Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Crystallization and preliminary X-ray diffraction analysis of a novel trypsin inhibitor from seeds of *Copaifera langsdorffii*

S. Krauchenco,[a] J. A. Silva,[b] R. A. P. Nagem,[a] J. R. Brandão Neto,[a] V. P. Forrer,[a] R. Carmona e Ferreira,[a] M. L. R. Macedo,[c] J. C. Novello,[b] S. Marangoni[b] and I. Polikarpov[a]*

[a]Laboratório Nacional de Luz Síncrotron, Campinas, SP, Brazil, [b]Departamento de Bioquímica, Instituto de Biologia, Unicamp, Campinas, SP, Brazil, and [c]Departamento de Ciências Naturais, CEUL, Universidade Federal de Mato Grosso do Sul, MS, Brazil

Correspondence e-mail: igor@lnls.br

Received 15 February 2001
Accepted 3 July 2001

A novel trypsin inhibitor isolated from seeds of *Copaifera langsdorffii* was purified to homogeneity and crystallized. Crystals suitable for X-ray analysis were grown using the hanging-drop vapour-diffusion method at 291 K in sodium acetate buffer at pH values near 4.3 using PEG 4000 as precipitant. The crystals presented symmetry compatible with the space group $P4_12_12$ or $P4_32_12$, with unit-cell parameters $a = b = 58.71$, $c = 93.75$ Å, and diffracted to 1.83 Å resolution at the synchrotron source.

1. Introduction

Proteinase inhibitors may be defined as proteins capable of strongly inhibiting hydrolytic enzymes both *in vitro* and *in vivo* by forming stoichiometric and stable complexes (Breddam *et al.*, 1991; Bode & Huber, 1991). They are present in multiple forms in numerous tissues and fluids of plants, animals and microorganisms (Laskowisk & Kato, 1980; Richardson, 1991). Of the inhibitors presents in plants, many are active against exogenous rather than endogenous enzymes, suggesting that they play a role in plant defence, conferring a broad spectrum of resistance to pests and pathogens (Bowles, 1990; Broadway & Duffey, 1986, 1988; Ryan, 1990; Christeller *et al.*, 1992; Shewry & Lucas, 1997; Vigers *et al.*, 1991).

Animal tests and medical experiments have shown that proteinase inhibitors of certain types are anticarcinogenic (Yavelow *et al.*, 1983; Troll *et al.*, 1987; Troll & Kennedy, 1989; Kennedy, 1994; Liener, 1995). The anticarcinogenic properties include the ability to reduce oxygen-radical formation (Yavelow *et al.*, 1982; Frenkel *et al.*, 1987), to suppress the growth of chemical-induced colon and anal gland tumours in rats (Billings *et al.*, 1990), breast tumours in rats and humans (Troll *et al.*, 1980; Tamir *et al.*, 1990) and lung tumours in mice (Witschi & Kennedy, 1989), to suppress chemical- or radiation-induced cell transformation (Billings *et al.*, 1987, 1989) and to reduce spontaneous chromosome abnormality (Afzal *et al.*, 1989). Epidemiological studies suggest that human populations which are known to have high concentration of certain proteinase inhibitors, mainly Bowman–Birk and Kunitz families, in their diet have lower rates of colon, breast, prostate and skin cancers (Correa, 1981; Kennedy, 1998; Chen *et al.*, 1992).

Proteinase inhibitors are divided according to the class of the enzyme they inhibit. Sequencing and X-ray crystallographic studies have shown that the inhibitors of serine proteinases can be further subdivided into several families. The major criteria for establishing a family are extensive homology among its members, topological relationships between the disulfide bridges and location of the reactive site (Laskowisk & Kato, 1980; Weder, 1992; Ryan, 1990; Bowles, 1990; Bode & Huber, 1992; Birk, 1985, 1994). Of the many types of trypsin inhibitors, the most important are the Kunitz and the Bowman–Birk proteinase inhibitor families. The Bowman–Birk inhibitors have molecular weights of 8–10 kDa and usually possess seven disulfide bridges. The Kunitz-type inhibitors usually have a molecular weight of about 20 kDa and only two disulfide bridges.

