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

Crystallization and preliminary X-ray analysis of carnein, a serine protease from *Ipomoea carnea*

Ashok Kumar Patel,[a]‡ Niels van Oosterwijk,[b]‡ Vijay Kumar Singh,[a] Henriëtte J. Rozeboom,[b] Kor H. Kalk,[b] Roland J. Siezen,[c] Medicherla V. Jagannadham[a] and Bauke W. Dijkstra[b]*

[a]Molecular Biology Unit, Institute of Medical Sciences, Banaras Hindu University, Varanasi 221005, India, [b]Laboratory of Biophysical Chemistry, University of Groningen, Nijenborgh 4, 9747 AG Groningen, The Netherlands, and [c]Center for Molecular and Biomolecular Informatics, Radboud University Medical Centre, PO Box 9901, 6500 HB Nijmegen, The Netherlands

‡ These authors contributed equally.

Correspondence e-mail: b.w.dijkstra@rug.nl

Received 16 January 2009
Accepted 6 March 2009

Carnein is an 80 kDa subtilisin-like serine protease from the latex of the plant *Ipomoea carnea* which displays an exceptional resistance to chemical and thermal denaturation. In order to obtain the first crystal structure of a plant subtilisin and to gain insight into the structural determinants underlying its remarkable stability, carnein was isolated from *I. carnea* latex, purified and crystallized by the hanging-drop vapour-diffusion method. A data set was collected to 2.0 Å resolution in-house from a single crystal at 110 K. The crystals belonged to the trigonal space group $P3_121$ or $P3_221$, with unit-cell parameters $a = b = 126.9$, $c = 84.6$ Å, $\alpha = \beta = 90$, $\gamma = 120^\circ$. Assuming the presence of one molecule per asymmetric unit, the Matthews coefficient is 2.46 Å^3 Da^{-1}, corresponding to a solvent content of 50%. Structure determination of the enzyme is in progress.

1. Introduction

The proper functioning of a cell is ensured by the precise regulation of protein levels through balancing the rates of protein synthesis and degradation. Protein degradation is mediated by proteases, which allow the recycling of amino acids into the cellular pool. However, proteases also play a major role in the regulation of diverse cellular processes by regulating protein activities *via* controlled proteolysis in concert with other cellular regulatory mechanisms such as post-translational modification (Callis, 1995; Schaller, 2004).

Serine proteases constitute one of the largest groups of proteolytic enzymes involved in regulatory processes. They catalyze the hydrolysis of specific peptide bonds in their substrates and this activity depends on a set of amino acids in the active site of the enzyme, one of which is always a serine. They include both exopeptidases that act on the termini of polypeptide chains and endopeptidases that act on nonterminal amino acids. In plants, serine proteases appear to be the largest class of proteases (Schaller, 2004).

Serine proteases belong to several different protein families (Barrett & Rawlings, 1995; Rawlings & Barrett, 1996). The subtilisin family is the second largest serine protease family identified to date, with over 500 known members across all lineages of eubacteria, archaebacteria, eukaryotes and viruses (Siezen & Leunissen, 1997). Crystal structures have revealed that subtilisins utilize a conserved catalytic triad consisting of a Ser, a His and an Asp residue, similar to members of the chymotrypsin and serine carboxypeptidase families. However, the order of these catalytic residues in the sequence (Asp137, His168 and Ser325 in subtilisin Carlsberg) is different from that of the other serine protease families and the three-dimensional structure of subtilisins is also different (Siezen & Leunissen, 1997; Siezen *et al.*, 2007), posing a case of convergent evolution towards a Ser-His-Asp catalytic triad. An interesting feature of the subtilisin family is that some members appear to be mosaic with little or no sequence similarity to any other known proteins (Siezen & Leunissen, 1997) and have large N- and C-terminal extensions.

