Two Reactive Site Locations and Structure−Function Study of the Arrowhead Proteinase Inhibitors, A and B, Using Mutagenesis†

Zhi-Wei Xie, Ming-Juan Luo, Wen-Feng Xu,§ and Cheng-Wu Chi*,‡

State Key Laboratory of Molecular Biology, Shanghai Institute of Biochemistry, Academia Sinica, Shanghai 200031, China

Received December 6, 1996; Revised Manuscript Received March 3, 1997⊗

ABSTRACT: The arrowhead (*Sagittaria sagittifolia*, Linn.) proteinase inhibitor A and B are double-headed and multifunctional, consisting of 179 amino acid residues with three disulfide bridges. Both their primary structures and cDNA sequences have been elucidated [Yang, H. L., Luo, R. S., Wang, L. X., Zhu, D. X., & Chi, C. W. (1992) *J. Biochem. 111,* 537; Xu, W. F., Tao, W. K., Gong, Z. Z., & Chi, C. W. (1993) *J. Biochem. 113*, 153; Luo, M. J., Lu, W. Y., & Chi, C. W. (1997) *J. Biochem*. (in press)]. Though they share 91% homology, they are different in inhibitory activities. Sequence analysis of their full-length cDNAs showed that there are seven extra residues in the C-terminal part which might be cleaved off by proteinase post-processing. To locate the reactive sites and study the structure−function relationship of the two forms A and B, the genes coding for the mature inhibitor B and its extended form were respectively cloned into the secretion expression vector, pVT102U/α, and expressed in *Saccharomyces cerevisiae* strain S-78. Both of the gene products were purified and characterized to have the same inhibitory activities as the natural one. The gene product of the extended form was a mixture with the extended C-terminal part of the inhibitor either completely or partially removed. The two previously predicted reactive site residues, Lys-44 and Arg-76 of inhibitor B, were then respectively substituted with Ala by site-directed mutagenesis and expressed. As compared with the natural inhibitor, each of the mutants could only inhibit one molecule, instead of two molecules of trypsin, and displayed an inhibitory activity against elastase, thus confirming the location of the two reactive sites in the inhibitors. The gene coding for inhibitor A, which for some reason could not be expressed in *S. cereviciae*, was successfully expressed in the reconstructed plasmid pET-1522bx in *Escherichia coli* strain BL21 with the expressed product existing in the inclusion body. After denaturation and renaturation, the active inhibitor A was obtained and purified by anhydrotrypsin affinity chromatography. Using site-directed mutagenesis, two residues of inhibitor A, namely, Ser-82 and Leu-87, prominently different from Leu-82 and Arg-87 in inhibitor B, were replaced by these two corresponding residues, respectively. As compared with the natural inhibitor A, its S82L mutant showed a lower inhibitory activity toward trypsin, whereas a higher activity was found in the L87R mutant. Meanwhile, both of their chymotrypsin inhibitory activities became weaker than the natural one. The important accessary role of the residue of position 87 in causing the difference in inhibitory properties between inhibitor A and B was discussed.

The arrowhead proteinase inhibitors, A and B, are the two major inhibitor components purified from the tubes of arrowhead (*Sagittaria sagittifolia*, Linn.) (Zhang *et al*., 1979). They are double-headed and multifunctional, capable of inhibiting many different serine proteinases, including trypsin, chymotrypsin as well as tissue kallikrein. The primary structures of both inhibitors A and B have been elucidated to be composed of 179 amino acid residues with three disulfide bridges (Yang *et al*., 1992; Luo *et al*., 1997). Their structures are quite unique, with no homologous sequence nor apparent domain boundary usually found in other double-headed proteinase inhibitors (Laskowski *et al*., 1980). Experiments with the chemical modification of the inhibitors suggested Lys and Arg to be essential for their inhibitory activities (Yang *et al*., 1991). Based on the comparison of amino acid sequences around the reactive sites of different types of proteinase inhibitors, it was predicted that the two reactive sites of the inhibitor were most likely located at Lys-44 and Arg-76. Since the arrowhead inhibitor shows no significant homology with any of other proteinase inhibitors, it might be classified as a new family, unique both in structure and function.

On the basis of primary structures of inhibitors A and B, their full-length cDNA and genomic structures have also been elucidated (Xu *et al*., 1993). Their open reading frame encode a pre-inhibitor, including a 24-residue signal peptide, a mature protein and seven extra residues in the C-terminal part of the inhibitor, which are cleaved off during post-translational processing immediately after protein synthesis.

Though inhibitor A and B share over 91% sequence homology, including two similar reactive sites, and most of the 17 nonidentical residues are conserved, their inhibitory features are quite different. Inhibitor A inhibits equimolar

† This work was supported by a state biological high technology research grant from the government of China and by the research funds provided by De Montfort University for the Chinese Centre of the Norman Borlaug Institute for Plant Science Research.

§ Present address: University of Washington, Department of Biochemistry, Seattle, WA 98195.

‡ Senior research fellow at De Montfort University Norman Borlaug Institute for Plant Science Research.

* To whom correspondence should be addressed at Shanghai Institute of Biochemistry, Academia Sinica, 320 Yue-Yang Road, Shanghai 200031, China. FAX: 86-21-6433 8357.

⊗ Abstract published in *Advance ACS Abstracts,* April 15, 1997.

