Crystallization and preliminary X-ray analysis of four cysteine proteases from *Ficus carica* latex

Sarah Haesaerts,^{a,b} John Alexander Rodriguez Buitrago,^c Remy Loris,^{a,b} Danielle Baeyens-Volant^d and Mohamed Azarkan^{d*}

^aStructural Biology Brussels, Vrije Universiteit Brussel, Pleinlaan 2, B-1050 Brussels, Belgium, ^bStructural Biology Research Center, VIB, Pleinlaan 2, B-1050 Brussels, Belgium, ^cLaboratoire de Bactériologie Moléculaire, Faculté de Médecine, Université Libre De Bruxelles, Route de Lennik 808, B-1070 Brussels, Belgium, and ^dProtein Chemistry Unit, Faculté de Médecine, Université Libre De Bruxelles, Route de Lennik 808, B-1070 Brussels, Belgium. *Correspondence e-mail: mazarkan@ulb.ac.be

Received 30 January 2015

Accepted 11 March 2015

Edited by N. Sträter, University of Leipzig, Germany

Keywords: ficin; cysteine protease; papain; *Ficus carica*.

The latex of the common fig (*Ficus carica*) contains a mixture of at least five cysteine proteases commonly known as ficins (EC 3.4.22.3). Four of these proteases were purified to homogeneity and crystals were obtained in a variety of conditions. The four ficin (iso)forms appear in ten different crystal forms. All diffracted to better than 2.10 Å resolution and for each form at least one crystal form diffracted to 1.60 Å resolution or higher. Ficin (iso)forms B and C share a common crystal form, suggesting close sequence and structural similarity. The latter diffracted to a resolution of 1.20 Å and belonged to space group $P3_121$ or $P3_221$, with unit-cell parameters $a = b = 88.9$, $c = 55.9$ Å.

1. Introduction

Proteases are vital biomolecules that control the turnover, localization and activity of a variety of proteins, regulate protein–protein interactions, contribute to the processing of cellular information, customize new bioactive (macro)-molecules and also produce, transduce and amplify molecular signals. Proteases thus influence DNA transcription and replication, cell differentiation, tissue morphogenesis and remodelling, fertilization, wound healing, inflammation, immunity, autophagy and apoptosis. Alterations in proteolytic machinery thus trigger many pathological situations such as inflammatory processes, cancer, neurodegenerative disorders and cardiovascular diseases. Many proteases are therefore potential drug targets (Turk, 2006) and constitute a major focus of attention for the pharmaceutical industry. In plants, proteases are involved in xylem formation, seed maturation, nutrient supply mobilization, degradation of misfolded proteins, responses to environmental stimuli and programmed cell death (Huffaker, 1990; van der Hoorn, 2008). They have been isolated from fruits, stems, seeds and latex (Boller, 1986). The activities of latex in plant defence and in alleviating various diseases have been attributed to the presence of several hydrolytic enzymes, among which proteases play a crucial role. Most of the proteases found in the latex of different plants belong to the cysteine and serine protease families (Singh *et al.*, 2008; El Moussaoui *et al.*, 2001; Azarkan *et al.*, 2004, 2011; Torres *et al.*, 2013; Ramos *et al.*, 2013; Yariswamy *et al.*, 2013). Why latex contains high amounts of such nonspecific proteases still remains puzzling. There is increasing evidence, however, that they participate in defence mechanisms by protecting ripening fruits against plant

pathogens (Baker & Drenth, 1987), which they can attack directly (Konno *et al.*, 2004). *In vitro* experiments clearly demonstrated that cysteine proteases act by degrading the protective cuticle of nematodes (Steppek *et al.*, 2004, 2005) or the peritrophic matrix of lepidopteran larvae (Pechan *et al.*, 2002; Mohan *et al.*, 2008). It has also been demonstrated that artificial diets supplied with cysteine proteases, such as papain from papaya latex and ficins from common fig latex, are toxic to silkworm larvae at the concentrations occurring in latex (Konno *et al.*, 2004). Cysteine proteases are also required for latex coagulation upon biotic or abiotic injuries (Azarkan *et al.*, 2004). This wound-healing property is probably the most efficient plant protection mechanism against the entry and further spread of pathogens (El Moussaoui *et al.*, 2001). All of these findings are in favour of a common defensive role of latex proteases, suggesting that these enzymes possess other functions than the simple digestive role that they are often assumed to play. The rapid action required for a defensive role may thus explain the lack of specificity and high abundance of cysteine proteases in latex.