A large number of subtilisin-like serine proteases have been identified in various plant species, where they have been implicated in diverse processes (Meichtry *et al.*, 1999; Beers *et al.*, 2004; Schaller, 2004; Rautengarten *et al.*, 2005). Plant subtilisins have been implicated in defence against pathogens, stomata and leaf development, lateral root emergence, xylem differentiation, stress response, pro-

grammed cell death, nodulation and nitrogen fixation, microsporogenesis and seed development (Barrett & Rawlings, 1995; Rawlings & Barrett, 1996; Adam *et al.*, 2001; Palma *et al.*, 2002; Rawlings *et al.*, 2006). However, their exact functions in these processes remain unclear, mainly owing to the lack of positive identification of their physiological substrates and the lack of three-dimensional structures of plant subtilisins.

Carnein, a 80 kDa subtilisin-like serine protease from *Ipomoea carnea*, is an exceptionally stable enzyme and has been proposed to be involved in basic functions associated with various physiological processes in plants (Patel *et al.*, 2007). The full amino-acid sequence of this protein is not yet known, but a search with the N-terminal sequence of the enzyme (TTHSPEFLGLAESSGLXPNS; Patel *et al.*, 2007) revealed a high homology (75% identity) of this 20-residue peptide to a subtilisin-like protein from grape. Using the full sequence of this latter protein, *BLAST* hits were found with other subtilisin-like proteins from grape, *Arabidopsis thaliana*, tobacco, rice, fern and swamp oak (A. K. Patel & N. van Oosterwijk, unpublished work). A multiple sequence alignment of these sequences showed that they are highly similar in sequence and length (50–90% identity over 720–780 residues). The 80 kDa size of carnein fits well with this length.

The three-dimensional structure of a plant subtilisin may prove beneficial for understanding proteolytic processes in plants and their regulation. In order to elucidate the three-dimensional structural organization of carnein and its mechanism of action and to gain insights into the structural determinants underlying its exceptional stability, we have initiated the determination of its three-dimensional structure by X-ray crystallography. Here, we report the crystallization and preliminary X-ray study of the serine protease carnein isolated from *I. carnea* latex.

2. Materials and methods

2.1. Purification

The enzyme was purified from the latex of *I. carnea* essentially as described by Patel *et al.* (2007), using ammonium sulfate precipitation, cation-exchange chromatography and size-exclusion chromatography. Fresh latex secretions induced by superficial incisions made on young stems and apical buds of *I. carnea* plants were collected into 50 m*M* MES [2-(*N*-morpholino)ethanesulfonic acid] buffer pH 6.0 and stored at 253 K for 24 h. The latex was thawed at room temperature and centrifuged at 12 000*g* (Sorvall RC-5C Plus) for 20 min to remove all insoluble materials. The crude protein solution was treated with ammonium sulfate to 85% saturation and the resultant precipitate was collected by centrifugation at 12 000*g* for 30 min and redissolved in 50 m*M* acetate buffer pH 4.0 followed by dialysis against the same buffer. The dialysate was applied onto an SP Sepharose (GE Healthcare) column pre-equilibrated with the same buffer. The bound protein was eluted with an NaCl gradient in the same buffer and fractions were assayed for protease activity (Patel *et al.*, 2007). The active fractions were desalted, concentrated using a 10 kDa cutoff Vivaspin (Vivascience) concentrator and subjected to size-exclusion chromatography on a Sephacryl S-200 column (Pharmacia, 1.2 × 120 cm) with a buffer consisting of 50 m*M* MES pH 6.0 containing 0.5 *M* NaCl at a flow rate of 0.25 ml min^{-1}. The active peak fractions were pooled, concentrated, desalted and stored at 277 K in 50 m*M* MES buffer pH 6.0.

Prior to crystallization screening, the purity and conformational homogeneity of the protein was determined by SDS–PAGE and dynamic light scattering (DynaPro 801, Wyatt Technology). The purified carnein was concentrated to 15 mg ml^{-1} with a concomitant buffer exchange to 20 m*M* MES buffer pH 6.0 containing 1 m*M* $CaCl_2$ using an Amicon ultrafiltration centrifugal device (Millipore).

