Phytochemistry 119 (2015) 5–18

Contents lists available at ScienceDirect

Phytochemistry

journal homepage: www.elsevier.com/locate/phytochem

Crystal structure of an antifungal osmotin-like protein from *Calotropis procera* and its effects on *Fusarium solani* spores, as revealed by atomic force microscopy: Insights into the mechanism of action

Marcio V. Ramos [a], Raquel S.B. de Oliveira [a], Humberto M. Pereira [b], Frederico B.M.B. Moreno [c], Marina D.P. Lobo [c], Luciana M. Rebelo [d], José Brandão-Neto [e], Jeanlex S. de Sousa [d], Ana C.O. Monteiro-Moreira [c], Cléverson D.T. Freitas [a], Thalles Barbosa Grangeiro [f,*]

[a] Departamento de Bioquímica e Biologia Molecular, Centro de Ciências, Universidade Federal do Ceará, Fortaleza, Ceará, Brazil
[b] Instituto de Física de São Carlos, Universidade de São Paulo, 13563-120 São Carlos, São Paulo, Brazil
[c] Núcleo de Biologia Experimental, Universidade de Fortaleza, Fortaleza, Ceará, Brazil
[d] Departamento de Física, Centro de Ciências, Universidade Federal do Ceará, Caixa Postal 6030, Campus do Pici, 60440-900 Fortaleza, Ceará, Brazil
[e] Diamond Light Source, Harwell Science and Innovation Campus Didcot, Oxfordshire OX11 0DE, United Kingdom
[f] Departamento de Biologia, Centro de Ciências, Universidade Federal do Ceará, Fortaleza, Ceará, Brazil

ARTICLE INFO

Article history:
Received 8 June 2015
Received in revised form 25 August 2015
Accepted 30 September 2015
Available online 5 October 2015

Keywords:
Sodom apple
Calotropis procera
Apocynaceae
Protein
Thaumatin-like protein

ABSTRACT

CpOsm is an antifungal osmotin/thaumatin-like protein purified from the latex of *Calotropis procera*. The protein is relatively thermostable and retains its antifungal activity over a wide pH range; therefore, it may be useful in the development of new antifungal drugs or transgenic crops with enhanced resistance to phytopathogenic fungi. To gain further insight into the mechanism of action of CpOsm, its three-dimensional structure was determined, and the effects of the protein on *Fusarium solani* spores were investigated by atomic force microscopy (AFM). The atomic structure of CpOsm was solved at a resolution of 1.61 Å, and it contained 205 amino acid residues and 192 water molecules, with a final *R*-factor of 18.12% and an R_{free} of 21.59%. The CpOsm structure belongs to the thaumatin superfamily fold and is characterized by three domains stabilized by eight disulfide bonds and a prominent charged cleft, which runs the length of the front side of the molecule. Similarly to other antifungal thaumatin-like proteins, the cleft of CpOsm is predominantly acidic. AFM images of *F. solani* spores treated with CpOsm resulted in striking morphological changes being induced by the protein. Spores treated with CpOsm were wrinkled, and the volume of these cells was reduced by approximately 80%. Treated cells were covered by a shell of CpOsm molecules, and the leakage of cytoplasmic content from these cells was also observed. Based on the structural features of CpOsm and the effects that the protein produces on *F. solani* spores, a possible mechanism of action is suggested and discussed.

1. Introduction

Pathogenesis-related (PR) proteins are defense molecules of plants that are induced in response to biotic and abiotic stresses. Based on their primary structures, serological relationships, and enzymatic or biological activities, these PR proteins have been classified into 17 families (PR-1 to PR-17) (van Loon et al., 2006). PR-S, an acidic 25-kDa protein that was found to be induced in tobacco leaves reacting hypersensitively to tobacco mosaic virus, is the type member of the PR-5 family (van Loon et al., 1987). Osmotin, a cationic 26-kDa tobacco protein that accumulates in response to salt stress (Singh et al., 1987), is another typical member of the PR-5 family. These two proteins (Cornelissen et al., 1986; Pierpoint et al., 1987), as well as all other PR-5 members characterized thus far, share high amino acid sequence similarity with the sweet-tasting protein thaumatin, isolated from the fruits of the West African shrub *Thaumatococcus daniellii* (van der Wel and Loeve, 1972). Therefore, members of the PR-5 family are also known as thaumatin-like proteins (TLPs).

Osmotin expression is modulated by salt and by other environmental factors such as drought and cold, and transgenic plants that overexpress osmotin genes are tolerant to these abiotic stresses

* Corresponding author at: Departamento de Biologia, Centro de Ciências, Campus do Pici, Universidade Federal do Ceará, Fortaleza, Ceará CEP 60440-900, Brazil.
E-mail address: tbgrangeiro@gmail.com (T.B. Grangeiro).

http://dx.doi.org/10.1016/j.phytochem.2015.09.012
0031-9422/

(Zhu et al., 1996; Guo et al., 2004). In addition, many osmotins and other TLPs possess antifungal activity, and this property has been associated with their ability to cause permeabilization of fungal plasma membranes (Abad et al., 1996). For example, transgenic potato plants expressing the tobacco osmotin delayed the development of disease symptoms after inoculation with spore suspensions of *Phytophthora infestans*, which causes late blight disease in potatoes (Liu et al., 1994). Thus, osmotin genes have the potential to be exploited for the improvement of crops, providing higher tolerance or resistance to stresses that reduce crop viability and productivity.

CpOsm is a 22-kDa osmotin-like protein purified from the latex of the Sodom apple, *Calotropis procera* (Apocynaceae) that has antifungal activity against *Fusarium solani*, *Neurospora* sp. and *Colletotrichum gloeosporioides* (de Freitas et al., 2011a). The protein is relatively thermostable and retains its antifungal activity over a wide pH range (de Freitas et al., 2011b). CpOsm induces membrane permeabilization of spores and hyphae of *F. solani*, and although its mechanism of action is not yet fully understood, the antifungal activity of CpOsm is likely the result of its binding to the cell wall and/or a cell membrane receptor, as previously suggested (Abad et al., 1996; Yun et al., 1997). Based on these features, CpOsm may be useful in the development of new antifungal drugs or transgenic crops with enhanced resistance to phytopathogenic fungi. However, a better understanding of its mode of action is needed.

