Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Preliminary crystallographic studies of EcTI, a serine proteinase inhibitor from *Enterolobium contortisiliquum* seeds

I. F. C. Batista,[a]* M. C. Nonato,[b] M. R. Bonfadini,[b] L. M. Beltramini,[c] M. L. V. Oliva,[a] M. U. Sampaio,[a] C. A. M. Sampaio[a] and R. C. Garratt[b]

[a]Departamento de Bioquímica, Universidade Federal de São Paulo (UNIFESP/EPM), Rua Três de Maio 100, CEP 04044-020 Sao Paulo SP, Brazil, [b]Laboratório de Cristalografia de Proteínas e Biologia Molecular Estrutural, Instituto de Física de São Carlos (IFSC/USP), Universidade de São Paulo, San Carlos SP, Brazil, and [c]Departamento de Biofísica do Instituto Física de São Carlos (IFSC/USP), Universidade de São Paulo, San Carlos SP, Brazil

Correspondence e-mail: bel.bioq@epm.br

Received 12 July 2000
Accepted 23 January 2001

Enterolobium contortisiliquum trypsin inhibitor (EcTI) belongs to the Kunitz family of plant inhibitors, which are widely distributed in nature, especially in plant seeds. EcTI is composed of two polypeptide chains with a total of 174 residues, homologous to other inhibitors from the same family. EcTI crystals, which were obtained with the acupuncture-gel technique, diffract to 2.0 Å resolution and belong to space group $P2_1$, with unit-cell parameters a = 37.12, b = 38.42, c = 54.08 Å, β = 98.08°. Molecular-replacement techniques using *Erythrina caffra* trypsin inhibitor (PDB code 1tie) as the search model indicate one monomer in the asymmetric unit. The secondary-structure content of EcTI was determined by circular dichroism spectroscopy, yielding values compatible with the expected topology.

1. Introduction

Protein enzyme inhibitors are distributed in a wide range of biological materials, but are especially common in plant seeds and have been studied in detail because of their application as tools in the investigation of enzyme mechanism and recently in the development of drugs for the treatment of a variety of pathologies (Oliva *et al.*, 1999; Patick & Potts, 1998). The serine protease inhibitors are probably the most intensely investigated proteins in plant seeds and are divided according to their structural properties into the Kunitz, Bowman–Birk, potato, squash and cereal superfamilies of inhibitors (Richardson, 1991; Oliva *et al.*, 2000).

The Kunitz-type family was originally named after the purification of the soybean trypsin inhibitor in the 1940s (Kunitz, 1947*a*,*b*); the structural characteristics of the inhibitor, its complex formation with trypsin, kinetic constants *etc.* have subsequently been well studied and reviewed (Blow *et al.*, 1974; Sweet *et al.*, 1974; Kim *et al.*, 1985; Song & Suh, 1998, De Meester *et al.*, 1998). The Kunitz inhibitors purified from Leguminosae seeds are proteins characterized by a molecular weight of around 20 kDa and are composed of approximately 180 amino-acid residues, including four cysteine residues which form two disulfide bridges. However, the trypsin inhibitors purified from the Mimosoideae subfamily consist of two polypeptide chains linked by one disulfide bridge, a peculiarity which is not shared by the inhibitors from the Caesalpinoideae and Papilionoideae subfamilies (Richardson, 1991).

The overall sequence similarity between all members of the family strongly suggests that the origin of the two chains is a post-translational proteolytic cleavage (Creighton & Darby, 1989).

E. contortisiliquum is a member of the Leguminosae, belonging to the Mimosoideae subfamily, and is a native species found in South America. It is popularly known as 'black ear', owing to the shape of its fruits (Sánchez de Lorenzo Cácere, 1998–2000). Several proteins have been purified from *E. contortisiliquum* seeds, including enzymes, storage proteins and protease inhibitors (Oliva *et al.*, 1988; Sousa & Morhy, 1989; Silva *et al.*, 1994).

Among the inhibitors, a member of the plant Kunitz inhibitor family, the *E. contortisiliquum* trypsin inhibitor (EcTI), has been previously purified and sequenced (Batista *et al.*, 1996). The aim of this report is to describe for the first time preliminary crystallographic and structural studies of a member of the plant Kunitz trypsin inhibitors which is composed of two polypeptide chains, *i.e.* EcTI. The three-dimensional structure of this inhibitor is expected to shed light on its important properties which include the inhibition of blood clotting and of fibrinolytic enzymes such as factor XIIa and plasmin, properties which are not shared by other members of the same family of known structure.

