Structure prediction of protease inhibitor region in amyloid precursor protein of Alzheimer's disease

K. Toma, N. Kitaguchi* and H. Ito*

*Computer Science Department and *Bio-Science Laboratory, Life Science Research Laboratories, Asahi Chemical Industry Co. Ltd., Fuji, Shizuoka Japan*

Recent findings of the protease inhibitor domain in amyloid precursor protein of Alzheimer's disease (APPI) raised a novel hypothesis on the mechanism of amyloid deposition in the brain. APPI has significant amino acid sequence homology with Kunitz-type basic trypsin inhibitor superfamily proteins, and the gene expression product showed real inhibitory activity. Since the three-dimensional model of APPI would help in understanding biological phenomena in molecular detail, we constructed an atomic model of APPI based on the structure of bovine pancreatic trypsin inhibitor (BPTI). The substitution of BPTI side chains by best-fitting corresponding amino acid structures was followed by the removal of van der Waals overlappings by molecular mechanics energy minimization with the AMBER force field, to give the feasible model of APPI. We also built serine protease models based on the structure of trypsin and investigated the target enzyme specificity of the inhibitory activity by the active-site mapping method. The models can explain the relative enzyme spectra of APPI and BPTI.

Keywords: protein structure prediction, amyloid precursor protein of Alzheimer's disease, Kunitz-type trypsin inhibitor

INTRODUCTION

Amyloid β-protein is a peptide present in the senile plaques and the cerebrovascular amyloid deposits of patients with Alzheimer's disease and Down's syndrome.[1–3] Several groups have been investigating the precursor of β-protein, which may offer a clue to the mechanism of either amyloid deposition or Alzheimer's disease. Recent molecular cloning of amyloid β-protein precursor (APP) cDNA has revealed the existence of at least three types of APP mRNAs.[4–7] The three APP mRNAs appeared to result from the alternative splicing of single APP gene transcript, based on the determined genomic structure.[5]

Among them, the larger two APPs were shown to contain regions that bore strong amino acid sequence homology with Kunitz-type trypsin inhibitors. The extract of COS-1 cells transfected with the longest precursor cDNA exhibited real inhibitory activity against trypsin.[5] Further study of the inhibitory activity of the molecule has revealed a basically similar but somewhat different enzyme specificity from BPTI.[8]

Potential biological significance of APPI in the pathogenesis of Alzheimer's disease and our interest in protease-inhibitor interaction in common[9] drove us to construct a model three-dimensional (3D) structure of APPI based on the structure of BPTI and models of target enzymes from the structure of trypsin. Those models predict that the inhibitory spectrum of APPI is basically similar to the wide inhibitory spectrum of BPTI and that they differ in affinity toward some enzymes mainly due to the P2′ residue difference.

METHODS

The structures of BPTI and bovine trypsin were taken from the Protein Data Bank (entry name: 2PTC).[10,11] The amino acid sequence of APPI and its alignment with that of BPTI were taken from the literature.[5] Amino acid sequences and alignments of the serine protease catalytic domains were taken from the literature.[12] Bovine and human factor Xa sequences were updated according to the literature,[13,14] and manually aligned. Since there was no difference between bovine and human factor Xa at the important regions covered in the following discussion, we used human factor Xa sequence in the modeling study.

All 3D models of proteins were constructed based on the fact that the tertiary structure is generally more conservative than the primary structure.[15] An all-atom model of APPI was constructed first by changing amino acid side chains

Address reprint requests to Dr. Toma at the Computer Science Department, Asahi Chemical Industry Co. Ltd., 2-1, Samejima, Fuji, Shizuoka 416, Japan.
Received 6 June 1989; accepted 27 June 1989

of BPTI with best-fitting corresponding amino acid structures, and then by removing steric hindrance with molecular mechanics energy minimization using the AMBER (version 3.0) force field.[16] The calculation was performed on a VAX-11/780. All α-carbon based comparative models were built without changing initial structures of BPTI and trypsin and without comprising insertion parts. The color codings of amino acid residues are the same as previously reported.[9]

Three-dimensional models were shown on an Evans and Sutherland PS340, and Color Plates were taken from the screen. FRODO[17] was used for the all-atom model, and the modified version of ALPHA[9,18] was used for the comparative models.