To illuminate the mechanism of CpOsm antifungal activity, the three-dimensional structure of the protein was determined, and its effects on *F. solani* spores were investigated by atomic force microscopy (AFM). In the present work, the structural features of CpOsm and the morphological changes in *F. solani* cells induced by the protein are described, and a mechanism of action is suggested and discussed.

2. Results

2.1. cDNA cloning and sequence analysis of the encoded protein

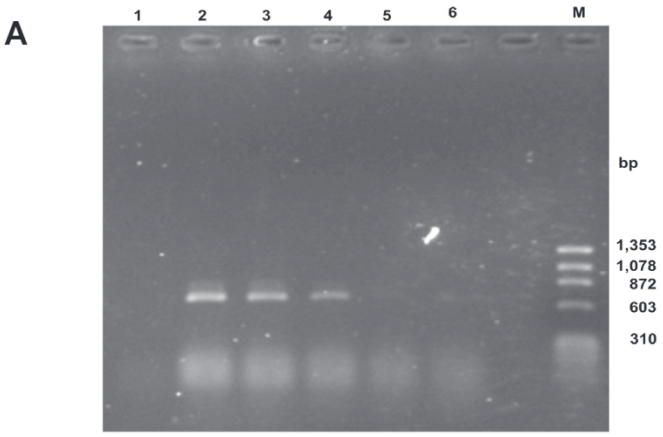
The CpOsm primary structure, which was used in the resolution of its three dimensional structure, was initially obtained by cDNA cloning. Using a pair of degenerate oligonucleotide primers, a cDNA fragment, presumptively encoding CpOsm, was successfully amplified by RT-PCR. A major DNA band of approximately 620–630 bp was obtained when PCR products were resolved by agarose gel electrophoresis (Fig. 1A). Sequencing of the cloned PCR products revealed a DNA sequence of 609 nucleotides, encoding a polypeptide chain with 203 amino acid residues (Fig. 1B). The protein sequence encoded by this cDNA fragment contained 16 Cys residues and the amino acid sequence ^{56}GRGRCQTGDCNGVLEC^{71}, which contains the thaumatin family signature G-x-[GF]-x-C-x-T-[GA]-D-C-x(1,2)-[GQ]-x(2,3)-C (Prosite accession number PS00316). The N-terminal sequence (40 amino acid residues) of the mature CpOsm, as previously determined by Edman degradation (de Freitas et al., 2011b), was in agreement with the sequence obtained from the cDNA, except for Gly^{29}, which was replaced by Gln. To confirm that this cDNA encoded the protein previously characterized from the latex of *C. procera*, the purified CpOsm was digested with trypsin, and the digestion products were analyzed by LC–ESI–MS/MS. Nine peptides were identified (Table 2), and their sequences matched exactly with specific segments in the protein primary structure derived from the cDNA fragment. The peptide labeled as 2 in Table 2 confirmed that Gln occurs at position 29, as found in the sequence obtained from the cDNA. Altogether, approximately 62.0% of the amino acid sequence encoded by the cDNA was covered by these peptide sequences.

RPS-BLAST searches against the CDD showed that CpOsm contains a single domain (201 amino acid residues, from Phe^{3} to Pro^{203}) of the GH64-TLP-SF superfamily (CDD accession number cl02511). This superfamily includes glycoside hydrolases of family 64 (GH64), thaumatin and TLPs. More specifically, significant matches were found between the CpOsm primary structure and the alignment models of the 3 subfamilies within the TLP family, TLP-P (Bit score = 222.26; e-value = 7.38e-74), TLP-PA (242.92; 5.88e-81) and TLP-F (115.49; 1.97e-31). Subfamily TLP-P (cd09217) is represented by thaumatin, allergenic/antifungal TLPs and other related plant proteins. The thaumatin-like xylanase inhibitors (TLXIs) such as those found in wheat and other cereals (Fierens et al., 2007) are also included in the subfamily TLP-P. The subfamily TLP-PA (cd09218) contains cherry (*Prunus avium*) allergen Pru Av 2 TLP (Dall'Antonia et al., 2005), peach PpAZ44 (*Prunus persica* abscission zone 44) TLP (Ruperti et al., 2002), *Caenorhabditis elegans* thaumatin family member THN-6, and other plant and animal homologs. The third subfamily, TLP-F (cd09219), is composed of TLG1, a TLP from the fungus *Lentinula edodes* (Sakamoto et al., 2006), as well as other Basidiomycota homologs. To determine the TLP subfamily to which CpOsm belongs, an NJ phylogenetic tree was generated from multiple alignments between the amino acid sequence of CpOsm and representative primary structures of the 3 TLP subfamilies. As depicted in Fig. 2, the sequences from CpOsm and the TLP-P members clustered in a supported clade (93% bootstrap), thus showing that the *C. procera* osmotin is a member of the TLP-P subfamily.

2.2. Crystal structure of CpOsm

The amino acid sequence (203 residues; calculated molecular mass of 22,093.17 Da) deduced from the cDNA fragment encoding CpOsm fit into the electron density map over the whole molecule. Moreover, the high-quality electron-density map allowed the identification of a stretch of two residues (^{204}PG^{205}) at the C-terminal end of the protein, which were not encoded in the cDNA sequence. Therefore, the CpOsm polypeptide found in the crystal has 205 amino acid residues. The calculated molecular mass of the crystalized protein is ~22,245.8 Da, and its theoretical pI is approximately 8.2. CpOsm purified from the latex of *C. procera* is a mixture of two naturally occurring isoforms (de Freitas et al., 2011a), herein named as A and B. Mass spectrometry analysis showed molecular masses of ~22,340 (isoform A) and 22,536 Da (isoform B), whereas two-dimensional gel electrophoresis revealed isoelectric points of 8.9 (isoform A) and 9.1 (isoform B) (de Freitas et al., 2011a). The calculated molecular mass of the CpOsm polypeptide chain found in the crystal (~22,245.8 Da) is different from the experimentally determined masses of the isoforms A and B (de Freitas et al., 2011a). However, the molecular mass of the crystalized CpOsm is closer to the mass of isoform A (~22,340 Da), and the observed difference (~94.2 Da) could be explained by some type of post-translational modification such as methyl-phosphorylation, which would cause a mass shift of ~94.0 Da. Likewise, a single amino acid replacement, Ala ↔ Tyr, would cause a mass change of *ca* 92.1 Da and could be another reasonable explanation. Therefore, the crystallized protein chain is likely to correspond to the isoform A of CpOsm or a closely related variant chain due to sequence microheterogeneity, thus representing a third minor isoform. Plant TLPs usually occur as mixtures of isoforms, which are very often the products of different mRNAs encoded by multigene families. High sequence identity (>95%) between TLP isoforms from the same plant source has been reported, and there are examples of isoforms that differ from each other by two or three amino acid residues (Lin et al., 1996).

