Acta Crystallographica Section F
Structural Biology and Crystallization Communications
ISSN 1744-3091

Expression, purification, crystallization and preliminary X-ray diffraction analysis of *Sagittaria sagittifolia* arrowhead protease inhibitor API-A in complex with bovine trypsin

Chunhui Jiang,[a]‡ Rui Bao[a]‡ and Yuxing Chen[a,b]*

[a]Institute of Protein Research, Tongji University, Shanghai 200092, People's Republic of China, and [b]Hefei National Laboratory for Physical Sciences at Microscale and School of Life Sciences, University of Science and Technology of China, Hefei, Anhui 230026, People's Republic of China

‡ These authors contributed equally.

Correspondence e-mail: cyxing@ustc.edu.cn

Received 27 June 2008
Accepted 8 October 2008

Protease inhibitors play key roles in physiological processes. Arrowhead protease inhibitor A (API-A), a member of the serine protease inhibitor family, can inhibit two trypsin molecules simultaneously. In the present work, API-A from *Sagittaria sagittifolia* has been cloned, expressed, purified and crystallized in complex with bovine trypsin. The crystals were obtained by the sitting-drop method. A data set was collected to 2.48 Å resolution from a single crystal. The crystal belonged to space group $C222_1$, with unit-cell parameters a = 76.63, b = 110.86, c = 152.99 Å, $\alpha = \beta = \gamma = 90^\circ$.

1. Introduction

Protease inhibitors are ubiquitously present in a variety of species that includes plants and animals as well as microorganisms (Laskowski & Kato, 1980). They play vital roles in many physiological processes such as protein digestion, cell signal transmission, inflammation, apoptosis, blood clotting, embryogenesis *etc.* (Leung *et al.*, 2000). As proteases are involved in many diseases, such as cancer (Kobayashi *et al.*, 2004), nephrotoxicity (Umeki *et al.*, 1989) and AIDS (Ng *et al.*, 2003), their inhibitors have been extensively studied over the last few decades as starting points for potential therapeutic agents. One of the most studied specific protease-inhibitor families, the Kunitz-type superfamily from serine protease inhibitors, has been well characterized (Laskowski & Kato, 1980; Nielsen *et al.*, 2004; Onesti *et al.*, 1991; Schmidt *et al.*, 2005; Song & Suh, 1998; Sweet *et al.*, 1974). Members of the family usually contain 170–200 residues and one or two intrachain disulfide bonds. They have a strong affinity for their target proteases and form stable complexes that are difficult to separate into the constituent proteases and inhibitors (Baillargeon *et al.*, 1980).

A few members of this family have two active sites and can thus act on two molecules of protease simultaneously; these are the so-called 'double-headed inhibitors' (Yang *et al.*, 1992). The arrowhead protease inhibitors A and B (API-A and API-B) were first purified from the tubers of *Sagittaria sagittifolia* L. and characterized by Chang *et al.* (1979). API-A and API-B share 91% sequence homology, but their inhibitory activities are distinctly different from each other. API-A inhibits equimolar amounts of trypsin and chymotrypsin simultaneously, whereas API-B inhibits two molecules of trypsin at the same time. Both of them can inhibit pig tissue kallikrein in a 1:1 ratio (Xie *et al.*, 1997). There has been some controversy about the exact locations of the active sites (Li *et al.*, 2002; Xie *et al.*, 1997). Therefore, high-resolution crystallographic structures are necessary to answer these and other questions.

In the present work, we report the crystallization and preliminary X-ray diffraction analysis of API-A in complex with two trypsin molecules. The structure will enable us to confirm the locations of the active sites and to better understand the characteristics of API-A and its structure–function relationship.

2. Materials and methods

2.1. Cloning and purification

Total RNA was isolated from 100 mg fresh arrowhead (*S. sagittifolia* L.) buds using Trizol reagent (Invitrogen) according to the protocol of Chomczynski & Mackey (1995). The first chain of cDNA was reverse-transcribed from the total RNA using a specific primer 5′-CGCCGCGGCCGCTTAGAGTGCGTCGRACTTMTG-3′ with a *Not*I site (where R represents G/A and M is A/C) and Superscript II reverse transcriptase (Invitrogen) according to the manufacturer's protocols. The coding sequence of API-A was PCR-amplified from cDNA using the forward primer 5′-CTCGCATATGGATCCCGTCGTCGACAGC-3′ containing an *Nde*I site. The PCR product was cloned into a pET28a-derived expression vector, which gives a protein with a hexahistidine tag just after the start codon. The resulting construct was then co-transformed with PKY206 (a plasmid containing the *groESL* genes of *Escherichia coli*, resulting in the synthesis of chaperones GroEL and GroES; Mizobata *et al.*, 1992) into the *E. coli* BL21 (DE3) strain and the transformed cells were grown at 310 K in LB containing 10 µg ml^{-1} kanamycin until the OD_{600} reached about 0.6; the cells were then induced with 0.2 m*M* IPTG at 291 K overnight.