The latex of *Ficus carica* contains a mixture of cysteine proteases that belong to the papain family (family C1, clan CA) and are known under the general term ficins (EC 3.4.22.3). We recently purified to homogeneity and characterized five ficin isoforms, named ficins A, B, C, D1 and D2. All ficin (iso)forms are not glycosylated and have similar molecular masses (Azarkan *et al.*, 2011). Ficins constitute attractive tools in many research areas. Two partially purified ficin forms have been reported to activate human coagulation factor X, probably underlying the haemostatic potency of fig latex proteases (Richter *et al.*, 2002). When compared with other cysteine proteases such as papain, bromelain and the aspartic protease pepsin, ficin as a mixture of all forms was shown to specifically cleave mouse IgG1, giving better yields and immunoreactivity of fragments suitable for clinical use (Milenic *et al.*, 1989; Mariant *et al.*, 1991). We previously demonstrated that the purified ficin (iso)forms display different specific amidase activities against small synthetic substrates (Azarkan *et al.*, 2011), as well as against natural proteinaceous substrates (our unpublished results), suggesting differences in their active-site organization. Interestingly, when tested separately, purified ficin (iso)forms also show different *in vitro* selective antitumour activities, with no significant cytotoxicity against normal cell lines (results to be published). To obtain further insight into the molecular basis of their differing biological and enzymological properties, a detailed structural characterization of the purified ficins, for which no sequence data are available so far, is required. In the present work, we report the crystallization and preliminary X-ray analysis of ficins A, B, C and D2.

2. Materials and methods

2.1. Purification of ficins

The ficin (iso)forms were purified as described previously (Azarkan *et al.*, 2011). Briefly, the latex was suspended in

100 mM sodium acetate buffer containing 1 mM EDTA at pH 5.0 and stirred for 1 h at room temperature before being subjected to centrifugation (40 000g, 4°C, 30 min). The supernatant constituting the mixture of ficins was applied onto an SP-Sepharose Fast Flow (13 × 2.5 cm internal diameter; GE Healthcare) home-made column pre-equilibrated with 100 mM sodium acetate buffer pH 5.0. The unbound material was washed away with ten column volumes of the pre-equilibrating buffer and elution of the bound proteins was performed with a linear concentration gradient from 100 to 800 mM sodium acetate buffer pH 5.0 (total volume 1000 ml, 60 ml h⁻¹). The chromatographic fractions were assayed for amidase activity using DL-BAPNA as a substrate. Each test tube (total volume 2.0 ml) contained 5.0% DMSO, 1.0 mM DL-BAPNA, 1.0 mM EDTA and 100 µl sample in a buffer containing sodium citrate, borate and phosphate (100 mM each) at pH 6.8 (Azarkan, Wintjens *et al.*, 1996). The sample was pre-incubated for 15.0 min in the buffer at 37.0°C in the presence or absence of DTT before the reaction was started by addition of the substrate. The reaction proceeded at 37.0°C

Figure 1

(a) Fractionation on SP-Sepharose Fast Flow of the whole soluble fraction of *F. carica* latex protein (dashed line, Na⁺ concentration; filled circles, absorbance at 280 nm; empty circles, amidase activity measured with DL-BAPNA in the presence of DTT). (b) Purification of pegylthio-ficin B (taken as an example) on SP-Sepharose Fast Flow (dashed line, Na⁺ concentration; filled circles, absorbance at 280 nm; empty circles, amidase activity measured with DL-BAPNA in the presence of DTT). Inset, silver-stained SDS-PAGE of *S*-methylthio-ficin B obtained from its *S*-pegylthio adduct, taken as an example. *S*-Methylthio derivatives of ficins A, C and D2 displayed the same purity (not shown) as for *S*-methylthio-ficin B.