RESULTS AND DISCUSSION

We decided to make a 58 amino acid fragment model of amyloid precursor protein that corresponded to the BPTI structure, since physiological processing points are not yet known at present. No insertion or deletion was found be-

```
          1        10        20        30        40        50        58
BPTI      RPDFCLEPPYTGPCKARIIRYFYNAKAGLCQTFVYGGCRAKRNNFKSAEDCMRTCGGA
APPI      VREVCSEQAETGPCRAMISRWYFDVTEGKCAPFFYGGCGGNRNNFDTEEYCMAVCGSA
           290       300       310       320       330       340 344
```

Figure 1. Sequence alignment of BPTI and APPI. Conserved residues are shown in boldface letters. One letter code of amino acids is used

```
            16  20        30        40        50        60
Trypsin       IVGGYTCGANTVPYQVSLNS---GYHFCGGSLINSQWVVSAAHCYKS--------
Factor X(b)   IVGGRDCAEGECPWQALLVNEE-NEGFCGGTILNEFYVLTAAHClHQAK------
Factor X(h)   IVGGQECKDGECPWQALLINEE-NEGFCGGTILSEFYILTAAHCLYQAK------
Plasmin       VVGGCVAHPHSWPWQVSLRTRF-GMHFCGGTLISPEWVLTAAHCLEKSPRPS---
Thrombin      IVEGSDAEIGMSPWQVMLFRKSPQELLCGASLISDRWVLTAAHCLLYPPWDKNFT
                                |---|                 Δ    |--------
                                 (a)                          (b)

                    70        80        90        100       110
Trypsin       --GIQVRLGEDNINV-VEGNEQFISASKSIVHPSYNSNTL-NNDIMLIKLKSAAS
Factor X(b)   --RFTVRVGDRNTEQ-EEGNEMAHEVEMTVKHSRFVKETY-DFDIAVLRLKTPIR
Factor X(h)   --RFKVRVGDRNTEQ-EEGGEAVHEVEVVIKHNRFTKETY-DFDIAVLRLKTPIT
Plasmin       --SYKVILGAHQEVN-LEPHVQEIEVSRLFLEPTRK-------DIALLKLSSPAV
Thrombin      ENDLLVRIGKHSRTRYERNIEKISMLEKIYIHPRYNWRENLDRDIALMKLKKPVA
              --|           |-|                 |-----|Δ
                            (c)                   (d)

                120       130       140       150       160
Trypsin       LNSRVASISLPT-----SCASAGTQCLISGWGNTKSSG-----TSYPDVLKCLKA
Factor X(b)   FRRNVAPACLPEKDWAEATLMTQKTGIVSGFGRTHEK------GRLSSTLKMLEV
Factor X(h)   FRMNVAPACLPERDWAESTLMTQKTGIVSGFGRTHEK------GRQSTRLKMLEV
Plasmin       ITDKVIPACLPS---PNYVVADRTECFITGWGETQG-------TFGAGLLKEAQL
Thrombin      FSDYIHPVCLPDRETAASLLQAGYKGRVTGWGNLKETWTANVGKGQPSVLQVVNL
                        |-----|                 |------|
                         (e)                     (f)

                    170       180       190       200
Trypsin       PILSDSSCKS--AYPGQITSNMFCAGYLEG---GKDSCQGDSGGPVVCS------
Factor X(b)   PYVDRSTCKL--SSSFTITPNMFCAGYDTQ---PEDACQGDSGGPHVTRFKDT--
Factor X(h)   PYVDRNSCKL--SSSFIITQNMFCAGYDTK---QEDACQGDSGGPHVTRFKDT--
Plasmin       PVIENKVCNRYEFLNGRVQSTELCAGHLAG---GTDSCQGDSGGPLVCFEKDK--
Thrombin      PIVERPVCKD--STRIRITDNMFCAGYKPDEGKRGDACEGDSGGPFVMKSPFNNR
                      |--|                 |---|       Δ      |------
                       (g)                   (h)                 (i)

                210       220       230       240  245
Trypsin       GKLQGIVSWGSGCAQKNKPGVYTKVCNYVSWIKQTIASN
Factor X(b)   YFVTGIVSWGEGCARKGKFGVYTKVSNFLKWIDKIMKARAG*
Factor X(h)   YFVTGIVSWGESCARKGKYGIYTKVTAFLKWIDRSMKTRGL*
Plasmin       YILQGVTSWGLGCARPNKPGVYVRVSRFVTWIEGVMRNN
Thrombin      WYQMGIVSWGEGCDRDGKYGFYTHVFRLKKWIQKVIDQFGE
              |
```

*Figure 2. Sequence alignment of bovine trypsin and catalytic domains of factor Xa (b: bovine; h: human; residues after * are not shown), human plasmin and human thrombin. Residue number is shown with the chymotrypsinogen numbering. Conserved residues are shown in boldface letters. Reactive triad residues are indicated with triangles. Insertions are shown with bars and termed (a) through (i)*

tween amino acid sequences of BPTI and the corresponding APPI fragment and six Cys residues that form three disulfide bonds exactly matched as shown in Figure 1. In the following discussion, we call this 58-residue segment APPI for the sake of convenience.