S0006-2960(96)02993-5 CCC: $14.00

Table 1: Oligonucleotides Used for Site-Directed Mutagenesis

sequence of oligonucleotide 5′−3′ (mutate codons are bold)	amino acid residue location		mutation
	inhibitor B	inhibitor A	
1 GGACGTCTGC**GCT**AGCTACGTCT	DAC**A**SYV 41−47		K44A
2 GAACTTGTA**GGC**GGACCCCAGC	FKY**A**SGL 79−73		R76A
3 CCGGCAT**CAA**GAAGGAG		SF**L**MP 80−84	S82L
4 CGCAGAT**GCG**CGGTACC		VP**R**IC 85−89	L87R

amount of trypsin and chymotrypsin simultaneously and is weakly active on kallikrein, whereas inhibitor B acts on two molecules of trypsin at the same time and inhibits kallikrein more strongly than inhibitor A does.

To further elucidate the characteristics of these inhibitors, the exact location of the two reactive sites and the structure−function relationship between the two forms, A and B, became imminent. Since the arrowhead inhibitor is very difficult to be modified either by trypsin or chymotrypsin at low pH, the conventional approach used to determine the reactive site of inhibitors by modification with the corresponding enzyme was not feasible. In the modified inhibitor the reactive site peptide bond is specifically cleaved, which then can be elucidated by amino acid sequencing. The recombinant gene expression and site-directed mutagenesis techniques were then employed for the arrowhead inhibitor.

In this paper, we describe the gene expression of the arrowhead inhibitors and their mutants by site-directed mutagenesis with results confirming the previously predicated Lys-44 and Arg-76 as the reactive site residues. The possible accessary role of some residues in the neighborhood of the reactive site in determining the differential activities of the two forms is also discussed.

MATERIALS AND METHODS

Materials. All of the restriction enzymes, DNA polymerase I (large fragment), T4 DNA ligase, T4 kinase, isopropyl α-D-thiogalactopyranoside (IPTG),[1] and X-gal were purchased from Bethesda Research Laboratories (BRL). The Sequenase Version 2.0 DNA sequencing system was from United States Biochemical (USB). [α-^{32}P]dATP (3000 Ci/mmol) was from Amersham. The natural arrowhead proteinase inhibitors A and B were prepared according to the previously described method (Zhang *et al*., 1979). Bovine trypsin and porcine elastase were purchased from Sigma Chemical, and porcine pancrease kallikrein and bovine chymotrypsin were prepared in this laboratory. Immobilized trypsin was prepared according to the described method (Yang *et al*., 1992). *N*-Succinyl-L-alanyl-L-alanyl-L-alanyl-*p*-nitroanilide (Suc-A-A-A-*p*NA), tosylarginine methyl ester (TAME), benzoyltyrosine ethyl ester (BTEE), and ATP were from the Shanghai Dong-feng Biochemical Reagent Factory. All other reagent were of analytical grade. Mutagenic oligonucleotides and PCR primers were synthesized with an Applied Biosystems 380A DNA synthesizer.

Escherichia coli strain TG1 was kindly given by Dr. En-Duo Wang. The phages M13 mp18 and mp19 were purchased from BRL and used for sequencing with Sequenase (Version 2.0) using the Sanger dideoxynucleotide chain termination method (Sanger *et al*., 1977). *E. coli* strain RZ1032 (*supE*, *duf*, *ung*$^-$) was used for site-directed mutagenesis (Kunkel *et al*., 1987). *Saccharomyces cerevisiae* strain S-78 (*leu2*, *ura3*, *rep4*), and the yeast secretion expression vector, pVT102U/α, were gifts from Dr. You-Shang Zhang. *E. coli* strain BL21 and expression plasmid pET were kindly given by Dr. Xun Xu. DNA manipulation was carried out by standard procedures (Sambrook *et al*., 1989).

Site-Directed Mutagenesis. Oligonucleotide-directed mutagenesis was performed as described (Kunkel *et al*., 1987). The phage M13 mp18 or mp19 containing the gene encoding the arrowhead inhibitor was used to transform *E. coli* strain RZ1032. The single-stranded DNA, in which some thymidines were substituted with uracil, could survive in this strain. The synthetic deoxyoligonucleotides (Table 1) were annealed to this substituted DNA prior to synthesis of the double-stranded DNA. The recombinants were used to transform *E. coli* strain TG1. All of the mutant sequences were confirmed by the DNA sequence analysis using the Sanger dideoxynucleotide chain termination method.

Polymerase Chain Reaction. PCR was used to amplify the genes encoding the mature inhibitor and its extended form with seven extra residues at the C-terminus. In each PCR reaction, the forward primer with an *Eco*RI site at the 5′-end was designed and synthesized corresponding to the N-terminal part of the natural inhibitor B, Asp1−Ser6: 5′-CG**GAATTC**TGATCCCGTCGTCGACAGCG-3′ (the *Eco*RI site is bolded). In order to make the reading frame of inhibitor B compatible with the expression vector, pVT102U/α, in the primer an extra nucleotide T was inserted between the *Eco*RI site and the first codon GAT. The reversed primers were designed and synthesized corresponding to the C-terminal part of the natural inhibitor, His174−Leu179, and to that of its extended form with seven extra residues, Lys182−Gln186, both flanked with a *Hin*dIII site at the 5′-end: 3′-GTATTCAAGCTGCGTGAGA**TTCGAA**GG-5′ and 3′-CTTTTGACGTGGCGTCA**TTCGAA**CC-5′ (the *Hin*dIII site is bolded), respectively. PCR was carried out by using the GeneAmp DNA amplification kit on a Perkin Elemer Cetus DNA thermal cycler. The PCR products were recovered from a 6% polyacrylamide gel after electrophoresis as described (Maxam *et al*., 1977).