and was stopped by the addition of 500 μl 50.0% acetic acid. The release of *p*-nitroaniline was measured spectrophotometrically using $\varepsilon_{410} = 8800 \text{ M}^{-1} \text{ cm}^{-1}$ (Erlanger *et al.*, 1961). One unit of activity (nkat) is the amount of protease that is able to hydrolyze one nanomole of substrate per second under the abovementioned conditions. The chromatographic fractions constituting ficin forms A, B, C and D (indicated by solid bars in Fig. 1*a*) were subjected to chemical derivatization (S-pegylation) with a thiol-specific polyethylene glycol derivative, methoxy polyethylene glycol *ortho*-pyridine disulfide (mPEG-OPSS 5kDa; Shearwater Polymers Inc., USA) to prevent the irreversible oxidation of thiol groups and to further improve their purification. Cysteine proteases can undergo oxidation (for example by air), leading to irreversible conversion of the catalytic thiol groups to $-\text{SO}_x\text{H}$, with $x = 1, 2$ or 3 . The oxidized forms may constitute up to 50% of the enzymatic material (Azarkan, Maes *et al.*, 1996; Azarkan, Wintjens *et al.*, 1996). Furthermore, the chromatographic behaviour of these oxidized forms and the reactive forms is exactly the same on ion-exchange supports. On the contrary, thiol-pegylated proteases behaved differently on the ion-exchanger, allowing their efficient separation from the irreversibly oxidized protein material. To perform thiol-pegylation, each pool was activated in the presence of 5 mM DTT for 20 min at room temperature under stirring and loaded onto an SP-Sepharose Fast Flow column (12.5 $\times$ 1.5 cm internal diameter) pre-equilibrated with ten column volumes of 100 mM sodium acetate buffer pH 5.0. The column was exhaustively washed with the same buffer to remove excess DTT. The bound proteins were eluted with 1 M sodium acetate buffer pH 5.0 and directly collected in solid mPEG-OPSS. The reaction of mPEG-OPSS with activated ficins is presented in Fig. 2. The reaction was complete when the liberated chromophore, 2-thiopyridone, reached a stable value at 343 nm. The reaction mixture, containing S-pegylated and underivatized (irreversibly oxidized) ficin forms, was exhaustively dialyzed against water (membrane cutoff 3.5 kDa) at 4°C before being subjected to fractionation on a SP-Sepharose Fast Flow column following the same experimental conditions as for the first chromatographic step. This allowed separation of the thiol-pegylated (active; peak I in Fig. 1*b*) ficins from underivatized (irreversibly oxidized; peak II in Fig. 1*b*) forms. It is interesting to note that the flow-through fraction corresponds to mPEG-OPSS in excess which could not be eliminated after the dialysis step.

To carry out the crystallization screening, for all ficins, the $-\text{SH}$ group of the essential catalytic cysteine was converted from the pegylated form to its $-\text{SS}-\text{CH}_3$ adduct. The reaction was carried out in 100 mM sodium acetate buffer containing 2.5 mM DTT at pH 5.0 for 20 min at room temperature. Methylmethanethiolsulfonate (MMTS; 20 mM) was then added to the reaction mixture and the reaction was left for 1 h under stirring. The resulting mixture was applied onto an SP-Sepharose Fast Flow (12.5 $\times$ 1.5 cm internal diameter) column pre-equilibrated with 100 mM sodium acetate buffer pH 5.0. The column was washed with the same buffer to eliminate the released mPEG moiety as well as MMTS and DTT until the

absorbance at 280 nm reached the baseline. Bound proteins were eluted with a linear gradient from 100 to 800 mM sodium acetate buffer pH 5.0 (total volume 1000 ml; 44.5 ml h^{-1}). The purity, which was >95% as estimated by silver staining, and homogeneity of the different ficin forms were assessed by SDS-PAGE (inset in Fig. 1*b*, in which ficin B is taken as an example) and mass-spectrometric analyses. Moreover, the homogeneity of the ficins was checked by titrating the thiol function of their catalytic cysteine residues using 2,2'-dipyridyldisulfide (2PDs) as a titrant. Upon reaction with a thiol-containing polypeptide, this compound is able to form a mixed disulfide bond and to simultaneously release 2-thiopyridone. Spectrophotometric quantification of 2-thiopyridone at 343 nm provided an easy way to determine the number of free thiol groups per mole of protein (Brocklehurst & Little, 1973). The fractions corresponding to the different ficins were pooled and concentrated to the desired volume by ultrafiltration (Amicon system; membrane cutoff 3.0 kDa) and on a 5000 MWCO Vivaspin 15R concentrator. For all samples, the buffer was exchanged with water after exhaustive washing on the Vivaspin device. The protein samples were kept at -20°C until use.