Building an all-atom model of APPI was a straightforward process, since there is no insertion or deletion in the alignment of amino acid sequences. In the first part of the modeling, we tried to keep original side chain conformations as far as possible. The removal of steric hindrance with molecular mechanics structure optimization using the AMBER force field gave the feasible model structure as shown in Color Plate 1. The final energy was *ca.* −930 kcal/mol, and the model structure was free from van der Waals overlappings. However, it was slightly skewed from the initial BPTI-based structure, possibly due to the ignorance of tightly binding water molecules in the original BPTI structure. The model structure thus constructed is remarkably similar to BPTI structure, which implies a fairly broad spectrum of enzyme specificity of APPI like BPTI.

While APPI showed similar levels of inhibitory activities against factor Xa ($Ki = 1.2 \times 10^{-6}$ M) and plasmin ($Ki = 4.6 \times 10^{-8}$ M), BPTI was more specific toward plasmin ($Ki = 8.9 \times 10^{-11}$ M) than factor Xa ($Ki = 1.5 \times 10^{-3}$ M), which was the most distinctive difference between APPI and BPTI. Both APPI and BPTI were good inhibitors of trypsin ($Ki = 1.1 \times 10^{-10}$ and 1.3×10^{-11}, respectively), and neither showed significant inhibitory activity against thrombin.[8] It is clear from the comparative view of reactive sites of APPI and BPTI (Color Plate 2) that there is no distinctive difference between APPI and BPTI other than the P2′ site, which is Met in APPI and Arg in BPTI, if we consider Arg and Lys at the P1 site almost similar. Therefore, the target enzyme specificity must arise mainly from the P2′ difference.

Since further discussion on the inhibitory specificity seemed to be difficult only from the inhibitor model, we then went into the enzymatic subsite model study. As reported previously,[9] the enzymatic subsites of serine proteases can be estimated from the combined knowledge of their model structures and substrate specificities. We constructed comparative models of trypsin, factor X, plasmin and thrombin based on Greer's alignment[12] (Figure 2) as shown in Color Plate 3.

We stressed with colors the important residues in the following discussion on the simple graphics model (Color Plate 3a). Insertion (d) in Figure 2, which only thrombin has, may cause its stringent substrate specificity at the P2 site; as a result, neither APPI nor BPTI is a good inhibitor toward thrombin. A common main component of the S2′ site, Phe 41, in trypsin, factor Xa and plasmin may explain similar magnitude of binding affinity of APPI toward factor Xa and plasmin with hydrophobic interactions between P2′-Met and S2′-Phe. In the case of BPTI, Arg at P2′ has a relatively long side chain, and the hydrophobic nature of methylene fragments must be considered. From the inspection of the trypsin–BPTI complex structure, the side chain of Arg 17 of BPTI appeared to contact first with hydrophobic interaction at the S2′ site, and the guanidine group leans to the direction of Asn 143 and Tyr 151 of trypsin because of the lack of space enough to hold all side chain atoms of Arg. In the model of factor Xa, basic Arg 143 and Arg 150 come to this extended S2′ site for BPTI, and acidic Glu 143 does in the model of plasmin. These extended S2′ subsites may cause favorable interaction with plasmin and unfavorable interaction with factor Xa for BPTI, and P2′ Met of APPI is small enough to fit the S2′ site and remains neutral in this respect. Thus, the subsite model can explain relative target enzyme selectivities of these inhibitors fairly well.

Recently, Kido *et al.* reported strong factor Xa inhibitory activity of trypstatine,[19] which also belongs to the Kunitz-type trypsin inhibitor superfamily. The reactive site amino acid sequences of BPTI, APPI and trypstatine are as follows:

	P3	P2	P1	P1′	P2′	P3′
BPTI	Pro	Cys	Lys	Ala	Arg	Ile
APPI	Pro	Cys	Arg	Ala	Met	Ile
trypstatin	Pro	Cys	Arg	Ala	Phe	Ala

The P2′ site Phe of trypstatine may interact with the postulated S2′ site Phe 41 of factor Xa, and trypstatine may thus become a good inhibitor of it. This further supports our enzymatic subsite model. Since Phe 41 is conserved between trypsin and factor Xa, small Gly 40 of factor Xa would help trypstatine to gain affinity with wider aromatic interaction space.