2. Experimental

2.1. Purification

The purification of EcTI was performed as described previously (Batista *et al.*, 1996). The seed proteins were extracted by swelling 30 g of the cotyledons with 300 ml 0.15 *M* NaCl. After homogenization, the material was centrifuged at 3000*g* and the supernatant was

taken as the starting material (crude extract), which was initially precipitated with 80%(*v*/*v*) acetone. The resulting precipitate was redissolved in 50 m*M* Tris–HCl buffer pH 8.0 and applied to a DEAE-Sepharose ion-exchange chromatography column which was equilibrated in the same buffer and eluted using a linear gradient of NaCl from 0 to 0.5 *M*, EcTI being eluted at a concentration of approximately 0.15 *M* NaCl. Protein elution was followed by measuring the absorbance at 280 nm and the inhibitory activity was assayed by bovine trypsin inhibition using benzoyl-argininyl-*p*-nitroanilide as substrate. The fractions showing inhibitory activity were pooled and applied to a Superdex G75 gel-filtration column equilibrated with 0.1 *M* Tris–HCl buffer pH 8.0. The protein and inhibitory activity of the eluted fractions were measured as described above and the partially purified EcTI thus obtained was further chromatographed on a Resource Q column coupled to an FPLC. In this step, the column was equilibrated with 50 m*M* Tris–HCl buffer pH 8.0 and EcTI was eluted using a linear gradient of NaCl (0–0.5 *M*). The homogeneity of the final preparation and intermediate stages were checked by SDS–PAGE (Laemmli, 1970) and isoelectric focusing using the PhastSystem equipment from Pharmacia Biotech (Heukeshoven & Dernick, 1988).

2.2. Circular dichroism

The CD spectra of EcTI were recorded on a Jasco J-720 spectropolarimeter (wavelength range 190–300 nm) in quartz cuvettes of 1 mm path length. The mdegree ellipticity obtained was transformed into molar ellipticity by the *Selcon-1* program using the mean weight residue and concentration prior to secondary-structure analysis (Sreerama & Woody, 1993).

2.3. Crystallization and X-ray data collection

EcTI was dialyzed against 0.02 *M* Tris–HCl buffer pH 8.0 and then concentrated using a Centricon YM-3 membrane (Millipore) to 7.5 mg ml^{-1}. The protein concentration was determined by measuring the sample absorbance at 280 nm using the extinction coefficient of 1.0 for a 1.0 mg ml^{-1} solution of EcTI (Janson & Rydén, 1998) The solution obtained was used to perform crystallization screens with the Hampton Research Crystal Screen (Crystal Screen I) factorial kits employing the hanging-drop method using multiwell plates at a temperature of 292 K (McPherson, 1982).

Crystals grew as clusters in the drops containing factorial solution #50 (Crystal Screen 1): 15% PEG 8000, 0.5 *M* lithium sulfate. This was the only factorial solution in which promising crystals were obtained. In order to optimize the initial conditions in an attempt to obtain single crystals suitable for X-ray diffraction, a cassette for the acupuncture-gel technique was prepared (Moreno & Soriano-Garcia, 1999). Under these conditions single crystals of EcTI (Fig. 1) were obtained after three weeks at 292 K using an identical precipitant solution to that used in the hanging drops.

The X-ray data were collected on a MAR Research 345 imaging plate at the Protein Crystallography (PCr) beamline at the Laboratório Nacional de Luz Síncrotron (LNLS), Campinas, Brazil (Polikarpov *et al.*, 1998). A single crystal was transferred to mother liquor enriched with 25%(*v*/*v*) glycerol and frozen at 100 K in a nitrogen-gas stream. The image plate was operated in the 300 mm scanning mode and the crystal-to-detector distance was set to 90 mm. Even though observable diffraction extended to 1.7 Å resolution, acceptable data statistics were only obtained to 1.96 Å (Table 1).