The structure of CpOsm was solved by molecular replacement using the model of tobacco osmotin. The atomic model was refined



B

```
1    GCCACGTTCACTATCCGGAACAATTGCCCCTACACAATTTGGGCAGCAGCAGTGCCAGGC
      A  T  F  T  I  R  N  N  C  P  Y  T  I  W  A  A  A  V  P  G   20

61   GGCGGAAGACGCCTAAATTCCGGACAAACATGGACCATCAACGTGGCACCCGGTACCGCT
      G  G  R  R  L  N  S  G  Q  T  W  T  I  N  V  A  P  G  T  A   40

121  GGAGCACGTATTTGGCCCAGAACTAACTGCAACTTCGACGGCGCCGGACGTGGCAGATGC
      G  A  R  I  W  P  R  T  N  C  N  F  D  G  A  G  R  G  R  C   60

181  CAAACCGGCGACTGCAACGGAGTTCTCGAATGCAAAGGTTACGGTCAGCCGCCGAACACC
      Q  T  G  D  C  N  G  V  L  E  C  K  G  Y  G  Q  P  P  N  T   80

241  TTGGCCGAATACGCATTAAATCAGTTCCAAAATTTGGACTTTTTCGACATATCCTTAGTT
      L  A  E  Y  A  L  N  Q  F  Q  N  L  D  F  F  D  I  S  L  V  100

301  GACGGGTTCAACGTGCCTATGGAATTTAGTCCAGTTTCTGGTTCCGGTGACAAATGCCGT
      D  G  F  N  V  P  M  E  F  S  P  V  S  G  S  G  D  K  C  R  120

361  GCTATACGTTGTACGGCAGATATAAACGGACAGTGTCCGAATGAACTTAGAGCTCCCGGT
      A  I  R  C  T  A  D  I  N  G  Q  C  P  N  E  L  R  A  P  G  140

421  GGATGTAATAATCCGTGCACGGTGTTTAAGACCGATAAGTATTGCTGTAATTCCGGTAGT
      G  C  N  N  P  C  T  V  F  K  T  D  K  Y  C  C  N  S  G  S  160

481  TGTGGACCCACTACTTATTCGAGGTTTTTTAAGGAGAGGTGTTGGGATGCTTATAGTTAT
      C  G  P  T  T  Y  S  R  F  F  K  E  R  C  W  D  A  Y  S  Y  180

541  CCGAAAGATGATCCGACCAGTACGTTTACTTGTCCCAGTGGGACCAACTATAGAGTTATC
      P  K  D  D  P  T  S  T  F  T  C  P  S  G  T  N  Y  R  V  I  200

601  TTTTGTCCC
      F  C  P   203
```

Fig. 1. Amplification and sequence characterization of a cDNA fragment encoding the *Calotropis procera* osmotin-like protein (CpOsm). (A) Agarose gel electrophoresis of a cDNA fragment amplified by RT-PCR using two degenerate oligonucleotide primers, as described in Section 5. Amplifications were performed at different annealing temperatures: 55.0 (lane 2), 56.6 (lane 3), 60.5 (lane 4), 63.1 (lane 5) and 65.5 °C (lane 6). Lane 1: control reaction (without template). Lane M: molecular weight markers. (B) DNA sequence of the amplified PCR products, and the deduced amino acid sequence. Primer sequences are underlined. The stretch of amino acid residues corresponding to the thaumatin family signature is shaded in gray. Numbers for the first nucleotide and the last amino acid residue in each row are shown on the left and right, respectively.

to a resolution of 1.61 Å, with a final *R*-factor of 18.12% and an R_{free} of 21.59%. The final model had good stereochemistry, with no residues in the disallowed regions of the Ramachandran plot. The refined structure comprises 1561 non-hydrogen atoms from 205 amino acid residues and 192 water molecules, and a single CpOsm molecule is contained in the asymmetric unit. Overall, the protein has 14 β-strands and 6 α-helixes (Fig. 3).

A structural search against the CATH database (Sillitoe et al., 2013) showed that CpOsm structure belongs to the thaumatin superfamily fold (CATH 2.60.110.10) and is classified in the thaumatin-like protein 1 functional family (CATH FunFam 913). The CpOsm structure was also compared with the whole PDB archive using the PDBeFold service (Table 3). Significant structural alignments (*z*-scores ranging from 14 to 19.1) were recovered between

Table 1
Summary of data collection (a) and final refinement statistics (b).

(a)

Data collection[a]	Value
Space group	$P6_122$
Cell dimensions (Å) *a*, *b*, *c*	*a* = 41.83, *b* = 64.64, *c* = 87.11
Detector	PILATUS3 6M
X-ray source	DLS I24
Wavelength (Å)	0.9686
Resolution range (Å)	39.29–1.60 (1.64–1.60)
Redundancy	22.4 (16.6)
R_{meas} (%)[b]	8.5 (76.4)
R_{pim}[c]	1.8 (19.3)
Completeness (%)	99.9 (91.1)
Total reflections	617,180 (33,903)
Unique reflections	28,434 (2,040)
I/σ (I)	23.4 (3.9)

(b)

Refinement parameters	Value
Reflections used for refinement	28,131
R-factor (%)[d]	18.12
R_{free} (%)[d]	21.59
Number of protein atoms	1561
Number of ligand atoms	0
B (Å^b)	23.27
Coordinate error (ML based) (Å)	0.15
Phase error (°)	21.84
Ramachandran plot	
Favored (%)	96.57
Allowed (%)	3.43
Outliers (%)	0
All-atom clash score	4.32
RMSD from ideal geometry	
r.m.s. bond lengths (Å)	0.007
r.m.s. bond angles (°)	1.175

r.m.s.: root-mean-square.

