D. W. (1985) *Science 228*, 815–822.
Suter, U., Bastos, R., & Hofstetter, H. (1987) *Nucleic Acids Res. 15*, 7295–7308.
Takahashi, T., Nakada, H., Okumura, T., Sawamura, T., & Tashiro, Y. (1985) *Biochem. Biophys. Res. Commun. 126*, 1054–1060.
Thiel, S., & Reid, K. B. M. (1989) *FEBS Lett. 250*, 78–84.
Tolleshaug, H., & Berg, T. (1979) *Biochem. Pharmacol. 28*, 2919–2922.
Tycko, B., Keith, C. H., & Maxfield, F. R. (1983) *J. Cell Biol. 255*, 5971–5978.
Van Lenten, L., & Ashwell, G. (1972) *J. Biol. Chem. 247*, 4633–4638.
Vercelli, D., Helm, B., Marsh, P., Padlan, E., Geha, R. S., & Gould, H. (1989) *Nature 338*, 649–650.
von Heijne, G. (1986) *EMBO J. 5*, 3021–3027.
von Heijne, G., & Gavel, Y. (1988) *Eur. J. Biochem. 174*, 671–678.
Wall, D. A., & Hubbard, A. (1981) *J. Cell Biol. 90*, 687–696.
Walter, P., Gilmore, R., & Blobel, G. (1984) *Cell 38*, 5–8.
Watts, C. (1985) *J. Cell Biol. 100*, 633–637.
Weigel, P. H. (1980) *J. Biol. Chem. 255*, 6111–6120.
Wessels, H. P., Geffen, I., & Spiess, M. (1989) *J. Biol. Chem. 264*, 17–20.
White, R. T., Damm, D., Miller, J., Spratt, K., Schilling, J., Hawgood, S., Benson, B., & Cordell, B. (1985) *Nature 317*, 361–363.
Wickner, W. T., & Lodish, H. F. (1985) *Science 230*, 400–407.
Yarden, Y., & Ullrich, A. (1988) *Biochemistry 27*, 3113–3119.
Zerial, M., Melancon, P., Schneider, C., & Garoff, H. (1986) *EMBO J. 5*, 1545–1550.

## *Accelerated Publications*

# X-ray Crystal Structure of the Protease Inhibitor Domain of Alzheimer's Amyloid β-Protein Precursor[†,‡]

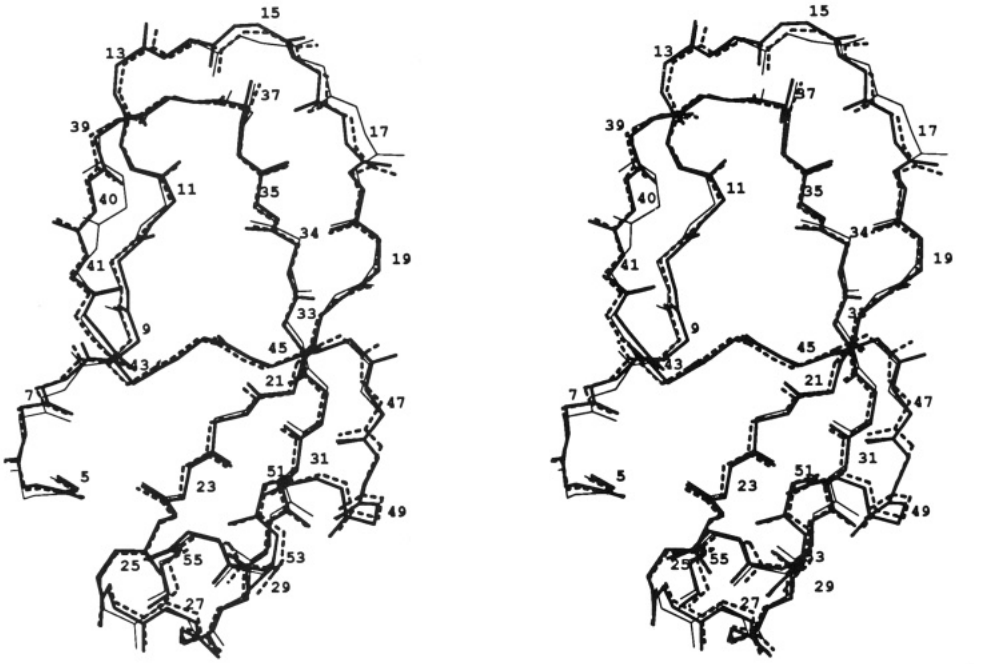
Thomas R. Hynes,* Michael Randal, Laura A. Kennedy, Charles Eigenbrot, and Anthony A. Kossiakoff
*Protein Engineering Department, Genentech Inc., 460 Point San Bruno Boulevard, South San Francisco, California 94080, and Department of Pharmaceutical Chemistry, University of California, San Francisco, San Francisco, California 94143*
*Received July 31, 1990; Revised Manuscript Received September 4, 1990*