Determination of Inhibitory Activities. The assay of trypsin inhibitory activity was performed in 3 mL of 0.02 M Tris, pH 7.8, 10 mM $CaCl_2$, containing 5 μg of trypsin and various amounts of the natural or recombinant inhibitors using 0.5 mM TAME as a substrate. The residual trypsin activity was measured at 247 nm. Chymotrypsin inhibitory activity was assayed essentially as described (Chu *et al*., 1965) by using 0.5 mM BTEE as a substrate with 6 μg of the enzyme. The inhibitory activity on elastase was assayed

[1] Abbreviations: PCR, polymerase chain reaction; HPLC, high-performance liquid chromatography; PAGE, polyacrylamide gel electrophoresis; bp, base pair(s); TAME, tosylarginine methyl ester; BTEE, benzoyltyrosine ethyl ester; BAEE, benzoylarginine ethyl ester; PBS, phosphate buffer solution; Suc-A-A-A-*p*NA, *N*-succinyl-L-alanyl-L-alanyl-L-alanyl-*p*-nitroanilide; IPTG, isopropyl 2-D-thiogalactopyranoside; X-gal, 5-bromo-4-chloro-3-indolyl β-D-galactoside; APIA, arrowhead protease inhibitor A; APIB, arrowhead protease inhibitor B.

as described (Rolka *et al*., 1989) using 0.6 mM Suc-A-A-A-*p*NA as a substrate with 10 μg of the enzyme. The inhibitory activity against kallikrein was assayed according to the method described previously (Zhang *et al*., 1979) using 0.8 mM BAEE as a substrate with 5 μg of the enzyme.

Gene Expression of Inhibitor B and Its Extended Form in the Yeast Secretion System. The gene fragments encoding the mature inhibitor B and its extended form with seven extra residues in the C-terminal part, obtained by the PCR method, were cleaved with *Eco*RI/*Hin*dIII, and ligated with the expression vector, pVT102U/α, through the *Xba*I/*Eco*RI linker, respectively, as previously described (Cheng *et al*., 1992). The vector has the mating α-factor secretion system to direct the expression of fused foreign products. The ligated mixture was used to transform TG1. The recombinant plasmid was confirmed with the DNA sequence determination and used to transform *S. cerevisiae* strain S-78. The transformant was then grown overnight in 2 mL of synthetic selected YSD medium, transferred to 50 mL of YPD medium, and further cultured at 30 °C for 48 h. The inhibitory activity was found in the supernatant.

Gene Expression of Inhibitor A and Its Mutants in E. coli. In order to express inhibitor A, which could not be successfully expressed in *S. cerevisiae*, an expression plasmid pET-1522bx was reconstructed from two plasmids pET15b and pET22b for gene expression in *E. coli* (Xie *et al*., 1997). The gene coding for inhibitor A was amplified by the PCR method using its positive clone from the arrowhead cDNA library as a template. The natural and the mutated genes of inhibitor A were then cloned into pET-1522bx with *Eco*RI/*Hin*dIII, respectively. Transformed cells with recombinant plasmids were grown in the LB medium containing 100 μg/mL of ampicillin at 30 °C for 2−2.5 h (OD_{600} = 0.8−1.0). After IPTG was added to a final concentration of 0.1 mM, cells were continuously grown at 37 °C for 4−5 h. The cells were harvested by centrifugation and lysed by lysozyme. The soluble or insoluble part of the lysate was analyzed by SDS−PAGE, which revealed that the expressed products existed in the inclusion body. After purification as described (Winkler *et al*., 1986), the inclusion body collected from 200 mL of cultured cells was suspended in 3 mL of 8 M urea, 50 mM mercaptoethanol, 0.1 M PBS, pH 7.5, and incubated with stirring overnight at room temperature. After centrifugation at 10000*g* for 20 min to remove insolubles, the supernatant was dialyzed against 400 mL of 8 M urea, 1 mM mercaptoethanol, 0.1 mM PBS, pH 7.5. The dialyzed solution was then dropped into 150 mL of the denaturation buffer, containing 2.5 M urea, 50 mM Tris-HCl, pH 8.0, 5 mM EDTA, 10 mM NaCl, 0.005% Tween-80, 1.25 mM reductive glutathione, and 0.25 mM oxidative glutathione at room temperature for over 20 h and then dialyzed against 1 M urea. The renaturation solution was applied to the anhydrotrypsin affinity chromatography column. The inhibitor fractions were finally eluted out with 0.02 N HCl. The yield of the purified recombinant inhibitor was around 3 mg per liter of media.