[a] Values in parentheses are for the outer resolution shell.

[b] $R_{meas} = \frac{\sum_{hkl}\sqrt{\frac{n}{n-1}}\sum_{j=1}^{n}|I_{hkl,j}-I_{hkl}|}{\sum_{hkl}\sum_{j}I_{hkl,j}}$.

[c] $R_{pim} = \frac{\sum_{hkl}\sqrt{\frac{1}{n-1}}\sum_{j=1}^{n}|I_{hkl,j}-I_{hkl}|}{\sum_{hkl}\sum_{j}I_{hkl,j}}$.

[d] $R = R_{free} = \frac{\sum_{hkl}|F_{hkl}^{obs}-F_{hkl}^{calc}|}{\sum_{hkl}F_{hkl}^{obs}}$.

the CpOsm structure and the thaumatin and TLPs structures. As shown in Table 3, CpOsm has a high overall structural similarity with these target proteins, with RMSD values of the aligned Cα-atoms in the pairwise comparisons ranging from 0.65 to 1.34 Å.

Similarly to other TLPs, the CpOsm structure contains three domains (Fig. 3A and B). Domain I, which constitutes the central core of the molecule, contains 11 β-strands forming a β-sandwich of two β-sheets. The front sheet is formed by six anti-parallel β-strands (β2, β3, β5, β7, β8 and β13), whereas the back sheet contains the other five β-strands (β1, β4, β9, β10 and β14). Moreover, a short α-helical segment (α1) occurs between the strands β7 and β8 of the front sheet of domain I. Domain II, on one side of the central domain I, is composed of a long α-helix (α6), four α-helical segments (α2–α5) and two short β-strands (β11 and β12). Domain III, on the other side of domain I, contains a long loop with 18 amino acid residues and a β-sheet formed by two anti-parallel β-strands (β5 and β6). The strand β5 is the longest β-strand of the CpOsm structure and is shared between domains I and III, whereas β6 is exclusive to domain III. The CpOsm fold is stabilized by eight disulfide bonds, four bonds in domain I (Cys^9-Cys^{202}, Cys^{50}-Cys^{60}, Cys^{119}-Cys^{191}, and Cys^{124}-Cys^{174}; residues are numbered in relation to Ala^1, the N-terminal residue of the mature protein), four bonds in domain II (Cys^{124}-Cys^{174}, Cys^{132}-Cys^{142}, Cys^{146}-Cys^{155}, and Cys^{156}-Cys^{161}), and two bonds in domain III (Cys^{50}-Cys^{60} and Cys^{65}-Cys^{71}). One of these disulfide bonds (Cys^{124}-Cys^{174}) connects domains I and II, and another (Cys^{50}-Cys^{60}) connects domains I and III. These eight disulfide bonds are conserved in thaumatin and in the representative TLPs with known three-dimensional structures (Fig. 4).

The CpOsm surface is predominantly hydrophilic (Fig. 5A–D), but one noteworthy structural feature is the presence of a prominent cleft that runs vertically along the whole front side surface of the molecule, between domains I and II (Fig. 5E). The cavity has a volume of ~1,543 Å^3 and an average depth of *ca* 8.6 Å. The topology of the rift resembles a left-hand hollow in which the front β-sheet of the domain I forms the palm and the domain II contains the thumb. The central portion of the cleft is very acidic, as revealed by electrostatic modeling (Fig. 5A). The electronegative character of this region is due to four acidic residues (Glu^{83}, Asp^{96}, Asp^{101} and Asp^{184}), the side-chains of which extend into the cavity of the cleft (Fig. 5F). Two of these acidic residues (Glu^{83} and Asp^{96}) of CpOsm are conserved in thaumatin and in all TLPs with known three-dimensional structure (Fig. 4). The other two acidic residues (Asp^{101} and Asp^{184}) of the electronegative cleft of CpOsm are conserved in the TLPs, but both are replaced by Lys residues in thaumatin (Fig. 4), which has a predominantly basic cleft (Batalia et al., 1996). Although one basic residue (Arg^{43}) occurs in the cleft of CpOsm, its side-chain is masked by an aromatic residue (Tyr^{74}); therefore, its contribution to the surface characteristics in this region is not significant. Indeed, the CpOsm acidic cleft is surrounded by four aromatic residues: Tyr^{74}, Phe^{89}, Phe^{94}, and Tyr^{178}. The side-chains of Phe^{89} and Phe^{94} form a hydrophobic patch at the bottom edge of the cleft, whereas the aromatic rings of Tyr^{74} and Tyr^{178} are located on the right and left borders of the acidic region of the rift, respectively (Fig. 5C and F). Excluding the acidic cleft, the rest of the front side surface and the whole back surface of CpOsm have a predominantly basic character, but these positively charged regions are evenly distributed and interspersed with patches of acidic regions (Fig. 5B). This is consistent with the biochemical behavior of the purified CpOsm, which is a mixture of two closely related isoforms with pI values of 8.9 and 9.1, respectively (de Freitas et al., 2011a).

2.3. The effect of CpOsm on F. solani spores

The effect of CpOsm on *F. solani* spores was investigated by AFM. The protein produced a remarkable effect on cell morphology, as depicted in Fig. 6. The dimensions (length, height and width) of untreated spores were 15.1 ± 1.6 μm, 2.3 ± 0.5 μm, and 5.6 ± 0.5 μm, respectively, with a cell volume estimated to be approximately 154.0 ± 23.0 μm^3 (Fig. 6A). The length, height and width of spores exposed to CpOsm were, on average, 11.6 ± 1.2 μm, 0.8 ± 0.2 μm and 5.8 ± 0.5 μm, respectively. Spores treated with CpOsm were highly shrunken (Fig. 6B), and the volume of these wrinkled cells was estimated to be, on average, 30.6 ± 6.0 μm^3, which is approximately one-fifth of the volume of untreated cells. Moreover, spores treated with CpOsm were covered by a shell of ring-like structures, which were likely oligomers or aggregates of CpOsm (Fig. 6D and E). An overflow of intracellular content from these wrinkled cells was also observed (Fig. 6F). Therefore, CpOsm drastically altered the morphology of *F. solani* spores, causing a *ca* 80% reduction in the cell volume.