CONCLUSION

Several types of models of APPI and enzymes can be used to explain present experimental results, especially those of enzyme specificities. Although these models are approximations, they will help to explain the biological meanings of APPI from a molecular structural point of view.

ACKNOWLEDGEMENT

This study was carried out in part through Special Coordination Funds of the Science and Technology Agency of the Japanese Government.

NOTE

APPI coordinates available upon request from K. Toma.

REFERENCES

1 Glenner, G. G. and Wong, C. W. Alzheimer's disease: initial report of the purification and characterization of a novel cerebrovascular amyloid protein. *Biochem. Biophys. Res. Commun.* 1984, **120,** 885–890
2 Glenner, G. G. and Wong, C. W. Alzheimer's disease and Down's syndrome: sharing of a unique cerebrovascular amyloid fibril protein. *Biochem. Biophys. Res. Commun.* 1984, **122,** 1131–1135
3 Masters, C. L., Simms, G., Weinman, N. A., Multhaup, G., McDonald, B. and Beyreuther, K. Amyloid plaque core protein in Alzheimer disease and Down syndrome. *Proc. Natl. Acad. Sci. USA* 1985, **82,** 4245–4249
4 Carrell, R. W. Enter a protease inhibitor. *Nature* 1988, **331,** 478–479
5 Kitaguchi, N., Takahashi, Y., Tokushima, Y., Shiojiri,

S. and Ito, H. Novel precursor of Alzheimer's disease amyloid protein shows protease inhibitory activity. *Nature* 1988, **331,** 530–532
6 Ponte, P., Gonzalez-DeWhitt, P., Schilling, J., Miller, J., Hsu, D., Wallace, W., Lieberburg, I., Fuller, F. and Cordell, B. A new A4 amyloid mRNA contains a domain homologous to serine proteinase inhibitors. *Nature* 1988, **331,** 525–527
7 Tanzi, R. E., McClatchey, A. I., Lamperti, E. D., Villa-Komaroff, L., Gusella, J. F. and Neve, R. L. Protease inhibitor domain encoded by an amyloid protein precursor mRNA associated with Alzheimer's disease. *Nature* 1988, **331,** 528–530
8 Kitaguchi, N., Takahashi, Y., Oishi, K., Shiojiri, S., Tokushima, Y., Utsunomiya, T. and Ito, H. Enzyme specificity of protease inhibitor region in amyloid precursor protein of Alzheimer's disease, submitted
9 Toma, K. and Suzuki, K. Mapping active sites of blood coagulation serine proteases—activated protein C and thrombin—on simple graphics models. *J. Mol. Graphics,* 1989, **7**(3), 146–149
10 Protein Data Bank, Chemistry Department, Brookhaven National Laboratory, New York, USA
11 Marquart, M., Walter, J., Deisenhofer, J., Bode, W. and Huber, R. The geometry of the reactive site and of peptide groups in trypsin, trypsinogen, and its complex with inhibitors. *Acta. Cryst.* 1983, **B39,** 480–490
12 Greer, J. Comparative model-building of the mammalian serine proteases. *J. Mol. Biol.* 1981, **153,** 1027–1042
13 Fung, M. R., Campbell, R. M. and MacGillivray, R. T. A. Blood coagulation factor X mRNA encodes a single polypeptide chain containing a prepro leader sequence. *Nuc. Acids Res.* 1984, **12,** 4481–4492
14 Fung, M. R., Hay, C. W. and MacGillivray, R. T. A. Characterization of an almost full-length cDNA coding for human blood coagulation factor X. *Proc. Natl. Acad. Sci. USA* 1985, **82,** 3591–3595
15 Blundell, T. L., Sibanda, B. L., Sternberg, M. J. E. and Thronton, J. M. Knowledge-based prediction of protein structures and the design of novel molecules. *Nature* 1987, **326,** 347–352
16 Weiner, S. J., Kollman, P. A., Case, D. A., Singh, U. C., Ghio, C., Alagona, G., Profeta, Jr., S. and Weiner, P. A new force field for molecular mechanical simulation of nucleic acids and proteins. *J. Am. Chem. Soc.* 1984, **106,** 765–784
17 Jones, T. A. Interactive computer graphics: FRODO. *Methods Enzymol.* 1985, **115,** 157–171
18 Toma, K. Simple protein model building tool. *J. Mol. Graphics* 1987, **5,** 101–102
19 Kido, H., Yokogoshi, Y. and Katunuma, N. Kunitz-type protease inhibitor found in rat mast cells. Purification, properties, and amino acid sequence. *J. Biol. Chem.* 1988, **263,** 18104–18108