2.2. Crystallization

Crystallization conditions for the different ficin forms were manually screened by the hanging-drop method using Crystal Screen and Crystal Screen 2 from Hampton Research. In all cases, drops were prepared by mixing 0.5 μl protein solution

Figure 2
Ficin purification protocol after chemical derivatization with mPEG-OPSS.

(5–10 mg ml⁻¹ in water, with 5 mg ml⁻¹ used in the initial screening) and 0.5 µl well solution and were equilibrated against 250 µl well solution. Crystals of ficin A were optimized by varying the protein concentration as well as the pH and by making use of Additive Screen from Hampton Research (in each case 0.1 µl additive solution was mixed with 0.4 µl protein solution and 0.5 µl well solution). For some hits for ficins B and C, optimization was carried out by varying the pH, protein concentration and precipitant concentration.

2.3. Data collection and processing

All crystals were mounted in CryoLoops (Hampton Research) and flash-cooled in liquid nitrogen. When required,

glycerol was added as a cryoprotectant. Crystals were transferred to the final glycerol concentration in steps of 5%(v/v) glycerol while keeping the buffer, salt and precipitant concentrations constant. The glycerol concentration was chosen as the minimum concentration required to avoid ice rings. All data were collected on the PROXIMA1 beamline at the SOLEIL synchrotron, Saint-Aubin, France using an ADSC Q315r CCD detector. For ficin B crystal forms I and II and ficin C crystal form I, a high-resolution pass and a low-resolution pass were combined to deal with overloaded reflections at low resolution. Typically, 0.5° frames were collected, although at low resolution 1.0° passes were also used. All data for ficins A, B and D2 were indexed, scaled and

Figure 3
(a) The crystal of ficin A used for data collection is indicated by an arrow. (b) Diffraction pattern of the single crystal indicated by the arrow in (a). (c)–(g) Crystal forms I–V of ficin B. (h)–(j) Crystal forms I–III of ficin C. (k) Crystals of ficin D2 forming a cluster of crystals corresponding to the conditions for crystal form I in Table 1. This cluster was broken with a needle and a more or less single-looking chunk was used to collect the data. (l) Crystal form II of ficin D2 obtained using the conditions described in Table 1. Despite their appearance, these crystals provided high-quality diffraction.

Table 1

Summary of data-collection statistics for ficins A, C and D2.

Values in parentheses are for the outermost resolution shell.