3. Discussion

Thaumatin, a sweet-tasting protein from the miracle berry (*T. daniellii*, Marantaceae), is the prototype of the thaumatin family; the three-dimensional structures of 2 thaumatin isoforms have been determined (Masuda et al., 2011a, 2011b). Structures from other members of the thaumatin family, the so-called

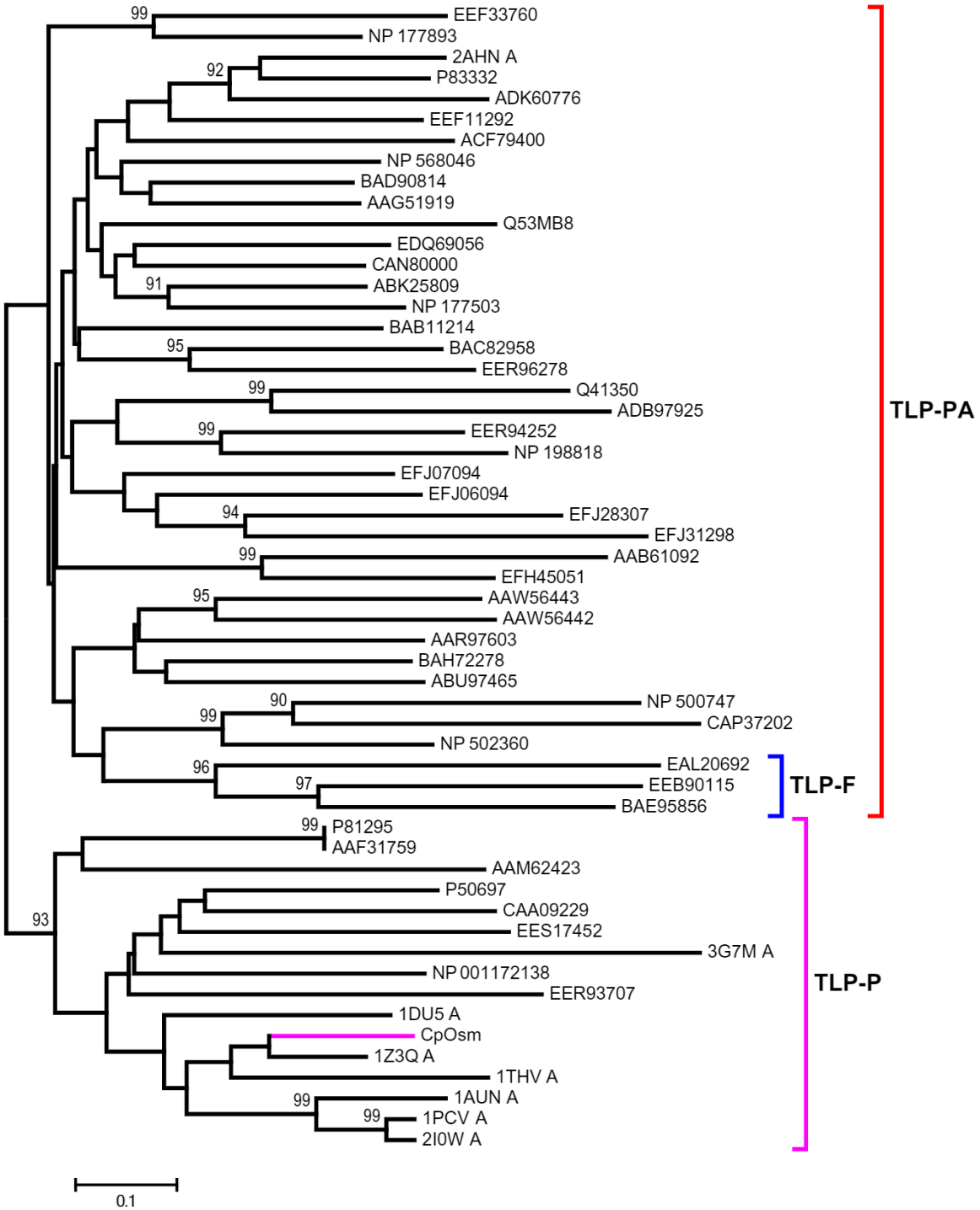


Fig. 2. Unrooted neighbor-joining (NJ) tree showing the phylogenetic relationship of CpOsm with representative members of the 3 subfamilies within the TLP family. The amino acid sequence of CpOsm (this work) was aligned to the sequences of representative members of the subfamilies TLP-P, TLP-PA and TLP-F of the TLP family, and the evolutionary distances were computed using the Poisson correction method. The optimal tree with the sum of branch length = 16.79344343 is shown. The percentages of replicate trees in which the associated sequences clustered together in the bootstrap test (1000 replicates) are shown next to the branches. The scale bar represents the branch lengths (the number of amino acid substitutions per site). The analysis involved 55 amino acid sequences and sites containing insertions/deletions in pairwise comparisons were ignored. There were a total of 284 positions in the final dataset. The PDB codes or GenBank accession numbers of the sequences used are shown.

thaumatin-like proteins (TLPs), have also been resolved. The TLPs with known three-dimensional structures are from species belonging to diverse plant families, as follows: maize (*Zea mays*, Poaceae; PDB ID of the structure: 1DU5) (Batalia et al., 1996), wheat (*Triticum aestivum*, Poaceae; PDB ID: 3G7M) (Vandermarliere et al., 2010), tobacco (*Nicotiana tabacum*, Solanaceae; PDB IDs: 1AUN and 1PCV) (Koiwa et al., 1999; Min et al., 2004), tomato (*Solanum lycopersicum*, Solanaceae; PDB ID: 2I0W) (Ghosh and Chakrabarti, 2008), banana (*Musa acuminata*, Musaceae; PDB ID: 1Z3Q) (Leone et al., 2006), wild cherry (*Prunus avium*, Rosaceae; PDB ID: 2AHN), kiwi-fruit (*Actinidia deliciosa*, Actinidiaceae; PDB ID: 4BCT), grape (*Vitis vinifera*, Vitaceae; PDB IDs: 4L5H and 4JRU) (Marangon et al., 2014), and apple (*Malus domestica*, Rosaceae; PDB ID: 3ZS3). Therefore, the resolved structures of