An unusual proteinase inhibitor forming a heterodimer of two non-covalently linked polypeptide chains with molecular weights of 11 and 9 kDa was extracted from the seeds of the *C. langsdorffii* tree (Leguminosae, Caesalpinioideae). Here, we present the results of purification, N-terminal characterization, crystallization and preliminary X-ray diffraction studies of this inhibitor.

2. Experimental

2.1. Protein extraction and purification

The whole seeds were ground in an electric mill and the heavily pigmented fragments of the tegument were separated by suspension in chloroform [1:3(*w*/*v*)]. A crude inhibitor preparation was obtained by extraction with 10 m*M* potassium phosphate buffer pH 7.6 [1:4(*w*/*v*)] for 4 h at 277 K. Centrifugation of the extracts was performed at 10 000 rev min^{-1} at 277 K for 30 min and was followed by

ammonium sulfate precipitation. The precipitate was dialyzed for 24 h at 277 K against distilled water and freeze-dried. All the following purification steps were performed at room temperature.

The precipitate was dissolved in 50 m*M* Tris–HCl buffer pH 8.0 and applied to a Sepharose DEAE column (1.6 × 19 cm) equilibrated with 50 m*M* Tris–HCl buffer pH 8.0. After 200 ml of buffer elution, a column was eluted with a concentration gradient of sodium chloride from 0 to 0.5 *M*. The flow rate was 30 ml h^{-1}, the collection volume was 3.0 ml and the protein absorption was monitored at 280 nm. The inhibitory fraction (DI) was dialyzed and freeze-dried. This fraction was applied to a Sepharose 4-B-anhydrotrypsin affinity column using 50 m*M* Tris buffer pH 8.0. The column was washed with the same buffer and the inhibitors were eluted with 1 m*M* HCl in 500 m*M* NaCl. The biological activity of the inhibitor was detected using bovine trypsin and *N*-α-benzoyl-DL-arginine *p*-nitroanilide (BAPNA) as described by Erlanger *et al.* (1961). The purity of the inhibitor was checked by SDS–PAGE (Fig. 1).

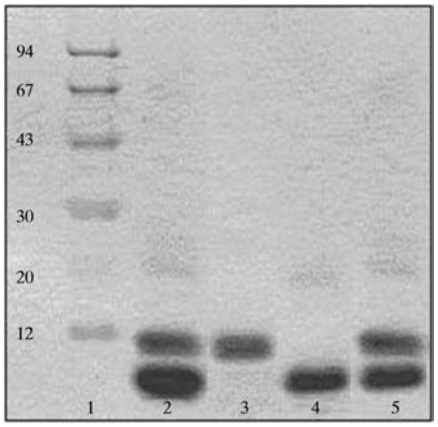

Figure 1
SDS–PAGE using a 16.5% tricine gel. Lane 1 shows the molecular-weight standards: phosphorylase (MW = 94 kDa), albumin (MW = 66 kDa), egg albumin (MW = 43 kDa), carbonic anhydrase (MW = 30 kDa), pancreas trypsin inhibitor (MW = 20 kDa) and cytochrome (MW = 12 kDa). Lane 2 shows the inhibitor extracted from *C. langsdorffii* seeds and reduced with 0.1 *M* DTT. Lanes 3 and 4 show the two non-covalently bound parts of the molecule separated with 0.1 *M* DTT and purified by HPLC using a reverse-phase column. Lane 5 shows the inhibitor from dissolved crystal.

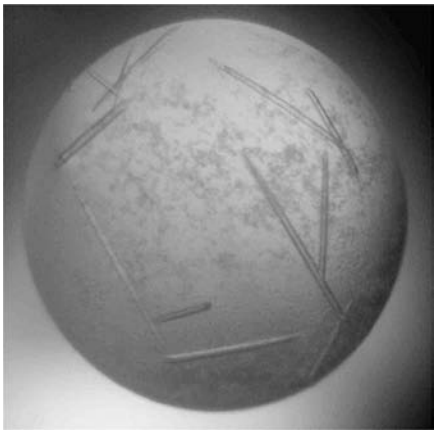

Figure 2
Crystals of *C. langsdorffii* protease inhibitor. The crystals were grown by the hanging-drop vapour-diffusion method at 291 K in 0.1 *M* sodium acetate buffer at pH values near 4 using PEG 4000 (20–25%) as precipitant.