The cell pellets were resuspended in 40 ml cold lysis buffer (20 m*M* Tris–HCl pH 7.5, 100 m*M* NaCl) and lysed by sonification. The lysate was clarified by centrifugation at 16 000*g* for 30 min and the supernatant was filtered and loaded onto 2 ml nickel-agarose affinity resin (Qiagen) equilibrated with lysis buffer. After washing with ten volumes of lysis buffer plus 10 m*M* imidazole, His-tagged protein was then eluted with lysis buffer plus 200 m*M* imidazole. The fractions containing the target protein were further purified by size-exclusion chromatography on a Superdex 200 column (GE Healthcare; Fig. 1*a*). The purity of the protein was checked by SDS–PAGE (Fig. 1*b*).

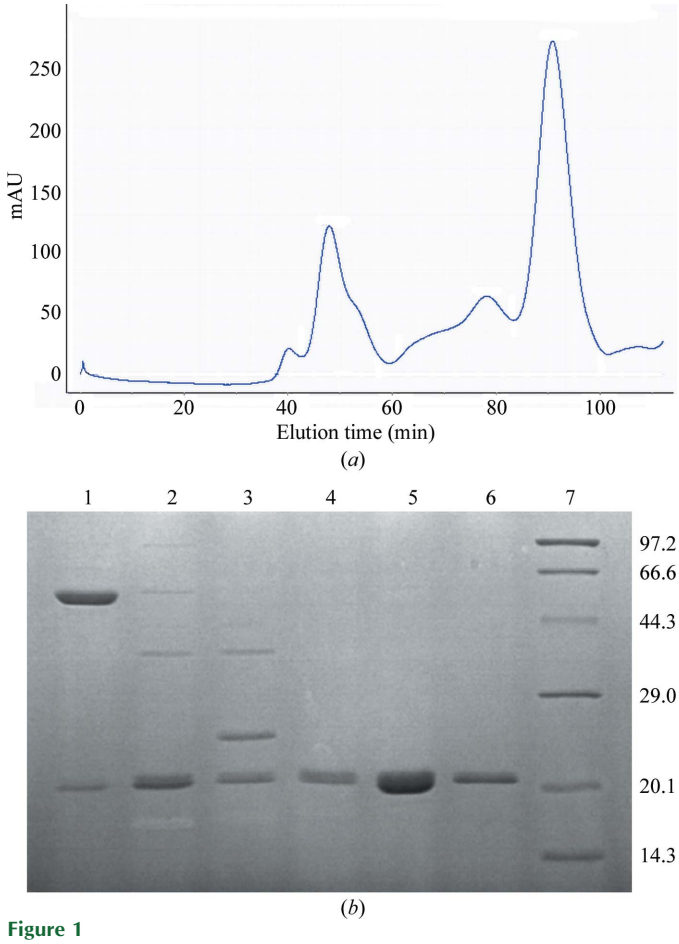


Figure 1
(*a*) Gel filtration of API-A using size-exclusion chromatography on a Superdex 200 column. The flow rate was 1 ml min^{-1}. The main peaks were collected as follows: peak I, 47.90 min; peak II, 78.21 min; peak III, 90.87 min. (*b*) 12% SDS–PAGE analysis of API-A. Lanes 2, 4 and 5, peaks I, II and III of API-A after gel filtration, respectively. Lane 7, low-molecular-weight markers (kDa).

2.2. Complex formation

For complex formation, the fractions corresponding to the highest peak, which contained about 5 mg protein, were pooled and mixed with 12 mg β-trypsin (bovine trypsin, USP grade; Ameresco) and $CaCl_2$ was then added to a final concentration of 20 m*M*. After incubation at 277 K for 2 h, the mixture was loaded onto a HiLoad 16/60 Superdex 200 (Amersham Biosciences) column equilibrated with buffer (20 m*M* Tris–HCl pH 7.5, 100 m*M* NaCl, 20 m*M* $CaCl_2$; Fig. 2). The fractions containing the API-A–trypsin complex were collected and concentrated to 15 mg ml^{-1} by ultrafiltration using an Amicon Ultra 10 kDa cutoff concentrator (Millipore) and stored at 193 K.

2.3. Crystallization experiments

Crystallization trials were carried out by the hanging-drop vapour-diffusion method at 291 K using Crystal Screens I and II (Hampton Research) for initial screening. Drops were prepared by mixing 1 µl protein solution with 1 µl precipitant solution and were equilibrated against 500 µl reservoir solution. For optimization, the sitting-drop method at 291 K was used.