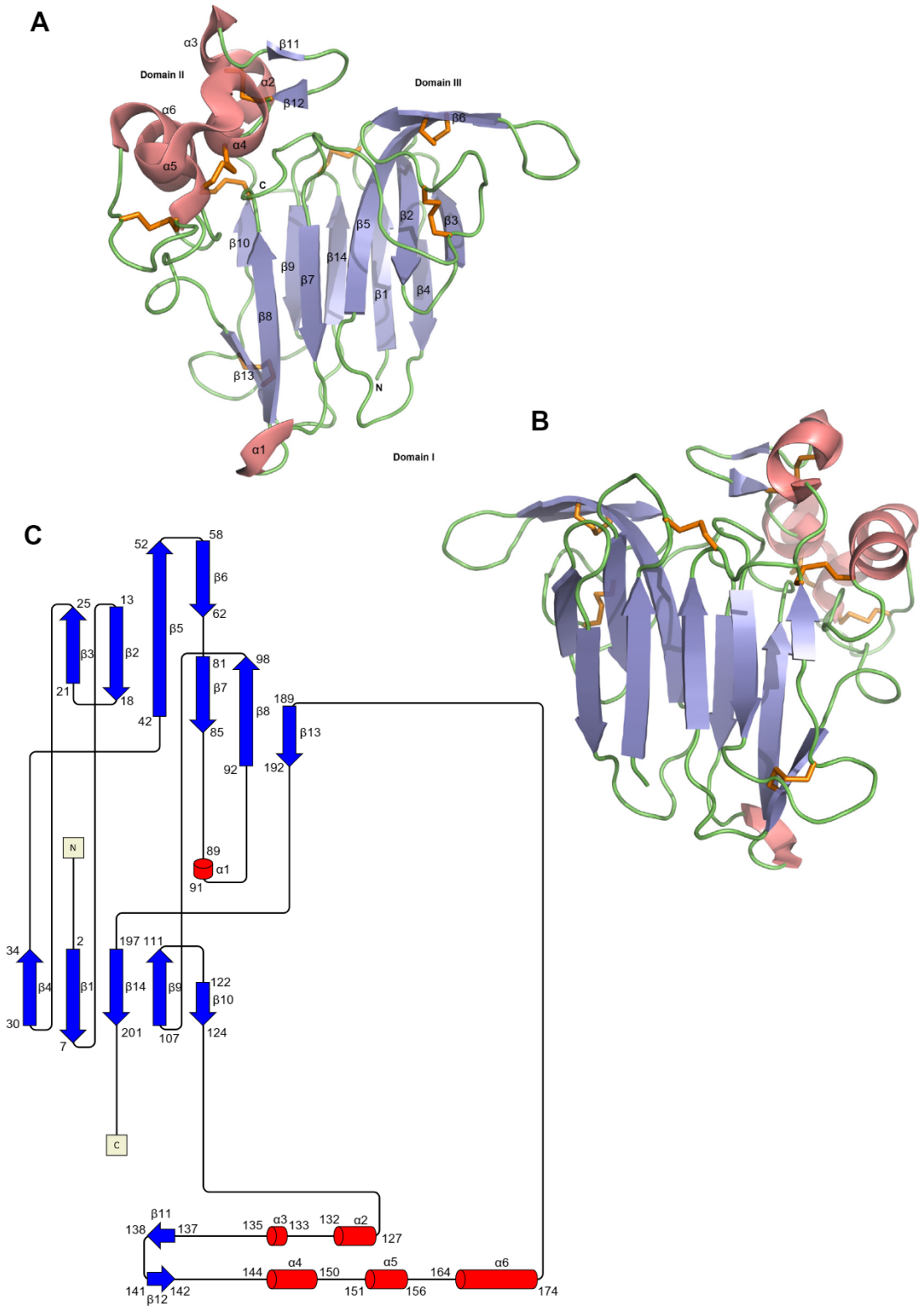


Fig. 3. Three-dimensional structure of CpOsm. Stereo ribbon diagram ((A): front view; (B): back view, rotated 180°) of the overall three-dimensional structure of CpOsm, which is composed of three domains (I, II and III). The eight disulfide bonds are shown in orange. (C) Topology of CpOsm. The residues (numbers are relative to Ala[1], the N-terminal residue of the mature protein) included in each β-strand or α-helix are shown. The topology diagram was generated with Pro-origami (Stivala et al., 2011) using DSSP to determine the secondary structural elements (Kabsch and Sander, 1983). (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

proteins with the thaumatin fold are from representative members of the two major clades of the flowering plants, with 4 members (miracle berry, maize, wheat and banana) of the monocots and 6 members (tobacco, tomato, wild cherry, kiwi, grape and apple) of the eudicots. The monocot structures are from members of the orders Poales (maize and wheat) and Zingiberales (miracle berry

Table 2
Amino acid sequences of tryptic peptides from *Calotropis procera* osmotin identified by LC–ESI–MS/MS. Peptides matching segments of the protein encoded by the cDNA cloned in the present study (GenBank accession number KR270529) are shown.