2.2. N-terminal characterization

N-terminal sequencing was performed by Edman degradation (Edman & Begg, 1967). 40 nmol of purified inhibitor reduced with 1 *M* DTT was transferred to a PVDF membrane and applied on the automatic sequencer Procise 491 (Applied Biosystems). Two different N-terminal sequences have been derived, one of which, corresponding to 11 kDa chain, shows a high degree of similarity with the members of the Kunitz family of inhibitors (Table 1). No homology was found, however, for the 9 kDa chain N-terminal amino-acid sequence.

2.3. Crystallization and X-ray data collection

All attempts to crystallize the separate polypeptide chains of the protein were unsuccessful. The intact protein inhibitor, however, was readily crystallizable. The crystals of the whole inhibitor used for data collection were grown by hanging-drop vapour diffusion at 291 K in 0.1 *M* sodium acetate buffer at a pH near 4.3 using PEG 4000 (20–25%) as a precipitant (Fig. 2). Drops consisted of equal volumes of protein at a concentration of 10 mg ml^{-1} and reservoir solution. Small crystals grew after 4 d. The presence of both chains in the crystallized material was confirmed by SDS–PAGE gel (Fig. 1).

X-ray diffraction data was collected from crystals mounted in a rayon loop, immersed for 30 s in a cryocooling solution (20% ethylene glycol mixed with the mother liquor) and flash-cooled to 80 K in a cold nitrogen stream (Fig. 3). Data collection was performed at the Protein Crystallography beamline (Polikarpov, Oliva *et al.*, 1997; Polikarpov, Perles *et al.*, 1997) at the Brazilian National Synchrotron Light Laboratory (Campinas, SP, Brazil) using a MAR345 image plate. Data was autoindexed and integrated with the program *DENZO* (Otwinowski, 1993). Scaling and merging of data were performed with the program *SCALEPACK* (Otwinowski, 1993).

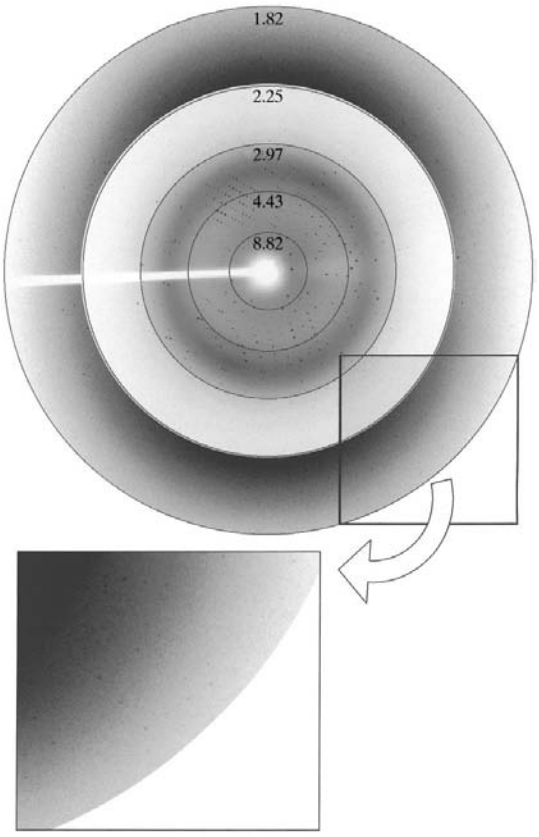

Figure 3
Diffraction pattern of the crystal. The outer area of the diffraction image (between resolution rings of 2.25 and 1.82 Å) is shown about four times more saturated than the inner part of the image. The close-up of the outer part of the image depicts reflections extending to a maximum resolution of 1.83 Å.

3. Results and discussion

The results from SDS–PAGE (Fig. 1) and N-terminal amino-acid sequencing (Table 1) indicate that the inhibitor under study has two chains, with molecular weights of approximately 11 and 9 kDa, that are non-covalently linked together to form a heterodimer of approximately 20 kDa. These two domains are separated under reducing conditions with 0.1 *M* DTT and have different N-terminal amino-acid composition.

Table 1
N-terminal amino-acid sequences of the two inhibitor polypeptide chains separated with 0.1 *M* DTT and isolated by purification on a HPLC reverse-phase column.