RESULTS

Gene Expression of Inhibitor B and Characterization of Its Gene Products. Since the gene of inhibitor B was fused with the DNA sequence encoding the leading peptide of α-mating factor in the expression plasmid, the gene product was first processed by the KEX2 proteinase inherent in yeast cells and then directly secreted into the culture supernatant. The procedure of purification of the recombinant inhibitor was mainly performed as previously described (Chen *et al*., 1992). The supernatant was concentrated, and the pH was adjusted to 8.0 with Tris base, followed by gel filtration on a Sephadex G-25 column. The fractions with the inhibitor activity were pooled and concentrated, and further purified by using affinity chromatography with immobilized trypsin. The yield of the purified recombinant inhibitor in each batch was around 4 mg per liter of media. Both the purified recombinants of inhibitor B and its extended form were homogeneous on SDS gel electrophoresis (Figure 1). The amino acid sequence analysis revealed that the first ten residues in the N-terminal part of the expressed inhibitor were Asn-Ser-Asp-Pro-Val-Val-Asp-Ser-Asp-Gly. They are consistent with those of the natural inhibitor except for the two additional residues, Asn-Ser, at the N-terminus, which were delivered from the oligonucleotide linker between the two restriction sites (*Xba*I/*Eco*RI). This proved that the products were accurately processed by KEX2 proteinase in the secretion plasmid as expected.

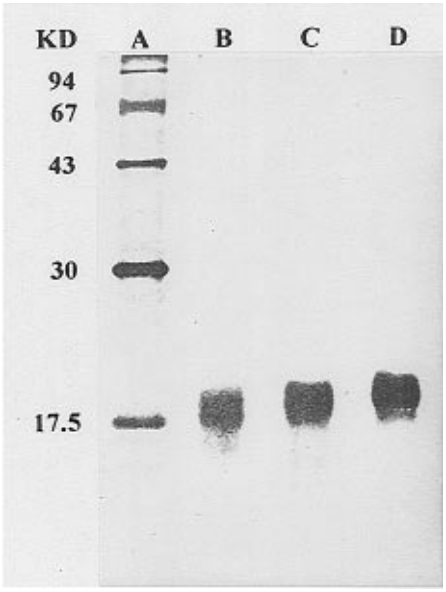

FIGURE 1: SDS−PAGE patterns of the purified arrowhead inhibitor B expressed in yeast. The lyophilized samples were dissolved in 0.66 M Tris-HCl, pH 6.8, 2% SDS, 10% glycerol, 5% mercaptoethanol, 0.1% bromophenol blue, and then subjected to electrophoresis on a 15% polyacrylamide gel, respectively. Lane A, standard marker proteins; lane B, the natural inhibitor B; lanes C and D, the recombinant inhibitor B with and without seven extra residues in the C-terminal part.

The molecular weight of the recombinant inhibitor B with seven extra residues, Ala-Met-Lys-Thr-Ala-Pro-Gln, in the C-terminal part of the inhibitor was precisely determined by the matrix-assisted laser-desorption mass spectrometry. The results showed that the gene product was actually a mixture, that included the extended inhibitor B and its completely or partially post-processed products. The deleted parts exactly corresponded to Thr-Ala-Pro-Gln and Lys-Thr-Ala-Pro-Gln (Figure 2).

The recombinant inhibitor B showed the same inhibitory activities against trypsin and chymotrypsin as the natural one (data not shown). It implied that the two additional amino acid residues at the N-terminal part, namely, Asn-Ser, had no effect on the inhibitory activities of the inhibitor.