Peptide	Mass (Da)		Score[a]	Expect value[b]	Sequence[c]	Sequence coverage (%)[e]	PDB ID or GenBank accession number
	Experimental	Calculated					
1	1802.8788	1802.8522	76	2.9e-04	R.^{7}NNCPYTIWAAAVPGGGR^{23}.R	8.4	4L2J
2	2069.0944	2069.0766	57	1.1e-02	R.^{24}RLNSGQTWTINVAPGTAGAR^{43}.I	9.8	4L2J
3	1912.9586	1912.9755	121	2.6e-09	R.^{25}LNSGQTWTINVAPGTAGAR^{43}.I	9.4	4L2J
4	1111.4168	1110.4513	60	6.7e-04	R.^{48}TNCNFDGAGR^{57}.G	4.9	4L2J
5	1540.6074	1539.6116	75	4.2e-05	R.^{60}CQTGDCNGVLECK^{72}.G	6.4	4L2J
6	1647.6870	1646.7141	62	8.5e-03	R.^{124}CTADINGQCPNELR^{137}.A	6.9	4L2J
7	1420.6112	1420.6228	101	2.6e-07	R.^{138}APGGCNNPCTVFK^{150}.T	6.4	4L2J
8	1768.6788	1768.6604	53	3.5e-03	K.^{154}YCCNSGSCGPTTYSR^{168}.F	7.4	4L2J
9	1188.4996	1188.4910	57	4.8e-04	R.^{174}CWDAYSYPK^{182}.D	4.4	4L2J
10	1787.7428	1787.7421	105	2.7e-08	K.^{183}DDPTSTFTCPGGTNYR^{198}.V (S)[d]	7.4	ABX71220

[a] Score values calculated by Mascot (score = $-10 \times$ Log (p)) express the probability p that a match of calculated and experimental mass is by chance; a score of 30, for example, accounts for $p \leqslant 0.001$.
[b] Expect value indicates the probability that the observed match between MS/MS spectra and peptide sequence occur by chance alone.
[c] The numbers before and after each sequence indicate the residue positions relative to Ala^{1}, which is the N-terminal residue of the mature protein; the sequences are separated from flanking amino acid residues by dots.
[d] The underlined residue (Gly) of this peptide is replaced by Ser^{193} (indicated in parenthesis) in the amino acid sequence deduced from the cDNA.
[e] Percentage of the protein amino acid sequence that is covered by the matched peptides.

Table 3
Structural comparison between the CpOsm structure (4L2J) and other TLP structures using the protein structure comparison service Fold at European Bioinformatics Institute (http://www.ebi.ac.uk/msd-srv/ssm).

Target structure (PDB ID)	*Q*-Score[a]	*P*-Score[b]	*Z*-Score[c]	RMSD (Å)[d]	N_{align}[e]	N_g[f]	Sequence identity (%)[g]	Matched SSEs (%)[h]	N_{res}[i]
1PCV	0.90	40.7	19.1	0.83	202	3	72	100/94	205
4JRU	0.90	40.4	19.0	0.65	197	1	73	88/88	201
1AUN	0.93	37.7	18.3	0.68	204	2	69	94/94	208
1Z3Q	0.91	37.3	18.2	0.65	198	2	78	88/93	200
1DU5	0.84	36.2	18.1	0.89	197	4	67	88/88	206
4BCT	0.90	35.7	17.8	0.77	199	2	80	94/88	201
1RQW	0.86	34.7	17.6	0.84	199	3	61	88/88	207
3AOK	0.87	34.7	17.6	0.84	199	3	62	69/69	207
1THV	0.88	33.1	17.2	0.82	200	3	62	81/87	207
2I0W	0.91	31.0	16.6	0.87	204	2	71	81/76	207
4L5H	0.91	29.9	16.3	0.60	196	1	76	75/67	198
2AHN	0.66	23.4	14.3	1.28	189	8	43	75/71	222
3ZS3	0.66	20.7	14.0	1.34	190	9	39	75/67	222

[a] Represents the quality function of Cα-alignment and is based on alignment length and RMSD.
[b] Probability of an equal hit by chance.
[c] Measures the statistical significance of a match in terms of Gaussian statistics.
[d] Root mean square deviation, calculated between Cα-atoms of matched residues at the best 3D superposition of the query and target structures.
[e] Length of alignment or number of matched residues, which is calculated at the best 3D superposition of query and target structures.
[f] Number of gaps of Cα-alignment.
[g] Percentage of pairs of identical residues among all aligned residues, calculated from structure alignment.
[h] Percentage of matched secondary structure elements (SSEs) of query chain identified in the target protein (before the slash) and percentage of SSEs of target chain identified in the query protein (after the slash).
[i] Number of residues of the target chain.

and banana) of the clade commelinids. The eudicots TLP structures are from representative species of the orders Solanales (tobacco and tomato), Ericales (kiwi), Rosales (wild cherry and apple) and Vitales (grape).

In the present work, the three-dimensional structure of CpOsm, an osmotin-like protein from the latex of the apple of Sodom (*C. procera*, Apocynaceae), was resolved by X-ray crystallography. The atomic structure of CpOsm reported here is the first one of a representative TLP of the order Gentianales, a member of the eudicots clade.

Approximately 99.95% of the extant angiosperms form Mesangiospermae, a supported monophyletic group, comprising five major lineages: eudicots, monocots, magnoliids, Chloranthaceae and Ceratophyllaceae (Cantino et al., 2007). Eudicots and monocots are the two largest groups of Mesangiospermae, containing *ca* 75% and 22% of angiosperm species, respectively (Drinnan et al., 1994). It has been hypothesized that monocots was the first group to diverge from the other Mesangiospermae (Lee et al., 2011; Zeng et al., 2014; Wickett et al., 2014); this event is estimated to have occurred ~165–170 million years ago (Zeng et al., 2014). Despite this divergence time, the CpOsm three-dimensional structure (this work) is very similar to the structures of thaumatin and of other TLPs from eudicots, as well as to those of monocots species (Table 3). This indicates a strong evolutionary pressure to conserve the thaumatin fold in flowering plants, which is very likely related to the function of these proteins as defense molecules against pathogens and to their role in the adaptation or tolerance of plants to abiotic stresses.