	Ficin A	Ficin C			Ficin D2	
		Form I	Form II	Form III	Form I	Form II
Reservoir solution	1.4 M ammonium phosphate, 0.1 M Tris-HCl pH 8.5	30% PEG 400, 0.1 M HEPES pH 7.5, 0.2 M MgCl ₂	30% PEG 8000, 0.1 M sodium cacodylate, 0.2 M ammonium sulfate	30% PEG 5000, 0.1 M HEPES pH 6.5, 0.2 M ammonium sulfate (Crystal Screen 2, No. 26)	2.0 M ammonium sulfate, 0.1 M sodium acetate pH 4.6 (Crystal Screen, No. 47)	2.0 M ammonium sulfate, 0.1 M HEPES pH 7.5, 2% PEG 400 (Crystal Screen 2, No. 14)
Cryoprotectant	20%(v/v) glycerol	None	None	None	17.5%(v/v) glycerol	17.5%(v/v) glycerol
Wavelength (Å)	0.9801	0.9801	0.9801	0.9801	0.9801	0.9801
Temperature (K)	100	100	100	100	100	100
Detector	ADSC Q315 CCD	ADSC Q315 CCD	ADSC Q315 CCD	ADSC Q315 CCD	ADSC Q315 CCD	ADSC Q315 CCD
Crystal-to-detector distance (mm)	256	156	256	326	171	220
Rotation range per image (°)	0.5	0.5	0.5	1.0	0.5	1.0
Total rotation range (°)	180	120	186	134	225	360
Exposure time per image (s)	0.8	0.8	0.8	0.8	0.8	0.8
Resolution range (Å)	36.30–1.60 (1.65–1.60)	31.80–1.18 (1.21–1.18)	43.60–1.90 (1.96–1.90)	49.70–2.10 (2.16–2.10)	37.00–1.45 (1.50–1.45)	42.70–1.90 (1.99–1.90)
Space group	C2	P ₃ ₁ 21 or P ₃ ₂ 21	P ₄ ₁ 2 ₁ 2 or P ₄ ₃ 2 ₁ 2	P ₂ ₁	C2	P1
Unit-cell parameters						
<i>a</i> (Å)	124.65	88.94	52.47	55.97	71.02	47.03
<i>b</i> (Å)	52.84	88.94	52.47	90.15	52.41	71.80
<i>c</i> (Å)	96.20	56.18	157.23	65.67	64.36	83.50
α (°)	90.00	90.00	90.00	90.00	90.00	66.37
β (°)	125.09	90.00	90.00	111.84	101.34	74.06
γ (°)	90.00	120.00	90.00	90.00	90.00	83.45
Observed reflections	577916 (20094)	619753 (41219)	235033 (19408)	133156 (7348)	179057 (13116)	267854 (38492)
Unique reflections	63065 (5064)	81464 (5646)	18158 (1460)	34396 (2362)	38225 (2976)	71715 (10368)
Multiplicity	9.20 (3.90)	7.60 (7.30)	12.90 (13.30)	3.90 (3.10)	4.70 (4.40)	3.70 (3.70)
$\langle I/\sigma(I) \rangle$	17.10 (3.10)	9.90 (3.50)	18.50 (4.60)	14.90 (3.30)	11.10 (3.40)	5.30 (2.10)
Completeness (%)	92.00 (74.00)	96.40 (93.60)	99.90 (100.00)	97.50 (81.00)	91.98 (72.00)	91.60 (89.20)
Mosaicity (°)	0.37	0.31	0.70	0.45	1.15	1.12
Wilson <i>B</i> factor (Å ²)	18.08	8.97	22.02	24.50	13.11	18.60
$R_{\text{meas}}^{\dagger}$	0.098 (0.520)	0.09 (0.33)	0.10 (0.60)	0.10 (0.28)	0.100 (0.37)	0.100 (0.35)
Solvent content (%)	53	44	42	41	48	51
No. of molecules in asymmetric unit	2	1	1	2	1	4

 $\dagger R_{\text{meas}} = \sum_{hkl} \{N(hkl)/[N(hkl) - 1]\}^{1/2} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$ (Diederichs & Karplus, 1997).

merged using *HKL-2000* (Otwinowski & Minor, 1997). Data from ficin C were indexed and integrated using *iMosflm* and scaled and merged using *SCALA* (Powell *et al.*, 2013; Collaborative Computational Project, Number 4, 1994). Structure-factor amplitudes were calculated from intensities using *TRUNCATE*. Twinning tests were equally performed by *phenix.xtriage* (Zwart *et al.*, 2005). No twinning was detected in any of the data sets. Analysis of the unit-cell contents was performed with *MATTHEWS_COEF*, which is part of the *CCP4* package (Winn *et al.*, 2011).

3. Results and discussion

Ficins A, B, C and D2 were purified from fig tree latex using an SP-Sepharose Fast Flow column. As shown in Fig. 1(a), fractionation of the whole soluble fraction of *F. carica* latex on an SP-Sepharose Fast Flow column shows five major peaks designated as the unretained fraction, ficin A, ficin B, ficin C and ficin D. According to amidase activity measurements, only the active peaks assigned to ficins were selected for further purification. Fig. 1(b) displays a symmetrical peak (indicated