2.2. Crystallization

Screening for suitable crystallization conditions was performed with a Douglas Instruments Oryx-6 crystallization robot using the commercial crystallization screening solutions JCSG+ and PACT (Qiagen), Cryo I and II and Wizard I and II (Emerald BioSystems), and Structure I and II (Molecular Dimensions). The first carnein crystals were obtained when the protein solution was incubated for 30 min with 1 µ*M* of the serine protease inhibitor PMSF (phenylmethanesulfonylfluoride) before mixing the protein with the preci-

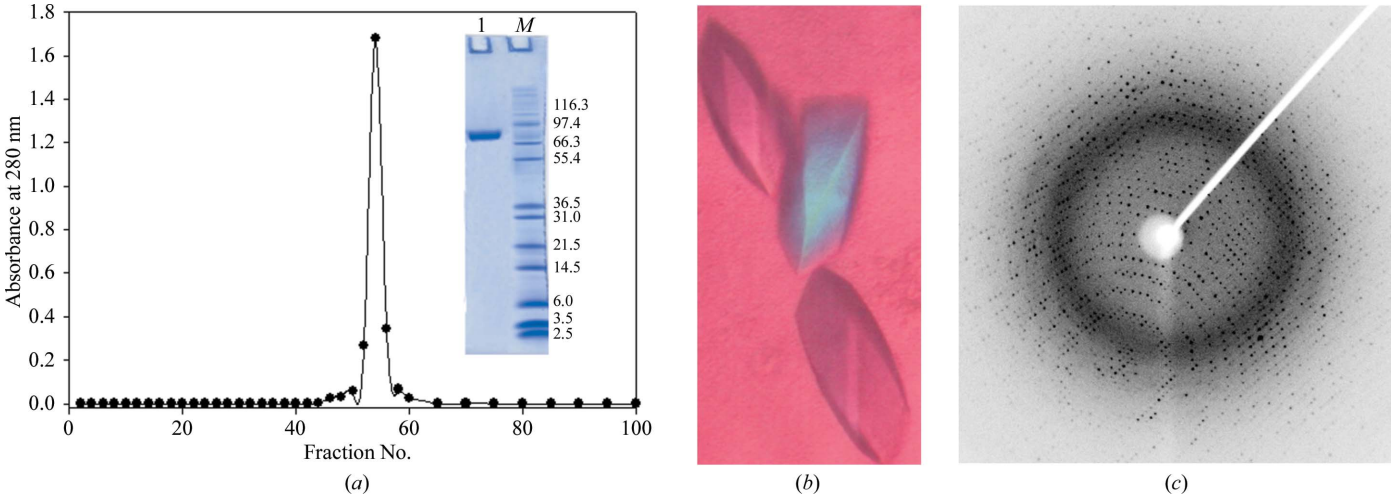

Figure 1
(*a*) Size-exclusion chromatogram of carnein obtained using a Sephacryl S-200 column with a buffer consisting of 50 m*M* MES pH 6.0 containing 0.5 *M* NaCl at a flow rate of 0.25 ml min^{-1} and a fraction volume of 1 ml. The peak fractions (52–57) were pooled and analyzed by SDS–PAGE. Insert: Coomassie Blue-stained 12.5% SDS–PAGE gel of purified carnein. Lane *M* contains molecular-weight markers (Mark 12 unstained, Invitrogen) and lane 1 contains carnein after the size-exclusion chromatography step. (*b*) Crystals of carnein grown in 4–5 d at 291 K in a hanging drop equilibrated against 750 µl reservoir solution containing 19%(*w*/*v*) PEG 6000, 0.2 *M* $CaCl_2$ in 0.1 *M* MES buffer pH 6.0. The crystals have typical dimensions of 400 × 300 × 100 µm. (*c*) Typical diffraction image collected to 2.0 Å resolution at 100 K on a DIP2030B image-plate detector (Bruker-AXS, Delft, The Netherlands) using Cu *K*α radiation from a Bruker-AXS FR591 rotating-anode generator. The crystal-to-detector distance was 150 mm and a 30 min exposure per frame of 1° oscillation was used to collect a total of 145 frames.