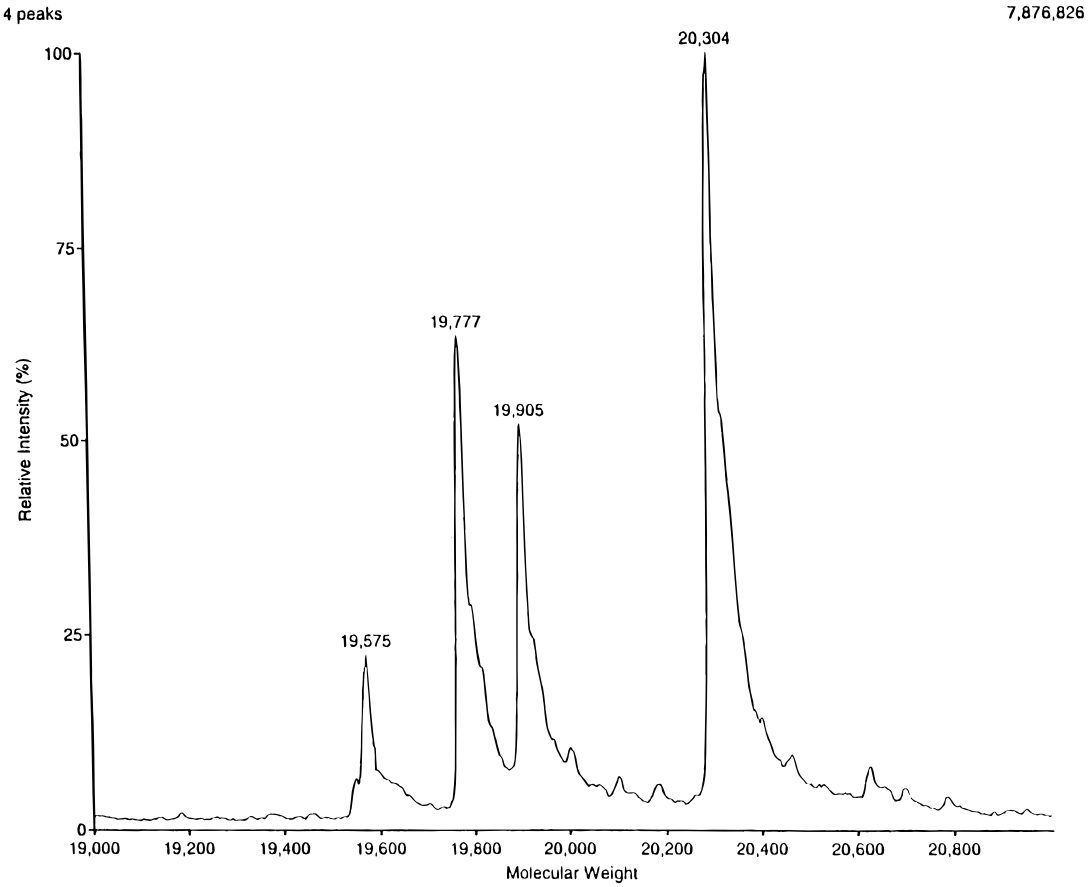

FIGURE 2: Mass spectrographic analysis of the recombinant inhibitor B with seven extra residues at the C-terminal part. The peak with molecular weight 19 575 corresponds to inhibitor B without additional seven extra residues in the C-terminal part; the peaks with molecular weights 19 777, 19 905, and 20 304 exactly correspond to the products with two (Ala-Met, MW 202), three (Ala-Met-Lys, MW 330), and seven extra residues (Ala-Met-Lys-Thr-Ala-Pro-Gln, MW 728) in the C-terminal part, respectively.

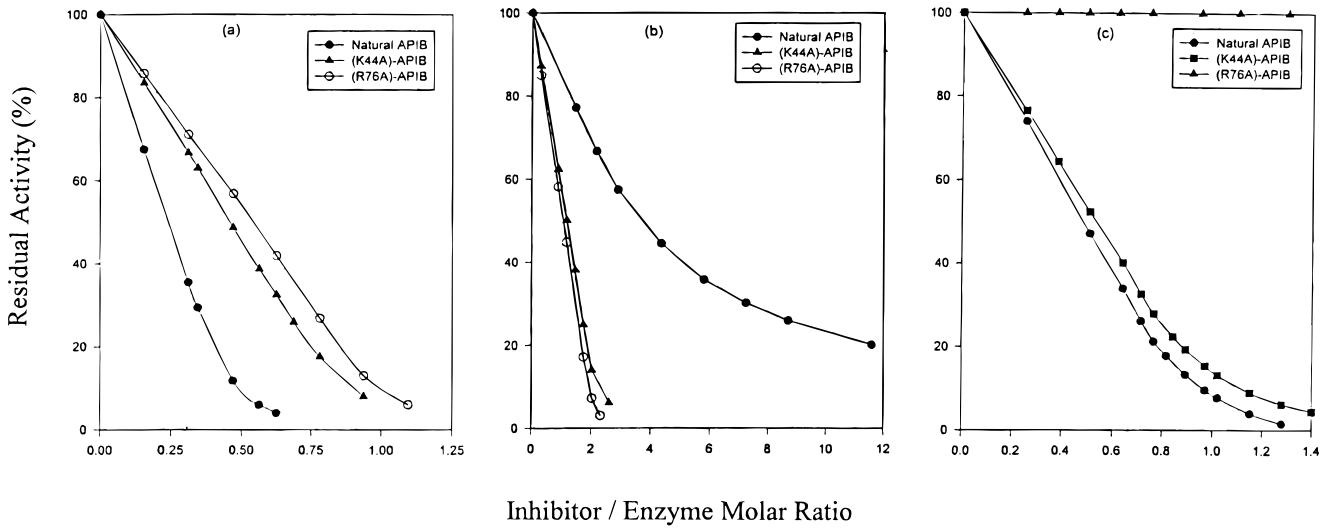

FIGURE 3: Inhibition curves of the natural inhibitor B and its recombinant mutants toward trypsin (a), elastase (b), and kallikrein (c). (a) Inhibitory activity toward trypsin with TAME as a substrate (0.5 mM in 20 mM Tris buffer, pH 7.8), trypsin 5 μg, reaction time 5 min at 37 °C, measured at 247 nm. (b) Inhibitory activity toward elastase with Suc-A-A-A-*p*NA as a substrate (0.6 mM in 0.1 M Tris buffer, pH 8.0), elastase 10 μg, reaction time 5 min at 37 °C, measured at 390 nm. (c) Inhibitory activity toward kallikrein with BAEE as a substrate (0.8 mM in 20 mM Tris buffer, pH 7.8), kallikrein 5 μg, reaction time 5 min at 25 °C, measured at 253 nm.