by the solid bar under peak I) obtained from the *S*-peglythio-ficin B conjugate, taken as an example, underlining the homogeneity of the protein samples. Their homogeneity was further checked by titrating the thiol function of their catalytic cysteine residues using 2,2'-dipyridyldisulfide (2PDS) as a titrant. Values for the number of free thiol groups per mole of enzyme ranging from 0.95 ± 0.04 to 1.03 ± 0.06 were determined. These data clearly demonstrated the usefulness of the thiol-pegylation technique to prepare highly active and pure cysteine proteases (Azarkan *et al.*, 2011). In addition, the purity of the *S*-methylthio-ficin derivatives regenerated from their thio-pegylated forms was checked by silver-stained SDS-PAGE and exemplified by ficin B (inset in Fig. 1b). The molecular masses, which were accurately determined by ESI-Q-TOF mass spectroscopy, were $23\,850 \pm 30$, $24\,130 \pm 7$, $24\,395 \pm 6$ and $24\,280 \pm 13$ Da for ficins A, B, C and D2, respectively. MALDI-TOF mass-spectrometric analysis of the purified proteins displays different peptide mass fingerprinting except for the D1 and D2 ficin forms. These two forms probably correspond to two isoforms (Azarkan *et al.*, 2011). Only the major form (D2) was thus selected for crystallization

Table 2
Summary of data-collection statistics for ficin B.
Values in parentheses are for the outermost resolution shell.

	Crystal form I	Crystal form II	Crystal form III	Crystal form IV	Crystal form V
Well solution	2.0 M ammonium sulfate, 0.1 M Na HEPES pH 7.5, 2% PEG 400	20%(w/v) PEG 4000, 0.1 M Na HEPES pH 7.5, 10%(v/v) 2-propanol	0.1 M Tris-HCl pH 8.5, 25% <i>t</i> -butanol	2.0 M sodium formate, 0.1 M sodium acetate pH 4.6	30% PEG 4000, 0.1 M Tris-HCl pH 8.5, 0.2 M lithium sulfate
Cryoprotectant	17.5%(v/v) glycerol	7.5%(v/v) glycerol	20%(v/v) glycerol	30%(v/v) glycerol	None
Wavelength (Å)	0.8856	0.9801	0.8856	0.8856	0.9801
Temperature (K)	100	100	100	100	100
Detector	ADSC Q315 CCD	ADSC Q315 CCD	ADSC Q315 CCD	ADSC Q315 CCD	ADSC Q315 CCD
Crystal-to-detector distance (mm)	220	200	290	250	292
Rotation range per image (°)	0.5	0.5	1.0	0.6	0.5
Total rotation range (°)	150	200	100	61.2	180
Exposure time per image (s)	0.8	0.8	0.8	0.8	0.8
Resolution range (Å)	26.00–1.35 (1.39–1.35)	34.80–1.35 (1.41–1.35)	36.00–1.80 (1.85–1.80)	28.00–1.70 (1.75–1.70)	35.00–2.05 (2.11–2.05)
Space group	<i>P</i> ₃ 21 or <i>P</i> ₃ 21	<i>P</i> ₂ ₁	<i>P</i> ₂ ₁ 2 ₁	<i>P</i> ₃ 21 or <i>P</i> ₃ 21	<i>P</i> ₄ 2 ₁ 2 or <i>P</i> ₄ 2 ₁ 2
Unit-cell parameters					
<i>a</i> (Å)	89.90	35.05	65.10	106.10	76.20
<i>b</i> (Å)	89.90	66.96	89.30	106.10	76.20
<i>c</i> (Å)	55.99	91.27	33.60	54.20	69.40
α (°)	90.00	90.00	90.00	90.00	90.00
β (°)	90.00	90.00	90.00	90.00	90.00
γ (°)	120.00	90.00	90.00	120.00	90.00
Observed reflections	466708 (21139)	358494 (27130)	55653 (4385)	166522 (13859)	166387 (14793)
Unique reflections	53284 (3155)	91483 (7536)	18047 (1401)	38643 (3208)	13346 (1101)
Multiplicity	8.70 (6.70)	3.90 (3.60)	3.00 (3.10)	4.30 (4.30)	12.50 (13.50)
$\langle I/\sigma(I) \rangle$	18.10 (3.30)	10.60 (2.80)	16.10 (3.80)	12.90 (2.20)	20.10 (5.30)
Completeness (%)	94.70 (66.60)	99.80 (99.90)	95.10 (90.30)	99.50 (99.50)	99.30 (100)
Mosaicity (°)	0.22	0.75	1.04	0.75	0.49
Wilson <i>B</i> factor (Å ²)	8.40	22.10	18.90	21.40	35.30
<i>R</i> _{meas} †	0.09 (0.37)	0.09 (0.45)	0.07 (0.34)	0.12 (0.58)	0.09 (0.67)
Solvent content (%)	52	67	38	40	40
No. of molecules in asymmetric unit	1	1	1	2	1