Table 1
Carnein diffraction data-collection and processing statistics.

Values in parentheses are for the outer resolution shell. The formulas for R_{merge} and $R_{p.i.m.}$ were taken from Weiss (2001).

Space group	$P3_121$ or $P3_221$
Unit-cell parameters	
$a = b$ (Å)	126.9
c (Å)	84.6
$\alpha = \beta$ (°)	90
γ (°)	120
V_M (Å^3 Da^{-1})	2.46
Solvent content (%)	50
Resolution range (Å)	19.8–2.0 (2.11–2.0)
Total No. of observations	447391 (55957)
No. of unique reflections	53147 (7544)
Multiplicity	8.4 (7.4)
Completeness (%)	99.6 (98.0)
R_{merge}†	0.094 (0.73)
$R_{p.i.m.}$‡	0.034 (0.28)
Mean $I/\sigma(I)$	17.7 (2.9)

† $R_{merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$. ‡ $R_{p.i.m.} = \sum_{hkl}[1/(N-1)]^{1/2} \times \sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$.

pitant solution. These crystals grew in a PACT screen condition containing 0.2 *M* $CaCl_2$ and 20%(*w*/*v*) PEG 6000 in 0.1 *M* MES buffer pH 6.0. To obtain diffracting crystals of sufficient size, further crystallization optimization was performed using the hanging-drop vapour-diffusion technique at 291 K.

2.3. Data collection and processing

A crystal was transferred to cryoprotectant solution [Paratone-N oil (Hampton Research) and paraffin oil in a 2:1 ratio] and data were collected in-house at 110 K to 2.0 Å resolution using a DIP2030B image-plate detector (Bruker AXS, Delft, The Netherlands) with Cu $K\alpha$ radiation from a Bruker AXS FR591 rotating-anode generator equipped with Franks mirrors. A crystal-to-detector distance of 150 mm and 30 min exposure per frame were used to collect a total of 145 frames with a 1° oscillation range per frame. The diffraction data set was indexed, integrated and scaled using *XDS* (Kabsch, 1993). *COMBAT* and *SCALA* (Evans, 2006) from the *CCP*4 suite (Collaborative Computational Project, Number 4, 1994) were used to convert the XDS ASCII file to an MTZ file and to graphically analyse the data-collection statistics.

3. Results and discussion

The serine protease carnein was purified by ammonium sulfate precipitation followed by SP Sepharose ion-exchange chromatography and size-exclusion chromatography. The additional size-exclusion chromatography step compared with the purification procedure of Patel *et al.* (2007) increased the specific activity of the enzyme from 0.60 to 0.86 units mg^{-1}. The protein was found to be highly pure on the basis of Coomassie R-250-stained SDS–PAGE (Fig. 1*a*). About 11.1 mg purified carnein was obtained from 187 mg crude latex. A dynamic light-scattering experiment showed that the protein was monodisperse and monomeric in solution (data not shown).

Crystallization conditions were found by screening over 500 different sitting-drop vapour-diffusion conditions. After optimization, well diffracting single crystals were grown at 291 K in 4–5 d from drops that were obtained by mixing 2 µl protein solution (15 mg ml^{-1}, preincubated for 30 min with 1 µ*M* PMSF) and 2 µl precipitant solution consisting of 750 µl 0.2 *M* $CaCl_2$, 19%(*w*/*v*) PEG 6000, 0.1 *M* MES buffer pH 6.0. The crystals had typical dimensions of 400 × 300 × 100 µm (Fig. 1*b*). Analysis of the content of the crystals by SDS–PAGE showed that they contained a protein of the expected mass of 80 kDa and a mass-spectrometric analysis of a tryptic digest of the crystallized protein was indistinguishable from a tryptic digest of the purified 80 kDa carnein.