The table also shows part of the soybean trypsin inhibitor primary structure, which presents 50% identity with chain I of the *C. langsdorffii* trypsin inhibitor. No significant homology was found for the chain II. *, identical; :, similar; ., different.

Inhibitor chains separated	N-terminal sequence
TI chain I (11 kDa)	RLVDTNGKPIENDGAEYYILPAVR
	-::**:*.*:.*.*..******:*
STI (residues 2–23)	VLDTDGNPLRN-GGTYYILPAIR
TI chain II (9 kDa)	WQLPSVTVGNPKVSAFGGPF

Table 2
Data collection and processing statistics.

Statistical values for the highest resolution shell are shown in parentheses.

Space group	$P4_{3(1)}2_12$
Unit-cell dimensions (Å)	
$a = b$	58.86
c	93.83
Wavelength (Å)	1.535
Crystal-to-detector distance (mm)	100
Resolution range (Å)	25.3–1.83 (1.87–1.83)
No. of reflections	123604
No. of unique reflections	15024
Redundancy	4.0
$\langle I/\sigma(I)\rangle$	23.6 (3.3)
Mosaicity (°)	0.5
Multiplicity	8.2 (5.4)
No. of images collected	100
Oscillation per image (°)	1.3
Completeness (%)	99.7 (97.6)
$R_{\rm merge}$ (%)	8.9 (50.8)

† $R_{\rm merge} = \sum_{hkl} |I - \langle I\rangle| / \sum_{hkl} I$.

An X-ray diffraction data set (Table 2) was collected from a flash-frozen crystal measuring 0.1 × 0.07 × 0.03 mm using synchrotron radiation with a wavelength of 1.38 Å. Diffraction pattern showed tetragonal Laue symmetry and systematic absences indicated that the crystal is compatible with the space group $P4_12_12$ or $P4_32_12$. The small size of the crystal limited diffraction intensities and the final resolution range. In order to optimize the X-ray diffraction data collection, the synchrotron-radiation wavelength was changed to 1.535 Å (Polikarpov, Teplyakov *et al.*, 1997). A second data set was collected from a different single crystal and evaluated. Its maximum resolution range extended to 1.83 Å. The crystal belonged to the same tetragonal space group. The unit-cell parameters were determined to be $a = b = 58.71$, $c = 93.75$ Å.

The calculated cell volume is 3.25×10^5 Å^3. Assuming the molecular weight of the protein to be 20 kDa, the calculated Matthews coefficient ($V_{\rm M}$; Matthews, 1968) was 2.03 Å^3 Da^{-1}, indicating the presence of one molecule in the asymmetric unit. A number of crystallographic models of Kunitz-type trypsin inhibitors available in Protein Data Bank were tested as search models for molecular replacement. However, all attempts to find molecular-replacement solution were unsuccessful. A systematic search for heavy-atom derivatives is currently under way. The quick cryoderivatization procedure recently introduced by Dauter *et al.* (2000) will be applied.

Financial support from the Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) *via* projects 99/03387-4 and 00/02317-1 and the Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq) is acknowledged.