Location of the Two Reactive Sites of Inhibitor B. Two mutants of inhibitor B, K44A and R76A, were expressed and purified using the same procedure as described above. Their inhibitory activities toward trypsin and elastase were determined (Figure 3). The results revealed that compared with the natural inhibitor, their inhibitory activity toward elastase become much stronger as expected, whereas their inhibitory activity toward trypsin was reduced by half with only one molecule instead of two molecules of trypsin affected, as in each case one of the two reactive sites remained unmutated in the mutants. The results definitely indicated that the inhibitor has two reactive sites located at positions 44 or 76. Compared with the natural inhibitor, the inhibitory activity on kallikrein remained the same in the

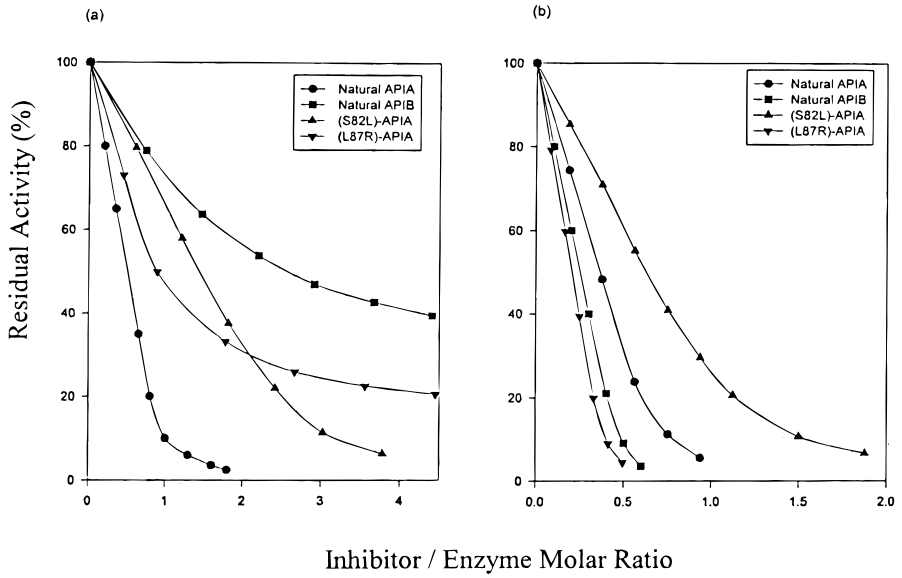

FIGURE 4: Inhibitory activities of the recombinant inhibitor A and its mutants against (a) chymotrypsin and (b) trypsin. (a) Inhibitory activity toward chymotrypsin with BTEE as a substrate (0.5 mM in 0.05M Tris buffer, pH 8.0), chymotrypsin 6 μg, reaction time 5 min at 25 °C, measured at 256 nm. (b) Inhibitory activity toward trypsin, same conditions as in Figure 3.

K44A mutant but was totally lost in the R76A mutant. It implied that the first reactive site, Lys-44, is specific for inhibiting trypsin only, while the second one, Arg-76, is multifunctional, capable of inhibiting not only trypsin, but also chymotrypsin as well as kallikrein. Once this multifunctional second reactive site, Arg-76, was mutated, its corresponding inhibitory specificities to different proteinases, trypsin and kallikrein, were also changed.

Characterization of the Recombinant Inhibitor A and Its Mutants. The inhibitory activities against trypsin and chymotrypsin of the recombinant inhibitor A displayed the same activity as the natural one (data not shown). The substitution of Ser 82 with Leu or Leu 87 with Arg in inhibitor A resulted in reducing the inhibitory activity toward chymotrypsin (Figure 4). Especially the L87R mutant behaving as inhibitor B, even at an excess amount, could not predominantly inhibit chymotrypsin, while the S82L mutant still kept the inhibitory feature as the natural one, however, the inhibitor/enzyme ratio at extrapolation to total inhibition became 3 instead of 1. The inhibitory activity against trypsin of the L87R mutant was obviously increased, and almost the same within the experimental error to that of inhibitor B, whereas the inhibitory activity of the S82L mutant was diminished (Figure 4). It implied that the hydrophilic residue, Ser-82, and the basic residue, Arg-87, neighboring the second reactive site, are important for the action of the inhibitor against trypsin, and the residue, Arg-87, might be one of the most important residues responsible for causing difference between the inhibitory activities of inhibitor A and B.

DISCUSSION

The arrowhead proteinase inhibitor A and B are double-headed and multifunctional, capable of inhibiting many different serine proteinases including trypsin, chymotrypsin, and tissue kallikrein. Compared with other proteinase inhibitors classified by Laskowski and Kato (1980), arrowhead inhibitors showed no significant homology with any of them. Considering their unique structure and function, the arrowhead inhibitor might be classified to a new family. Additionally, the inhibitor might have potential application value because of its multifunctional property against different proteinases, high specific inhibitory activity and stability. For example, the gene encoding inhibitor B has been successfully transferred into cotton and cauliflower, the preliminary results seemed promising that the transgenic plants became insect resistant (data not shown).

The cDNA deduced amino acid sequences of inhibitor A and B are consistent in principle with those determined by primary structure analysis (Yang *et al*., 1992; Xu *et al*., 1993; Luo *et al*., 1997) except that there are seven extra residues in the C-terminal part of the inhibitors, which might be cleaved off by proteinase post-processing immediately after protein synthesis. The same thing occurred for the bovine pancreatic trypsin inhibitor, which has 17 residues split off from the C-terminal part after gene expression (Anderson *et al*., 1983). When the yeast secretion expression vector, pVT102U/α, was used, the gene coding for the extended inhibitor was also expressed in yeast but the gene product was proved to be a mixture by mass spectrographic analysis, composed of the intact gene product with the seven extra residues and the completely or partially post-processed products. It indicated that yeast cells share the similar post translation process system with plant cells though their recognition sites of processing are not entirely the same.