Diffraction analysis of these crystals showed that they belonged to a primitive trigonal space group, with unit-cell parameters $a = b = 126.9$, $c = 84.6$ Å, $\alpha = \beta = 90$, $\gamma = 120°$. The symmetry of the diffraction pattern and the systematic absences (00*l* systematically absent except for $l = 3n$) indicated that the space group was $P3_121$ or $P3_221$. The crystals diffracted to 2.0 Å resolution (Fig. 1*c*) and the data statistics are summarized in Table 1. Assuming the presence of one 80 kDa molecule per asymmetric unit, the Matthews coefficient (V_M) of 2.46 Å^3 Da^{-1} falls in the generally observed range (Matthews, 1968). The solvent content is estimated to be 50%.

We are currently testing various subtilisin models for molecular replacement. In addition, we are now soaking crystals with heavy atoms in order to determine the structure of carnein by MIR phasing. Simultaneously, attempts are being made to determine the full amino-acid sequence of carnein.

Funding from the Council of Scientific and Industrial Research, India and the University Grants Commission, India is gratefully acknowledged. AKP and VKS received a travel grant from the Boehringer Ingelheim Fonds, Germany to carry out crystallization experiments and data collection in the Netherlands. Dr F. Fusetti and the Netherlands Proteomics Centre are thanked for the mass-spectrometric analysis of carnein. The authors thank Professor Jan Drenth, Tjaard Pijning, Marcel Bokhove, Pramod K. Madoori and Robert Floor (RUG, Groningen) for discussions and help during crystallization and data analysis, and Bernadet Renckens for sequence alignments.

References

Adam, Z., Adamska, I., Nakabayashi, K., Ostersetzer, O., Haussuhl, K., Manuell, A., Zheng, B., Vallon, O., Rodermel, S. R., Shinozaki, K. & Clarke, A. K. (2001). *Plant Physiol.* **125**, 1912–1918.
Barrett, A. J. & Rawlings, N. D. (1995). *Arch. Biochem. Biophys.* **318**, 247–250.
Beers, E. P., Jones, A. M. & Dickerman, A. W. (2004). *Phytochemistry*, **65**, 43–58.
Callis, J. (1995). *Plant Cell*, **7**, 845–857.
Collaborative Computational Project, Number 4 (1994). *Acta Cryst.* D**50**, 760–763.
Evans, P. (2006). *Acta Cryst.* D**62**, 72–82.
Kabsch, W. (1993). *J. Appl. Cryst.* **26**, 795–800.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Meichtry, J., Amrhein, N. & Schaller, A. (1999). *Plant Mol. Biol.* **39**, 749–760.
Palma, J. M., Sandalio, L. M., Corpas, F. J., Romero-Puertas, M. C., McCarthy, I. & del Rio, L. A. (2002). *Plant Physiol. Biochem.* **40**, 521–530.
Patel, A. K., Singh, V. K. & Jagannadham, M. V. (2007). *J. Agric. Food Chem.* **55**, 5809–5818.
Rautengarten, C., Steinhauser, D., Bussis, D., Stintzi, A., Schaller, A., Kopka, J. & Altmann, T. (2005). *PLoS Comput. Biol.* **1**, 297–312.
Rawlings, N. D. & Barrett, A. J. (1996). *Biochim. Biophys. Acta*, **1298**, 1–3.
Rawlings, N. D., Morton, F. R. & Barrett, A. J. (2006). *Nucleic Acids Res.* **34**, D270–D272.
Schaller, A. (2004). *Planta*, **220**, 183–197.
Siezen, R. J. & Leunissen, J. A. M. (1997). *Protein Sci.* **6**, 501–523.
Siezen, R. J., Renckens, B. & Boekhorst, J. (2007). *Proteins*, **67**, 681–694.
Weiss, M. S. (2001). *J. Appl. Cryst.* **34**, 130–135.