References

Afzal, V., Wiencke, J. K. & Wolff, S. (1989). *Carcinogenesis*, **10**, 1193–1196.
Billings, P. C., Morrow, A. R., Ryan, C. A. & Kennedy, A. R. (1989). *Carcinogenesis*, **10**, 687–691.
Billings, P. C., Newberne, P. M. & Kennedy, A. R. (1990). *Carcinogenesis*, **11**, 1083–1086.
Billings, P. C., St Clair, W., Ryan, C. A. & Kennedy, A. R. (1987). *Carcinogenesis*, **8**, 809–812.
Birk, Y. (1985). *Int. J. Pept. Protein Res.* **25**, 113–131.
Birk, Y. (1994). *Arch. Latinoam. Nutr.* **44**, 26S–30S.
Bode, W. & Huber, R. (1991). *Biomed. Biochim. Acta*, **50**, 437–446.
Bode, W. & Huber, R. (1992). *Eur. J. Biochem.* **204**, 433–451.
Bowles, D. J. (1990). *Annu. Rev. Biochem.* **59**, 873–907.
Breddam, K., Widmer, F. & Meldal, M. (1991). *Int. J. Pept. Protein Res.* **37**, 153–160.
Broadway, R. M. & Duffey, S. S. (1986). *J. Insect Physiol.* **8**, 673–680.
Broadway, R. M. & Duffey, S. S. (1988). *J. Insect Physiol.* **34**(12), 1111–1117.
Chen, P., Rose, J., Love, R., Wei, C. H. & Wang, B. (1992). *J. Biol. Chem.* **267**, 1990–1994.
Christeller, J. T., Laing, W. A., Markwick, N. P. & Burgess, E. P. J. (1992). *Insect Biochem. Mol. Biol.* **22**, 735–746.
Correa, P. (1981). *Cancer Res.* **41**, 3685–3690.
Dauter, Z., Dauter, M. & Rajashankar, K. R. (2000). *Acta Cryst.* D**56**, 232–237.
Edman, P. & Begg, G. (1967). *Eur. J. Biochem.* **1**, 80–91.
Erlanger, B. F., Kolowsky, N. & Cohen, N. (1961). *Arch. Biochem. Biophys.* **95**, 271–278.
Flavin, D. F. (1982). *Vet. Hum. Toxicol.* **24**, 25–28.
Frenkel, K., Chrzan, K., Ryan, C. A., Wiesner, R. & Troll, W. (1987). *Carcinogenesis*, **8**, 1207–1212.
Kennedy, A. R. (1994). *Cancer Res.* **54**, 1999S–2005S.
Kennedy, A. R. (1998). *Pharmacol. Ther.* **78**, 167–209.
Laskowisk, M. Jr & Kato, I. (1980). *Annu. Rev. Biochem.* **49**, 593–626.
Liener, I. E. (1995). *J. Nutr.* **125**, 744S–759S.
Lorenzi, H. (1992). *Árvores Brasileiras – Manual de Identificação e Cultivo de Plantas Nativas do Brasil.* Editora Platarum LTDA.
Matthews, B. M. (1968). *J. Mol. Biol.* **33**, 491–497.
Otwinowski, Z. (1993). *Proceedings of the CCP4 Study Weekend. Data Collection and Processing*, edited by L. Sawyer, N. Isaacs & S. Bailey, pp. 56–62. Warrington: Daresbury Laboratory.
Polikarpov, I., Oliva, G., Castellano, E. E., Garratt, R., Arruda, P., Leite, A. & Craievich, A. (1997). *Nucl. Instrum. Methods A*, **405**, 159–164.
Polikarpov, I., Perles, L. A., de Oliveira, R. T., Oliva, G., Castellano, E. E., Garratt, R. & Craievich, A. (1997). *J. Synchrotron Rad.* **5**, 72–76.
Polikarpov, I., Teplyakov, A. & Oliva, G. (1997). *Acta Cryst.* D**53**, 734–737.
Pusztai, A. (1989). *Toxicants of Plant Origin*, Vol. 2, edited by P. R. Sheek, pp. 29–71. Boca Raton, Florida: CRC Press.
Richardson, M. (1991). *Methods Plant. Biochem.* **5**, 259–305.
Ryan, C. A. (1990). *Annu. Rev. Phytopathol.* **28**, 425–449.
Shewry, P. R. & Lucas, J. A. (1997). *Adv. Bot. Res.* **26**, 135–192.
Tamir, S., Kadner, S., Katz, J. & Finlay, T. H. (1990). *Endocrinology*, **127**, 1319–1328.
Troll, W. & Kennedy, A. R. (1989). *Cancer Res.* **49**, 499- 502.
Troll, W., Wiesner, R. & Frenkel, K. (1987). *Adv. Cancer Res.* **49**, 265–283.
Troll, W., Wiesner, R., Shellabarger, C. J., Holtzman, S. & Stone, J. P. (1980). *Carcinogenesis*, **1**, 469–472.
Vigers, A. J., Roberts, W. K. & Selitrennikoff, C. P. (1991). *Mol. Plant-Microbe Interact.* **4**, 315–323.
Weder, J. K. P. (1992). *Annu. Rev. Biochem.* **16**, 239–279.
Witschi, H. & Kennedy, A. R. (1989). *Carcinogenesis*, **10**, 2275–2277.
Yavelow, J., Finlay, T. H., Kennedy, A. R. & Troll, W. (1983). *Cancer Res.* **43**, 2454–2459.
Yavelow, J., Gidlund, M. & Troll, W. (1982). *Carcinogenesis*, **3**, 135–138.