ABSTRACT: Alzheimer's amyloid β-protein precursor contains a Kunitz protease inhibitor domain (APPI) potentially involved in proteolytic events leading to cerebral amyloid deposition. To facilitate the identification of the physiological target of the inhibitor, the crystal structure of APPI has been determined and refined to 1.5-Å resolution. Sequences in the inhibitor–protease interface of the correct protease target will reflect the molecular details of the APPI structure. While the overall tertiary fold of APPI is very similar to that of the Kunitz inhibitor BPTI, a significant rearrangement occurs in the backbone conformation of one of the two protease binding loops. A number of Kunitz inhibitors have similar loop sequences, indicating the structural alteration is conserved and potentially an important determinant of inhibitor specificity. In a separate region of the protease binding loops, APPI side chains Met-17 and Phe-34 create an exposed hydrophobic surface in place of Arg-17 and Val-34 in BPTI. The restriction this change places on protease target sequences is seen when the structure of APPI is superimposed on BPTI complexed to serine proteases, where the hydrophobic surface of APPI faces a complementary group of nonpolar side chains on kallikrein A versus polar side chains on trypsin.

The major component of cerebral amyloid deposits in Alzheimer's disease is amyloid β-protein (AP), a proteolytically derived peptide of amyloid β-protein precursor (APP), an integral membrane glycoprotein (Glenner & Wong, 1984). Considerable work has focused on the nature of the proteolytic events leading to the accumulation of AP (Sisodia et al., 1990) following the discovery that alternative splicing of the APP message yields three products, two containing a Kunitz protease inhibitor domain (Kitaguchi et al., 1988; Ponte et al., 1988; Tanzi et al., 1988). Recent studies indicate that the inhibitor-containing species is overproduced in Alzheimer's patients (Johnson et al., 1989; Tanaka et al., 1989). It has been suggested that APPI may block the normal degradation of APP, leading to the deposition of cerebral amyloid (Müller-Hill & Beyreuther, 1989; Selkoe, 1990). The site of normal proteolytic cleavage occurs at a Lys residue within the AP peptide sequence (Esch et al., 1990). The primary specificity determining residue in APPI is an Arg, indicating that APPI blocks proteases that cleave at basic residues. The inhibition spectrum of APPI toward a range of serine proteases has been shown to be broad but distinct from basic pancreatic trypsin inhibitor (BPTI) (Kitaguchi et al., 1990).

[†] This work was supported by NIH Grant GM33571.
[‡] The coordinates of APPI have been deposited in the Brookhaven Protein Data Bank (entry name 1AAP).

0006-2960/90/0429-10018$02.50/0 



```
                5   10    15    20   25   30    35   40   45   50   55
APPI   VREVCSEQAE TGPCRAMIS RWYFDVTEGKCAPF FYGGCG GNRNNFDTEEYCMAVCGSA
BPTI   RPDFCLEPPY TGPCKARII RYFYNAKAGLCQTF VYGGCR AKRNNFKSAEDCMRTCGGA
```

FIGURE 1: Superposition of the backbone atoms of the two independent APPI molecules (thick lines, mol 1; dashed lines, mol 2) with BPTI (thin lines; 4PTI, Brookhaven Protein Data Bank) (Wlodawer et al., 1987; Bernstein et al., 1977). The alignment matched backbone atoms of residues 5–38 and 42–55 (Table I). The backbone conformation of APPI residues 39–41 differs from that of BPTI. The deformation of the backbone of APPI residues 15–17 relative to that of BPTI may be due in large part to crystal contacts in this region. The sequences of APPI and BPTI are shown below. Boxed residues comprise the two binding loops which lie in the inhibitor–protease interface.