† $R_{\text{meas}} = \sum_{hkl} \{N(hkl)/[N(hkl) - 1]\}^{1/2} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$ (Diederichs & Karplus, 1997).

experiments. Screening experiments for ficin A showed a condition in which intergrown plate-shaped crystals were obtained. This condition was further optimized by varying the precipitant concentration and screening additives. No condition could be found that systematically yielded single crystals. Occasionally a single crystal was nevertheless observed, and one such crystal was used for data collection (Figs. 3*a* and 3*b*). For ficin B, a large number of hits were obtained in a variety of conditions, resulting in the discovery of five distinct crystal forms (Figs. 3*c*–3*g*), all of which diffracted to at least 2.05 Å resolution. Two of the crystal forms provided data to near-atomic resolution (1.35 Å; Table 2). For ficin C, five hits were observed that were either harvested directly from the screen or optimized. This resulted in three crystal forms suitable for data collection (Figs. 3*h*, 3*i* and 3*j*). Form I has an identical space group and unit cell to form I of ficin B and diffracted to between 1.60 and 1.20 Å resolution depending on the condition from which the crystals were harvested. Data-collection statistics for each crystal form are given in Table 1. Finally, initial screening of ficin D2 using a concentration of 5 mg ml^{−1} did not yield any hits. Doubling the concentration resulted in two hits that could be used directly for data collection (Figs. 3*k* and 3*l*). The first condition corresponded to a large crystalline chunk consisting of several individual crystals. A piece of

single crystal was obtained by breaking the larger aggregate with a needle. The resulting species diffracted to 1.45 Å resolution. In a second condition, small plate-shaped crystals were observed that diffracted to 1.90 Å resolution (Table 1). Despite their appearance, which suggested a conglomerate of many individual crystals, the diffraction patterns were relatively clean. Matthews coefficient analyses for each ficin crystal are compiled in Tables 1 and 2. It is interesting to note that purified ficin forms remain poorly characterized both from the enzymological and the structural points of view. Although the amino-acid sequences of the different ficins are currently unknown, sequencing of some peptides has shown that they all belong to the family of cysteine proteases (Azarkan *et al.*, 2011; Liener & Friedenson, 1972), which includes papain as the best characterized member. The high resolution observed for each variant should allow us to determine the structures of the different ficins by molecular replacement and to obtain preliminary electron-density-based sequences that can then be used to design primers to pick up the genes or mRNA from the plant. Although in general different ficin (iso)forms crystallize under different conditions and in distinct crystal forms, one crystal form is common to ficins B and C. For both (iso)forms, it is also the crystal form that was most commonly observed, suggesting that these two

(iso)forms may be very closely related, possibly only by one or a few amino-acid substitutions or post-translational modifications, either naturally occurring or resulting from the purification procedure. Ficins A and D2, which also presented fewer crystal forms, may thus be somewhat more distant. However, variation in crystallization behaviour between closely related proteins is also common (Buts *et al.*, 2005).

Acknowledgements

This work was supported by grants from OZR-VUB, the Hercules Foundation and VIB. The authors thank Andrew Thompson and the PROXIMA1 team for outstanding beam-line support and Laetitia Bolle for technical assistance in ficin purification.