2.4. Data processing

All data were processed and scaled using the programs *DENZO* and *SCALEPACK* (Otwinowski & Minor, 1997). Data reduction was carried out to 1.96 Å resolution. Owing to poor statistical quality, higher resolution data were not used. Molecular replacement using the atomic coordinates of the *Erythrina caffra* trypsin inhibitor as a search model was performed with the program *AMoRe* (Navaza & Saludjian, 1997). For this purpose, non-conserved residues of the *Er. caffra* structure were reduced to alanines. Data were used in the resolution range 10–3 Å and the potential solutions were optimized by rigid-body refinement.

3. Results and discussion

The inhibitor from *E. contortisiliquum* seeds, EcTI, consists of 174 amino-acid residues with two disulfide bridges and is homologous to the Kunitz trypsin inhibitor from *Er. caffra*, whose three-dimensional structure has been previously determined (PDB code 1tie; Onesti *et al.*, 1991). However, the latter is composed of only a single polypeptide chain and shares only 39% sequence identity with EcTI.

The CD spectrum of EcTI was characterized by a minimum at 213 nm. Deconvolution of the spectrum using the *Selcon-1* program yielded estimates of 59% β-sheet and 41% irregular structures. These values are compatible with those for the single-chain inhibitor from *Er. caffra*, which is composed of 39% β-sheet, 57% irregular structure and 4% 3_{10} helix as assigned using the program *PROMOTIF* (Hutchinson & Thornton, 1996). Several factors may be responsible for the small discrepancy between the two sets of values, which include the relatively large error in determining secondary-structure content by deconvolution of CD spectra, the low overall sequence identity between the two molecules and the fact that EcTI has suffered an internal proteolytic cleavage.

EcTI crystals belong to space group $P2_1$, as determined from systematic absences along b^*, with unit-cell parameters $a = 37.12$, $b = 38.42$, $c = 54.08$ Å, $\beta = 98.08$, and diffract to 1.7 Å resolution. The structure determination *via* molecular-replacement techniques using the inhibitor from *Er. caffra* as a search model indicates the presence of one monomer in the asymmetric unit. Employing the known molecular weight of 19 458 Da (as determined from the amino-acid sequence), this corresponds to a Matthews coefficient of 1.96 Å^3 Da^{-1} and a solvent content of 36.8%, which are both within the

Table 1
Data-collection and processing statistics.

Values in parentheses refer to the last shell (2.01–1.96 Å).

Space group	$P2_1$
Unit-cell parameters (Å,°)	$a = 37.12$, $b = 38.42$, $c = 54.08$, $\beta = 98.08$
Resolution range (Å)	20–1.96
No. of collected reflections	23475
No. of unique reflections	11611
R_{sym}† (%)	10.3 (31.1)
Completeness (%)	87.0 (87.0)
Percentage $I/\sigma(I) > 2$ (%)	80.0 (48.5)

† $R_{sym} = \sum_h \sum_i |I(h)_i - \langle I(h)\rangle| / \sum_h \sum_i I(h)_i$, where $I(h)$ is the intensity of the reflection h, $\sum_h$ is the sum over all reflections and $\sum_i$ is the sum over all i measurements of the reflection h.

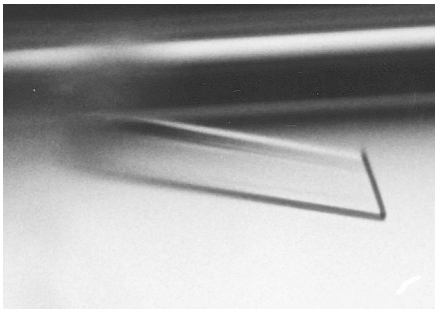

Figure 1
Crystal of EcTI.

expected normal ranges. Assuming two monomers per asymmetric unit leads to an unrealistic value of 0.98 Å^3 Da^{-1} for V_M and a negative solvent content. The best molecular-replacement solution (α = 63.2, β = 9.5, γ = 141.0°, x = 0.3999, y = 0.0000, z = 0.2396) yielded a correlation coefficient of 0.33, 4% above the subsequent potential solutions which were all clustered around 0.29. The relatively low correlation coefficient observed for the top solution is probably a consequence of the low overall sequence identity and the potential (but unknown) structural consequences of the internal cleavage of the polypeptide chain in EcTI.

The similarity between the EcTI and 1tie sequences suggests that their three-dimensional structures are similar, consisting of 12 antiparallel β-strands joined by long loops. The scissile bond (Arg63–Ile64) is expected to be located on an unconstrained external loop that protrudes from the surface of the molecule. However, whilst 1tie is known to inhibit t-PA, plasmin and trypsin, EcTI also shows high affinity for human plasma kallikrein (K_i = 6.15 nM) and factor XIIa (81.81 nM). The full refinement of its structure is expected to contribute to understanding these differences.