Although Kunitz inhibitors are numerous and widespread in a variety of tissues, APPI is only the second naturally occurring member of the family to be examined in structural detail by X-ray diffraction and complements the vast body of knowledge that exists for BPTI. Sequences of 30 Kunitz inhibitors are known, providing an extensive data base of sequence variation. In combination with serine protease sequences it is clear that the wide spectrum of affinity and specificity of Kunitz inhibitors results from sequence variation on both sides of the inhibitor–protease interface (Creighton & Darby, 1989). Comparison of the structure of APPI with BPTI permitted the detailed examination of the effects of side-chain alterations on the structure and chemistry of the inhibitor surface and the extent to which the tertiary structure in conserved. By using the BPTI–trypsin and –kallikrein complex structures as a template and superimposing the APPI model on BPTI, we have evaluated the significance of the differences on the inhibitor side of the interface. The nature of protease targets which would complement the APPI structure could then be evaluated in the context of the inhibitor–protease complex.

## Experimental Procedures

*Protein Purification and Crystallization.* APPI protein was expressed in *Escherichia coli* from a plasmid carrying a synthetic gene encoding the APPI sequence (Castro et al., 1990). Protein was purified on a trypsin affinity column followed by anion exchange on MonoQ (Pharmacia). The protein was desalted, lyophilized, and dissolved in crystallization buffer (10 mM sodium acetate, pH 6.5). Crystals suitable for high-resolution X-ray data collection were grown from a 9 mg/mL protein solution by vapor diffusion at 20 °C against buffer containing 25% saturated magnesium sulfate.

*X-ray Data Collection.* The space group of the APPI crystals is $P2_12_12_1$ ($a$ = 35.6 Å, $b$ = 38.9 Å, and $c$ = 73.6 Å) with two molecules in the asymmetric unit. Data to 1.5-Å resolution were collected from three APPI crystals (~0.25 × 0.25 × 0.40 mm) on a Rigaku AFC6R diffractometer with monochromatic Cu K$\alpha$ X-rays (180 mA, 50 kV). The merging R factor on intensity was 9.6% to 1.6 Å with a 1$\sigma$ cutoff.

*Structure Refinement.* Starting phases were obtained by the molecular replacement method with a 331-atom model derived from the structure of BPTI (4PTI) (Wlodawer et al., 1987), eliminating side chains that differed significantly between APPI and BPTI. A single strong rotation solution was obtained with the Crowther fast-rotation function as implemented in the MERLOT program package (Fitzgerald, 1988). A strong cross peak in the native APPI Patterson suggested that the two independent APPI molecules had similar orientation and were related by a half unit cell translation along the $a$ and $b$ axes. By use of a model composed of two 331-atom structures, with orientations corresponding to the rotation solution and related by half unit cell translations along the $a$ and $b$ axes, the correct translation solution was found with the program BRUTE (Fujinaga & Read, 1987). The starting model had an $R$ factor of 38% over the resolution range of 8–3 Å. The APPI structure was refined by Hendrickson–Konnert stereochemically restrained least-squares refinement (Hendrickson, 1985) and model building with FRODO (Jones, 1985). The final $R$ factor over the resolution range of 8–1.5 Å is 17.7% with a 2$\sigma$ cutoff on measured amplitudes (11 908 reflections). The structure includes 105 water molecules. The rms deviations in bond distances and bond angles are 0.018 Å and 0.033 Å, respectively.

## Results and Discussion

The sequence of APPI has 43% sequence identity with that of BPTI (Figure 1). Crystals of APPI grew in space group $P2_12_12_1$ with two molecules in the asymmetric unit. In the