2.4. Data collection

X-ray diffraction data were collected at 100 K using an in-house Rigaku MM007 X-ray generator (λ = 1.54179 Å) with a MAR Research 345 detector at the School of Life Sciences, University of Science and Technology of China (USTC, Hefei, People's Republic of China). The crystal was flash-frozen and maintained at 100 K using nitrogen gas during data collection. 20%(*v*/*v*) glycerol was added to the reservoir as a cryoprotectant. The cryoprotectant was introduced into the crystal by soaking.

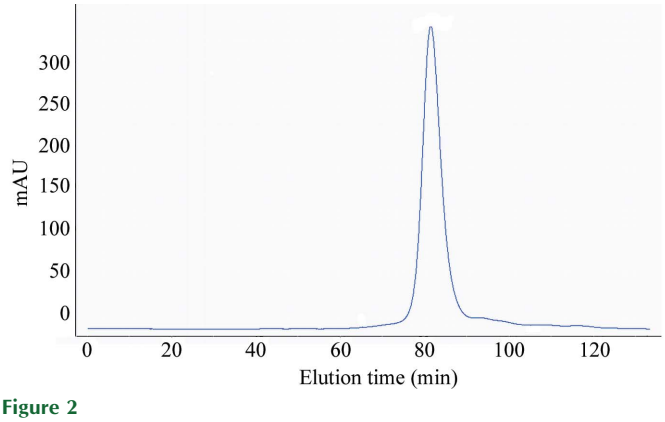
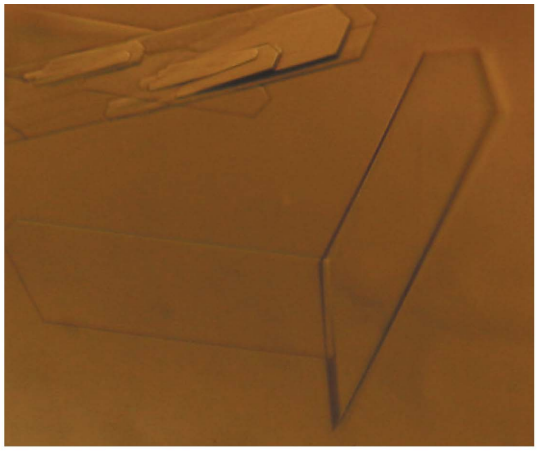
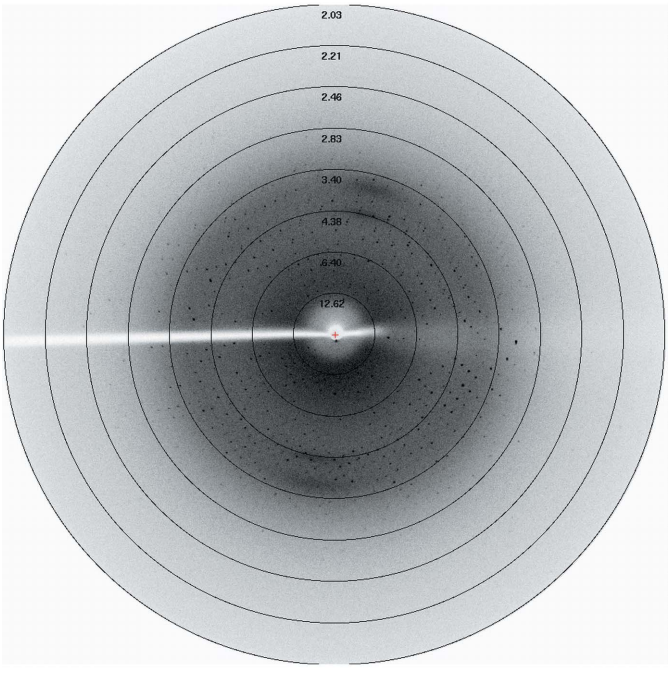


Figure 2
Gel filtration of API-A complexed with trypsin using a HiLoad 16/60 Superdex 200 column. The flow rate was 1 ml min^{-1}. The peak at 81.40 min corresponds to the complex.



(*a*)



(*b*)

Figure 3
(*a*) Crystals of *S. sagittifolia* API-A complexed with trypsin obtained by the sitting-drop vapour-diffusion method. The average dimensions of these crystals were 0.5 × 0.2 × 0.03 mm. (*b*) X-ray diffraction pattern at a resolution of 2.48 Å.