The previous prediction of two reactive sites of the arrowhead inhibitor was in part based on the sequence comparison. The neighboring sequence of the first predicted reactive site Lys-44 is -Cys-Lys-Ser-Tyr-, whose P_2 residues Cys conforms to that in the reactive site of bovine pancreas Kunitz inhibitor (Gebhard *et al*., 1986), whereas its P_2' residue is an aromatic amino acid, Tyr, the same as in many Kazal type inhibitors (Greene *et al*., 1969). The sequence around another predicted reactive site Arg-76 is -Arg-Tyr-Lys-Phe-, similar to that around the reactive site of the soybean Kunitz inhibitor (-Arg-Ile-Arg-Phe-). In the structure of the arrowhead inhibitor, only this sequence meets the rule that the P_2' residue in the reactive site sequence of all kallikrein inhibitors should be a basic residue, namely, Arg(Lys)-X-Arg(Lys) (Fritz *et al*., 1983). Studies on the

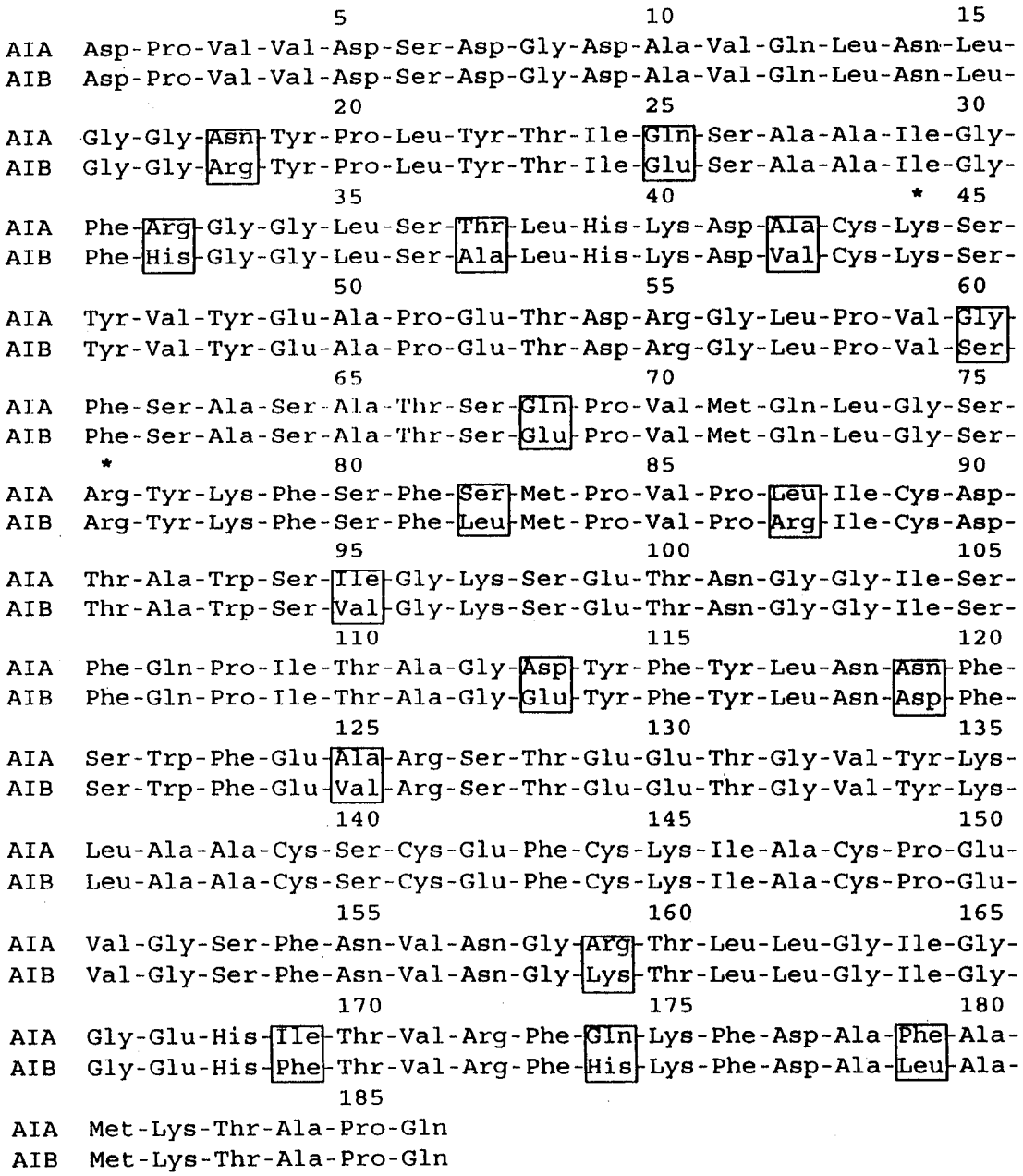

FIGURE 5: Comparison between the amino acid sequences of arrowhead inhibitors A and B. The different residues are framed. The asterisk indicates the reactive site of the inhibitor.

inhibitory activity characters of the inhibitor B mutants, K44A- and R76A-inhibitor B, definitely confirmed the previous assignment of this two reactive sites. Especially compared with the natural inhbitor, the R76A mutant not only displays an inhibitory activity toward elastase but also lost all its inhibitory activity toward kallikrein, indicating that Arg-76 is the second reactive site and is responsible for the multifunctional inhibitory feature.