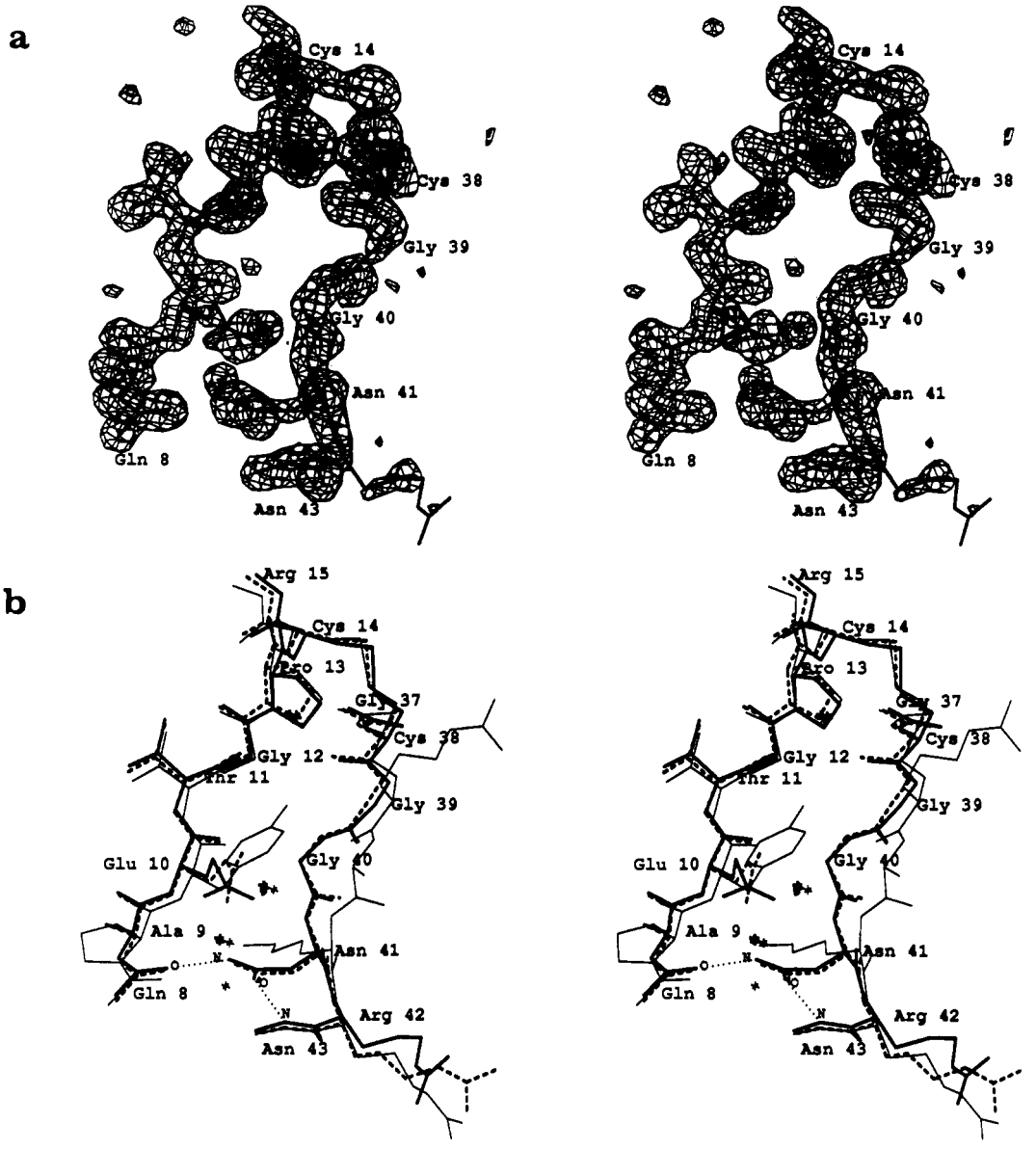



FIGURE 2: Structure of the protease binding loops of APPI in the region surrounding residues 39–41. (a) Residue delete $F_o$–$F_c$ electron density map contoured at $3\sigma$ ($\sigma$ is the rms electron density in the unit cell). The atoms shown, covering residues 8–15 and 37–43 and three buried water molecules, were not included in map calculations to eliminate bias from the model. Maps covering residues 8–15 and 37–43 were calculated separately and combined. Weak electron density is observed for the side chains of Glu-10 and Arg-42. (b) Superposition of the two APPI molecules (thick lines, mol 1; dashed lines, mol 2) with BPTI (thin lines, 4PTI) (Wlodawer et al., 1987). Hydrogen bonds between the Asn-41 side chain and the backbone oxygen of Gln-8 (2.65 Å) and the backbone nitrogen of Asn-43 (2.97 Å) are shown with dotted lines. The water molecule in the structure of BPTI which is displaced by the Asn-41 side chain is also shown. The structures were matched as described in Figure 1.

course of phase determination by molecular replacement, the crystal packing was found to resemble a C-centered cell ($C222_1$) having one molecule in the asymmetric unit with 2-fold symmetry broken by a 9° rotation between the two independent APPI molecules. Therefore, the packing environments of the two molecules are similar. The refined structures of the two APPI molecules (Figure 1) match closely with an rms deviation between backbone atoms of 0.38 Å (Table I).