3. Results and discussion

The API-A protein was co-expressed with the chaperones GroES and GroEL in a soluble form in *E. coli* BL21 (DE3) and was purified to homogeneity after size-exclusion chromatography. After incubation with excess trypsin for 2 h, the putative complex was applied onto a HiLoad 16/60 Superdex 200 column. The peak corresponding to the target complex appeared at 81.40 min (Fig. 2), which was earlier than API-A (90.87 min; Fig. 1*a*). The elution time and symmetry of the peak indicated that API-A had formed a stable complex with trypsin. Hampton Research Crystal Screens I and II were used for preliminary screening and gave crystals from seven conditions, three of which contained ammonium sulfate and the other four different precipitants. These initial hits gave needle-shaped crystals that were not suitable for X-ray diffraction. Optimization was carried out with the sitting-drop method. Eventually, crystals that were suitable for diffraction studies and had dimensions of 0.5 × 0.2 × 0.03 mm appeared in drops with reservoirs containing 0.1 *M* sodium cacodylate pH 6.5, 0.2 *M* ammonium sulfate, 20% PEG 8000 (Fig. 3*a*). These crystals diffracted to 2.48 Å resolution (Fig. 3*b*). The diffraction data were processed with the program *AUTOMAR* v.1.2 (Bartels & Klein, 2003). These crystals belonged to space group $C222_1$, with unit-cell parameters $a = 76.63$, $b = 110.86$, $c = 152.99$ Å, $\alpha = \beta = \gamma = 90^\circ$. The data-collection statistics are listed in Table 1. Structure solution is in progress and we hope that this will enable us to characterize the binding of API-A to trypsin more precisely.

Table 1
Data-collection statistics.

Values in parentheses are for the highest resolution shell.

Nominal resolution range (Å)	30–2.48 (2.57–2.48)
Space group	$C222_1$
Temperature (K)	100
Unit-cell parameters (Å, °)	$a = 76.63$, $b = 110.86$, $c = 152.99$, $\alpha = \beta = \gamma = 90.00$
Wavelength (Å)	1.54180
Unique reflections	22790 (2161)
Completeness (%)	99.7 (99.5)
$I/\sigma(I)$	9.9 (2.1)
$R_{\rm merge}$†	12.46 (47.8)

† $R_{\rm merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $I_i(hkl)$ is the intensity of an observation and $\langle I(hkl)\rangle$ is the mean value for the unique reflection.

This work was supported by the Ministry of Education of China (Talents Project of New Century NCET-06-0374 and Program PRA B07-02) and the National Natural Science Foundation of China (Grant No. 30670461).

References

Baillargeon, M. W., Laskowski, M. Jr, Neves, D. E., Porubcan, M. A., Santini, R. E. & Markley, J. L. (1980). *Biochemistry*, **19**, 5703–5710.
Bartels, K. & Klein, C. (2003). *The AUTOMAR User's Guide*. http://www.marresearch.com/automar/.
Chang, H. Y., Chi, C. W. & Lo, C. Q. (1979). *Sci. Sin.* **22**, 1443–1454.
Chomczynski, P. & Mackey, K. (1995). *Biotechniques*, **19**, 942–945.
Kobayashi, H., Yagyu, T., Inagaki, K., Kondo, T., Suzuki, M., Kanayama, N. & Terao, T. (2004). *Cancer*, **100**, 869–877.
Laskowski, M. Jr & Kato, I. (1980). *Annu. Rev. Biochem.* **49**, 593–626.
Leung, D., Abbenante, G. & Fairlie, D. P. (2000). *J. Med. Chem.* **43**, 305–341.
Li, J., Ruan, K. C. & Chi, C. W. (2002). *Sheng Wu Hua Xue Yu Sheng Wu Wu Li Xue Bao (Shanghai)*, **34**, 662–666.
Mizobata, T., Akiyama, Y., Ito, K., Yumoto, N. & Kawata, Y. (1992). *J. Biol. Chem.* **267**, 17773–17779.
Ng, T. B., Lam, S. K. & Fong, W. P. (2003). *Biol. Chem.* **384**, 289–293.
Nielsen, P. K., Bonsager, B. C., Fukuda, K. & Svensson, B. (2004). *Biochim. Biophys. Acta*, **1696**, 157–164.
Onesti, S., Brick, P. & Blow, D. M. (1991). *J. Mol. Biol.* **217**, 153–176.
Schmidt, A. E., Chand, H. S., Cascio, D., Kisiel, W. & Bajaj, S. P. (2005). *J. Biol. Chem.* **280**, 27832–27838.
Song, H. K. & Suh, S. W. (1998). *J. Mol. Biol.* **275**, 347–363.
Sweet, R. M., Wright, H. T., Janin, J., Chothia, C. H. & Blow, D. M. (1974). *Biochemistry*, **13**, 4212–4228.
Umeki, S., Tsukiyama, K., Okimoto, N. & Soejima, R. (1989). *Am. J. Med. Sci.* **298**, 221–226.
Xie, Z. W., Luo, M. J., Xu, W. F. & Chi, C. W. (1997). *Biochemistry*, **36**, 5846–5852.
Yang, H. L., Luo, R. S., Wang, L. X., Zhu, D. X. & Chi, C. W. (1992). *J. Biochem.* **111**, 537–545.