The arrowhead inhibitors A and B share 91% homology, but their inhibitory properties are different. Inhibitor B shows a stronger affinity for trypsin and kallikrein than A, whereas A inhibits chymotrypsin more efficiently than B does. Unlike inhibitor B, inhibitor A could not be expressed in the yeast secretion expression system. The reason might be related with its higher inhibitory activity toward chymotrypsin. It is reported that there is a chymotrypsin-like enzyme in yeast cells (Fujishiro *et al*., 1980; Kominami *et al*., 1981), which might affect the correct expression or secretion of the expressed inhibitor A in yeast. The same situation happened when we studied the structure−function relationship of the *Trichosanthes* trypsin inhibitor, a squash family inhibitor (Ling *et al*., 1993), its mutant with a chymotrypsin inhibitory activity also could not be expressed in this yeast secretion system.

Comparing the protein sequences of inhibitor A and B, among the 179 amino acid residues, only 17 of them are different, and most of the different residues are conserved, for example, Gln-Glu 25, Gln-Glu 68, Asp-Glu 113, Asn-Asp 119, Arg-Lys 159, etc. (Figure 5). Some of the unconservative residues, such as Thr-Ala 37, Ile-Val 95, Ala-Val 125, and other residues being situated in the N-terminal or the C-terminal part at positions Asn-Arg 18, Arg-His 32, Ile-Phe 169, Gln-His 174, Phe-Leu 179, are thought unlikely to influence the inhibitory features. Among the remaining different residues, Ser/Leu 82 and Leu/Arg 87, neighboring the second reactive site, seem to be important in causing the difference in inhibitory properties. The above predication was confirmed by the present results using mutagenesis to

study the inhibitory properties of the mutant S82L and L87R of inhibitor A. It was indicated that a hydrophilic residue Ser at position 82 and a basic residue Arg at position 87 near the second reactive site are important to the trypsin inhibitory activity, whereas a hydrophobic residue Leu at position 87 might play a crucial role in making the inhibitor more specific for chymotrypsin. For a better understanding of the structure−function relationship of the inhibitor, it is desirable to elucidate the crystal structures of inhibitor A and B. Fortunately, the crystal of the complex of inhibitor B and trypsin at a molar ratio of 1:2 has been obtained, and the study on the three-dimensional structure of this complex is in progress.

ACKNOWLEDGMENT

We thank Dr. Wu-yuan Lu for analyzing the molecular weights of the recombinant inhibitor B and its extended form by mass spectrometry.

REFERENCES

Anderson S., & Kingston, I. B. (1983) *Proc. Natl. Acad. Sci. U.S.A. 80*, 6838−6842.
Chen, X. M., Qian, Y. W., Chi, C. W., Gan, K. D., Zhang, M. F., & Chen, C. Q. (1992) *J. Biochem. 112*, 45−51.
Chu, H. M., & Chi, C. W. (1965) *Sci. Sin. 14*, 1441−1453.
Fritz, H., & Wunderer, G. (1983) *Arzneim. Forsch. Drug Res. 33*, 479−494.
Fujishiro, K., Sanada, Y., Tanaka, H., & Katunuma, N. (1980) *J. Biochem. 87*, 1321−1326.
Gebhard, W., Tshesche, H., & Fritz, H. (1986) in Proteinase Inhibitors (Barret, A. J., & Salvensen, G., Eds.) pp 375−388, Elsevier, Amsterdam.
Greene, L. J., & Byrteit, D. C. (1969) *J. Biol. Chem. 244*, 2646−2657.
Kominani, E., Hoffschulte, H., & Holzer, H. (1981) *Biochim. Biophys. Acta 66*, 124−135.
Kunkel, T. A., Roberts, J. B., & Zabour, R. A. (1987) *Methods Enzymol. 154*, 367−382.
Laskowski, M., Jr., & Kato, I. (1980) *Annu. Rev. Biochem. 49*, 593−626.
Ling, M. H., Qi, H. Y., & Chi, C. W. (1993) *J. Biol. Chem. 268*, 810−814.
Luo, M. J., Lu, W. Y., & Chi, C. W. (1997) *J. Biochem.* (in press).
Maxam, A. M., & Gilbert, W. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 560−564.
Rolka, K., Kupryszewski, G., Ragnarsson, U., Otlewski, J., Wilusz, T., & Polanowski, A. (1989) *Biol. Chem. Hoppe-Seyler 370*, 499−502.
Sambrook, J., Fritsch, E. F., & Maniatis, T. (1989) *Molecular Cloning: A Laboratory Manual*, 2nd ed., Cold Spring Harbor Laboratory Press, Plainview, NY.
Sanger, F., Nicklen, S., & Coulson, A. E. (1977) *Proc. Natl. Acad. Sci. U.S.A. 74*, 5463−5467.
Winkler, M. E., Blaber, M., Bennet, G. L., Holmes, W. E., & Vegar, G. (1985) *Bio/Technology 3*, 990−998.
Xu, W. F., Tao, W. K., Gong, Z. Z., & Chi, C. W. (1993) *J. Biochem. 113*, 153−158.
Yang, H. L., Wang, L. X., Zhu, D. X., & Chi, C. W. (1991) *Sci. Sin. 34*, 832−839.
Yang, H. L., Luo, R. S., Wang, L. X., Zhu, D. X., & Chi, C. W. (1992) *J. Biochem. 111*, 537−545.
Zhang, X. Y., & Chi, C. W. (1979) *Sci. Sin. 22*, 1443−1454.

BI962993C