With the exception of residues 39–41 (discussed below), the backbone conformation of APPI is almost identical with that of BPTI (Figure 1, Table I). Sequence differences in the hydrophobic core are conservative with the most significant change at residue 21, where there is a Trp in APPI and a Tyr in BPTI. Residue 21 fills a hydrophobic groove in the protein surface. The increased size of the Trp side chain in APPI is accommodated by greater extrusion into solvent space. The conformation of side chains that serve as structural determinants, including the three disulfide bridges, and residues in the hydrophobic core are conserved. The positions of three internal waters that are part of a structurally important hy-

Table I: Root Mean Square Deviations between APPI and BPTI (Å)[a]

| | APPI mol 1 | APPI mol 2 | BPTI (4PTI) | BPTI (2PTC) |
|---|---|---|---|---|
| APPI mol 1 | 0 | 0.38 | 0.41 | 0.38 |
| APPI mol 2 | 0.38 | 0 | 0.45 | 0.46 |
| BPTI (4PTI) | 0.58 | 0.60 | 0 | 0.36 |
| BPTI (2PTC) | 0.60 | 0.65 | 0.37 | 0 |

[a] Deviations between the backbone atoms (N, Cα, C, O) of the two independent molecules in the crystal structure of APPI (mol 1 and mol 2) and structures of BPTI alone (4PTI) (Wlodawer et al., 1987) and bound to trypsin (2PTC) (Rühlmann et al., 1973). Values below the diagonal correspond to the superposition of residues 5–55; values above the diagonal exclude residues 39–41, which differ significantly between APPI and BPTI.