References

- Azarkan, M., Maes, D., Bouckaert, J., Dao-Thi, M.-H., Wyns, L. & Looze, Y. (1996). *J. Chromatogr. A* **749**, 69–72.
- Azarkan, M., Matagne, A., Wattiez, R., Bolle, L., Vandenameele, J. & Baeyens-Volant, D. (2011). *Phytochemistry*, **2**, 1718–1731.
- Azarkan, M., Wintjens, R., Looze, Y. & Baeyens-Volant, D. (2004). *Phytochemistry*, **65**, 525–534.
- Azarkan, M., Wintjens, R., Smolders, N., Nijs, M. & Looze, Y. (1996). *J. Chromatogr. A*, **724**, 185–192.
- Baker, E. N. & Drenth, J. (1987). *Biological Macromolecules and Assemblies*, Vol. 3, edited by F. Jurnak & A. McPherson, pp. 313–368. New York: John Wiley & Sons.
- Boller, T. (1986). *Plant Proteolytic Enzymes*, edited by M. J. Dalling, Vol. 1, pp. 67–96. Boca Raton: CRC Press.
- Brocklehurst, K. & Little, G. (1973). *Biochem. J.* **133**, 67–80.
- Buts, L., Wellens, A., Van Molle, I., Wyns, L., Loris, R., Lahmann, M., Oscarson, S., De Greve, H. & Bouckaert, J. (2005). *Acta Cryst.* **D61**, 1149–1159.
- Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* **D50**, 760–763.
- Diederichs, K. & Karplus, P. A. (1997). *Nature Struct. Mol. Biol.* **4**, 269–275.
- El Moussaoui, A., Nijs, M., Paul, C., Wintjens, R., Vincentelli, J., Azarkan, M. & Looze, Y. (2001). *Cell. Mol. Life Sci.* **58**, 556–570.
- Erlanger, B. F., Kokowsky, N. & Cohen, W. (1961). *Arch. Biochem. Biophys.* **95**, 271–278.
- Hoorn, R. A. L. van der (2008). *Annu. Rev. Plant Biol.* **59**, 191–223.
- Huffaker, R. C. (1990). *New Phytol.* **116**, 199–231.
- Konno, K., Hirayama, C., Nakamura, M., Tateishi, K., Tamura, Y., Hattori, M. & Kohno, K. (2004). *Plant J.* **37**, 370–378.
- Liener, I. E. & Frieden, S. B. (1972). *Arch. Biochem. Biophys.* **149**, 169–174.
- Mariant, M., Camagna, M., Tarditi, L. & Seccamani, E. (1991). *Mol. Immunol.* **28**, 69–77.
- Milenic, D. E., Esteban, J. M. & Colcher, D. (1989). *J. Immunol. Methods*, **120**, 71–83.
- Mohan, S., Ma, P. W. K., Williams, W. P. & Luthe, D. S. (2008). *PLoS One*, **3**, e1786.
- Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
- Pechan, T., Cohen, A., Williams, W. P. & Luthe, D. S. (2002). *Proc. Natl Acad. Sci. USA*, **99**, 13319–13323.
- Powell, H. R., Johnson, O. & Leslie, A. G. W. (2013). *Acta Cryst.* **D69**, 1195–1203.
- Ramos, M. V., Araújo, E. S., Jucá, T. L., Monteiro-Moreira, A. C., Vasconcelos, I. M., Moreira, R. A., Viana, C. A., Beltramini, L. M., Pereira, D. A. & Moreno, F. B. (2013). *Int. J. Biol. Macromol.* **58**, 211–219.
- Richter, G., Schwarz, H. P., Dorner, F. & Turecek, P. L. (2002). *Br. J. Haematol.* **119**, 1042–1051.
- Singh, V. K., Patel, A. K., Moir, A. J. & Jagannadham, M. V. (2008). *Phytochemistry*, **69**, 2110–2119.
- Stepek, G., Behnke, J. M., Buttle, D. J. & Duce, I. R. (2004). *Trends Parasitol.* **20**, 322–327.
- Stepek, G., Buttle, D. J., Duce, I. R., Lowe, A. & Behnke, J. M. (2005). *Parasitology*, **130**, 203–211.
- Torres, M. J., Trejo, S. A., Natalucci, C. L. & López, L. M. (2013). *Planta*, **237**, 1651–1659.
- Turk, B. (2006). *Nature Rev. Drug Discov.* **5**, 785–799.
- Winn, M. D. *et al.* (2011). *Acta Cryst.* **D67**, 235–242.
- Yarishwamy, M., Shivaprasad, H. V., Joshi, V., Nanjaraj Urs, A. N., Nataraju, A. & Vishwanath, B. S. (2013). *J. Ethnopharmacol.* **149**, 377–383.
- Zwart, P., Grosse-Kunstleve, R. & Adams, P. (2005). *CCP4 Newsl. Protein Crystallogr.* **43**, contribution 7.