References

Batista, I. F. C., Oliva, M. L. V., Araujo, M. S., Sampaio, M. U., Richardson, M., Fritz, H. & Sampaio, C. A. M. (1996). *Phytochemistry*, **41**(4), 1017–1022.

Blow, D. M., Janin, J. & Sweet, R. M. (1974). *Nature (London)*, **249**, 54–57.

Creighton, T. E. & Darby, N. J. (1989). *Trends Biochem. Sci.* **14**(8), 319–324.

De Meester, P., Brick, P., Lloyd, L. F., Blow, D. M. & Onesti, S. (1998). *Acta Cryst.* D**54**, 589–597.

Heukeshoven, J. & Dernick, R. (1988). *Electrophoresis*, **9**(1), 28–32.

Hutchinson, E. G. & Thornton, J. M. (1996). *Protein Sci.* **5**(2), 212–220.

Janson, J. C. & Rydén, L. (1998). Editors. *Protein Purification: Principles, High Resolution Methods and Applications*, 2nd ed., pp. 28–30. New York: Wiley–VCH.

Kim, S.-H., Hara, S., Hase, S., Ikenaka, T., Toda, H., Kitamura, K. & Kaizuma, N. (1985). *J. Biochem. (Tokyo)*, **98**, 435–448.

Kunitz, M. (1947*a*). *J. Gen. Physiol.* **30**, 291–310.

Kunitz, M. (1947*b*). *J. Gen. Physiol.* **30**, 311–320.

Laemmli, U. K. (1970). *Nature (London)*, **227**(259), 680–685.

McPherson, A. (1982). *Preparation and Analysis of Protein Crystals*. New York: John Wiley.

Moreno, A. & Soriano-Garcia, M. (1999). *Acta Cryst.* D**55**, 577–580.

Navaza, J. & Saludjian, P. (1997). *Methods Enzymol.* **276**, 581–594.

Oliva, M. L. V., Andrade, S. A., Batista, I. F. C., Sampaio, M. U., Juliano, M., Fritz, H., Auerswald, E. A. & Sampaio, C. A. M. (1999). *Immunopharmacology*, **45**, 145–149.

Oliva, M. L. V., Sampaio, M. U. & Sampaio, C. A. M. (1988). *Biol. Chem. Hoppe-Seyler*, **369**, 229–232.

Oliva, M. L. V., Souza-Pinto, J. C., Batista, I. F. C., Araujo, M. S., Silveira, V. L. F., Auerswald, E. A., Mentele, R., Eckerskorn, C., Sampaio, M. U. & Sampaio, C. A. M. (2000). *Biochim. Biophys. Acta*, **1477**, 64–74.

Onesti, S., Brick, P. & Blow, D. M. (1991). *J. Mol. Biol.* **217**(1), 153–176.

Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.

Patick, A. K. & Potts, K. E. (1998). *Clin. Microbiol. Rev*. **11**, 614–627.

Polikarpov, I., Perles, L. A., de Oliveira, R. T., Oliva, G., Castellano, E. E., Garratt, R. C. & Craievich, A. (1998). *J. Synchrotron Rad.* **5**, 72–76.

Richardson, M. (1991). *Methods Plant Biochem.* **5**, 259–305.

Sánchez de Lorenzo Cácere, J. M. (1998–2000). *Enterolobuim contortisiliquum.* http://www.guiaverde.com/arboles/Enterolobiumcontortisiliquum.htm.

Silva, G. A., Oliva, M. L. V., Sampaio, M. U., Araujo, M. S. & Sampaio, C. A. M. (1994). *Braz. J. Med. Biol. Res.* **27**(6), 1299–310.

Song, H. K. & Suh, S. W. (1998). *Acta Cryst.* D**54**, 441–443.

Sousa, M. V. & Morhy, L. (1989). *Ann. Acad. Bras. Cienc.* **61**(4), 405–412.

Sreerama, N. & Woody, R. W. (1993). *Anal. Biochem.* **209**, 32–34.

Sweet, R. M., Wright, H. T., Janin, J., Chothia, C. H. & Blow, D. M. (1974). *Biochemistry*, **13**, 4212–4228.