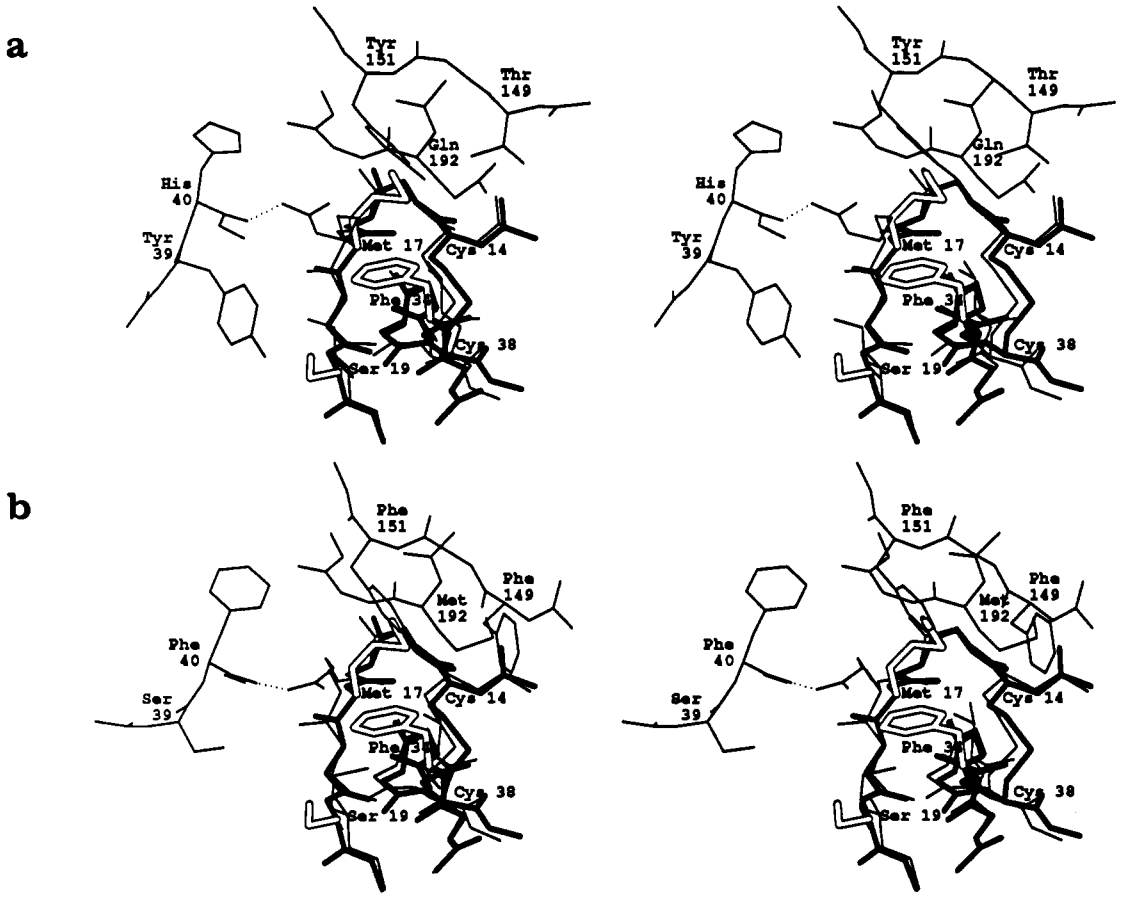



FIGURE 3: Changes in the inhibitor–protease interface of APPI involving residues 17, 34, and 19 in the context of trypsin and kallikrein A. Labeled and shown with open lines are the three APPI side chains (Met-17, Phe-34, and Ser-19) which differ from BPTI (Arg-17, Val-34, and Ile-19). (a) APPI (thick lines) superimposed on BPTI in the BPTI–trypsin complex structure (thin lines, 2PTC) (Rühlmann et al., 1973). Backbone and side-chain atoms of trypsin on the protease side of the interface are labeled and shown in thin lines (Tyr-39, His-40, Thr-149, Tyr-151, and Gln-192). (b) APPI (thick lines) superimposed on BPTI in the BPTI–kallikrein A complex structure (thin lines, 2KAI) (Chen & Bode, 1983). Backbone and side-chain atoms of kallikrein A on the protease side of the interface are labeled and shown in thin lines (Ser-39, Phe-40, Phe-149, Phe-151, and Met-192). In both figures the hydrogen bond between the side chain of Arg-17 in BPTI and the backbone oxygen of protease residue 40 is shown with a dotted line.

drogen-bonding network are identical in APPI and BPTI (Figure 2b).

An important physiological property of Kunitz inhibitors is their exceedingly strong binding to proteases (dissociation constant of BPTI–trypsin $6 \times 10^{-14}$ M) (Vincent & Lazdunski, 1972). Two binding loops (residues 11–19 and 34–39) define the specificity and form the interface between these inhibitors and their protease targets (Rühlmann et al., 1973). There are five hydrogen bonds in the interface between backbone atoms of the inhibitor and protease in the BPTI–trypsin and –kallikrein A complex structures (Rühlmann et al., 1973; Chen & Bode, 1983). The backbone atoms involved are unchanged in APPI so that altered target specificity presumably results from the side-chain substitutions in the interface. Six of the 15 interface residues show extensive sequence variation among 30 Kunitz inhibitor sequences (Creighton & Charles, 1987). Five of the six variable positions differ between APPI and BPTI (Figure 1). The changes occur in three regions of the inhibitor–protease interface. The center of the first loop mimics the bound peptide substrate and defines the primary specificity of inhibitors toward classes of proteases. Residues on each end of the protease binding loops interact with protease atoms around the perimeter of the catalytic site and confer additional specificity within a protease class (Creighton & Charles, 1987).

Residue 15 is located at the center of the first protease binding loop and is the principal determinant of Kunitz inhibitor specificity. In the inhibitor–protease complex, the side chain of residue 15 fills the P1 position preceding the scissile peptide bond. Residue 15 is an Arg in APPI and a Lys in BPTI. The substitution of Arg for Lys is conservative and common among inhibitors specific for the protease class which cleaves at basic residues (Creighton & Charles, 1987). On the basis of structural information from other systems, the guanidinium group of the Arg can make a direct salt bridge interaction with Asp-189 of the protease while the Lys interaction in BPTI is mediated by a water molecule (Bode et al., 1984).

Table II: Backbone $\phi$ and $\psi$ Angles of Residues 39–41

| | residue | | | | | |
|---|---|---|---|---|---|---|
| | 39 | | 40 | | 41 | |
| | $\phi$ | $\psi$ | $\phi$ | $\psi$ | $\phi$ | $\psi$ |
| APPI mol 1 | 62 | −146 | 84 | −152 | −140 | −176 |
| APPI mol 2 | 74 | −141 | 74 | −168 | −135 | −166 |
| BPTI (4PTI) | 62 | 36 | −56 | 156 | −104 | 172 |
| BPTI (2PTC) | 70 | 35 | −50 | 142 | −96 | 176 |

At the end of the second protease binding loop there is a significant alteration in the backbone fold (residues 39–41), which may represent a common feature among Kunitz inhibitors (Figure 2). A Gly-Gly-Asn sequence is found in APPI, while BPTI has an Arg-Ala-Lys sequence. Along with the shift in backbone position, Gly-39 and Gly-40 adopt backbone dihedral angles that are sterically unfavorable for non-Gly residues (Table II). The side chain of Asn-41 occupies the position of a well-ordered water molecule found in all BPTI

structures and forms hydrogen bonds to the backbone oxygen of Gln-8 and the backbone nitrogen of Asn-43, as well as one of the buried water molecules (Figure 2b). The peptide unit between residues 39 and 40 has flipped, exchanging the position of the hydrogen-bond donor and acceptor atoms. Protease residues that interact with the backbone atoms of this region should discriminate between the alternate conformations seen in APPI and BPTI.

Of 30 Kunitz inhibitor sequences, 26 have a Gly-Asn sequence for residues 40 and 41 (Creighton & Charles, 1987). Nine of the 26 sequences have Gly at residue 39 with a variety of side chains in the remaining 17. The nine inhibitors with the Gly-Gly-Asn sequence for residues 39–41 presumably adopt the backbone conformation seen in APPI. Although it is difficult to predict the backbone conformation for the 17 inhibitor sequences with non-Gly residues at position 39, the conservation of Gly-40 and Asn-41 argues for the adoption of the APPI backbone conformation. In the BPTI–trypsin and –kallikrein A structures the side chain of Arg-39 hydrogen bonds to protease atoms, indicating that the side-chain variation at this position contributes to target specificity. It is likely that Gly-40 and Asn-41 influence target specificity by altering the geometry of residue 39. We plan to address this question experimentally through structural analysis of site-directed mutants.

At the other end of the protease binding loops of APPI, Met-17 and Phe-34 create a large hydrophobic patch on the inhibitor surface in place of Arg-17 and Val-34 in BPTI. The significance of the change is illustrated by superimposing the structure of APPI onto BPTI complexed to trypsin and kallikrein A (Figure 3). For BPTI the primary contact is the side chain of Arg-17, which hydrogen bonds to the backbone oxygen of protease residue 40. The four kallikrein A side chains in this region of the interface (Phe-40, Phe-149, Phe-151, and Met-192) (Figure 3b) form a more compatible site for the hydrophobic patch on APPI than the polar residues of trypsin (His-40, Thr-149, Tyr-151, and Gln-192) (Figure 3a). This is consistent with binding data showing that, relative to BPTI, the $K_i$ of APPI is 4-fold lower for kallikrein A and 8-fold higher for trypsin (Kitaguchi et al., 1990). The final amino acid difference in the protease binding loops occurs at residue 19 where there is a Ser in APPI and a Ile in BPTI. In the BPTI–trypsin complex, the side chain of Tyr-39 from trypsin interacts with Ile-19, hydrogen bonding to the backbone nitrogen and packing against side-chain atoms. Kallikrein A does not form a contact with residue 19 of BPTI.

Given the conserved tertiary fold of Kunitz inhibitors and their serine protease targets, it is possible to evaluate candidate target sequences, as protease residues which face the bound inhibitor are known. Arg-15 is the primary determinant of specificity, suggesting that the in vivo target for APPI is a protease which cleaves at basic residues. APPI inhibits trypsin with a $K_i$ of $10^{-10}$ M (Kitaguchi et al., 1990). However APPI also inhibits chymotrypsin, which cleaves specifically at hydrophobic residues, in the nanomolar range of concentration (Kitaguchi et al., 1990). This points to the important contribution of residues along the margins of the active site to target affinity and specificity. The hydrophobic surface of APPI created by the Met-17 and Phe-34 indicates that protease targets will have nonpolar residues in the protease loops which face this region of the inhibitor. The altered backbone conformation of residues 39–41 could participate in specific hydrogen bonding with protease atoms, and the lack of a side chain at residue 39 could accommodate a larger protease loop or side chain in this region of the interface. In a general context, the change in backbone geometry demonstrates that the mutation of surface residues can cause substantial rearrangement of the backbone framework that underlies surface structure. An understanding of the impact of sequence variation on protease inhibitor structure will aid in the search for physiological targets and provides a model system for the study of protein–protein recognition at the atomic level.

## Acknowledgments

We thank Dr. Steven Anderson for providing the APPI expression system, Michael Covarrubias and John Altman for assistance with fermentation, and Dr. Bart de Vos for assistance with crystallographic programs.

## References

Bernstein, F. C., Koetzle, T. F., Williams, G. J. B., Meyer, E. F., Jr., Brice, M. D., Rodgers, J. R., Jennard, O., Shimanouchi, T., & Tasumi, M. (1977) *J. Mol. Biol. 112*, 535–542.

Bode, W., Walter, J., Huber, R., Wenzel, H. R., & Tschesche, H. (1984) *Eur. J. Biochem. 144*, 185–190.

Castro, M., Marks, C. B., Nilsson, B., & Anderson, S. (1990) *FEBS Lett. 267*, 207–212.

Chen, Z., & Bode, W. (1983) *J. Mol. Biol. 164*, 283–311.

Creighton, T. E., & Charles, I. G. (1987) *Cold Spring Harbor Symp. Quant. Biol. 52*, 511–519.

Creighton, T. E., & Darby, N. J. (1989) *Trends Biochem. Sci. 14*, 319–324.

Esch, F. S., Keim, P. S., Beattie, E. C., Blacher, R. W., Culwell, A. R., Oltersdorf, T., McClure, D., & Ward, P. J. (1990) *Science 248*, 1122–1124.

Fitzgerald, P. M. D. (1988) *J. Appl. Crystallogr. 21*, 273–278.

Fujinaga, M., & Read, R. J. (1987) *J. Appl. Crystallogr. 20*, 517–521.

Glenner, G. G., & Wong, C. W. (1984) *Biochem. Biophys. Res. Commun. 122*, 1131–1135.

Hendrickson, W. A. (1985) *Methods Enzymol. 115*, 252–270.

Johnson, S. A., Rogers, J., & Finch, C. E. (1989) *Neurobiol. Aging 10*, 267–272.

Jones, A. T. (1985) *Methods Enzymol. 115*, 157–171.

Kitaguchi, N., Takahashi, Y., Tokushima, Y., Shiojiri, S., & Ito, H. (1988) *Nature 331*, 530–532.

Kitaguchi, N., Takahashi, Y., Oishi, K., Shiojiri, S., Tokushima, Y., Utsunomiya, T., & Ito, H. (1990) *Biochim. Biophys. Acta 1038*, 105–113.

Müller-Hill, B., & Beyreuther, K. (1989) *Annu. Rev. Biochem. 58*, 287–307.

Ponte, P., Gonzalez-DeWhitt, P., Schilling, J., Miller, J., Hsu, D., Greenberg, B., Davis, K., Wallace, W., Lieberburg, I., Fuller, F., & Cordell, B. (1988) *Nature 331*, 525–527.

Rühlmann, A., Kukla, D., Schwager, P., Bartels, K., & Huber, R. (1973) *J. Mol. Biol. 77*, 417–436.

Selkoe, D. J. (1990) *Science 248*, 1058–1060.

Sisodia, S. S., Koo, E. H., Beyreuther, K., Unterbeck, A., & Price, D. L. (1990) *Science 248*, 492–495.

Tanaka, S., Nakamura, S., Ueda, K., Kameyama, M., Shiojiri, S., Takahashi, Y., Kitaguchi, N., & Ito, H. (1989) *Biochem. Biophys. Res. Commun. 157*, 472–479.

Tanzi, R. E., McClatchey, A. I., Lamperti, E. D., Villa-Komaroff, L., Gusella, J. F., & Neve, R. L. (1988) *Nature 331*, 528–530.

Vincent, J., & Lazdunski, M. (1972) *Biochemistry 11*, 2967–2977.

Wlodawer, A., Deisenhofer, J., & Huber, R. (1987) *J. Mol. Biol. 193*, 145–156.