Far-UV circular dichroism (CD) and fluorescence spectroscopy analyses of CpOsm solutions at various pH values (3.0, 5.0, 7.0 and 9.0) revealed CD and fluorescence spectra with essentially the same shape. Moreover, CD spectroscopy also showed that the CpOsm conformation remained stable at temperatures up to 75 °C, and its thermal denaturation was calculated to be 77.8 °C (de Freitas et al., 2011b). The conformational stability of CpOsm

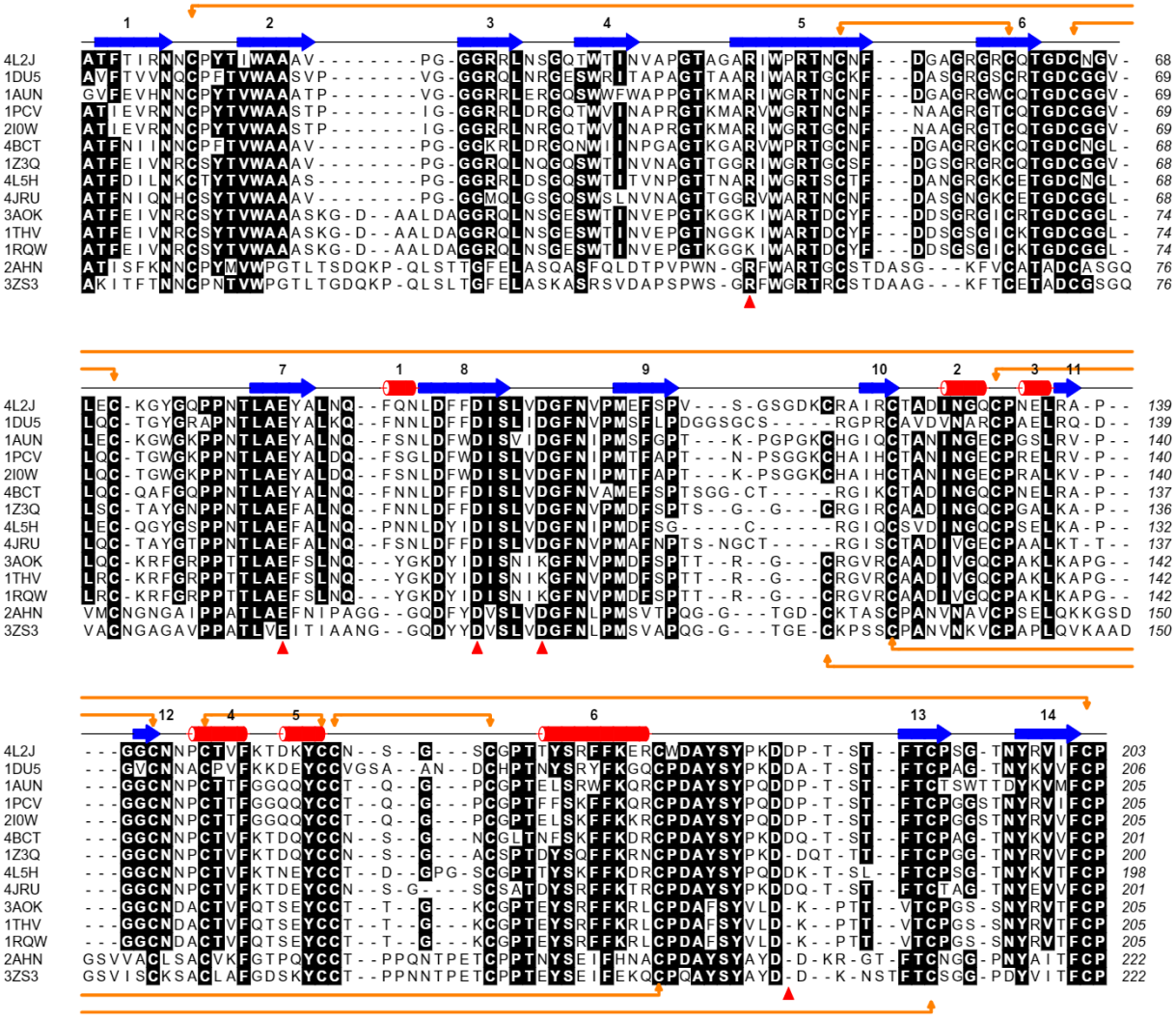


Fig. 4. Structure-based multiple sequence alignment of CpOsm, thaumatin and TLPs. The structure of CpOsm (4L2J) was superimposed onto the structures of thaumatin (3AOK, 1THV and 1RQW) and TLPs from different plant species using the tool MatchMaker of the program UCSF Chimera version 1.10 (Pettersen et al., 2004). The TLP structures were from *Zea mays* (1DU5), *Nicotiana tabacum* (1AUN and 1PCV), *Solanum lycopersicum* (2I0 W), *Musa acuminata* (1Z3Q), *Prunus avium* (2AHN), *Actinidia deliciosa* (4BCT), *Vitis vinifera* (4L5H and 4JRU), and *Malus domestica* (3ZS3). The multiple sequence alignment was then generated from the superimposed structures using the tool Match->Align of UCSF Chimera. Positions containing the same residue in at least 8 sequences are shaded. The α-helices and β-strands of CpOsm are shown as red cylinders and blue arrows, respectively, whereas the disulfide bonds are indicated by orange lines. Sites containing the charged residues of the central cleft of CpOsm are indicated by red triangles. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

to elevated temperatures and to pH variation can be explained by the fold adopted by its polypeptide chain (Fig. 3), which is stabilized by eight disulfide bonds, a characteristic feature of members of the TLP family.

CpOsm is an antifungal TLP that is capable of inhibiting the spore germination of *F. solani* (IC_{50} = 67.0 μg/mL), *Neurospora* sp. (IC_{50} = 57.5 μg/mL) and *C. gloeosporioides* (IC_{50} = 32.1 μ/mL), as previously reported (de Freitas et al., 2011a). Moreover, it has been demonstrated that CpOsm induces membrane permeabilization of spores and hyphae of *F. solani*, as revealed by a propidium iodide uptake assay (de Freitas et al., 2011b). In the present study, direct evidence of the interaction between CpOsm and the pathogen cell surface was obtained by AFM analysis (Fig. 6). The images of *F. solani* spores treated with CpOsm clearly show that the fungal cells were covered by a layer of ring-like structures, which could be interpreted as oligomers or aggregates of CpOsm. The wrinkled shape of the fungal cells treated with CpOsm immediately suggests that once the protein binds to cell surface components, it induces a dramatic change in the plasma membrane permeability, leading to an overflow of cytoplasm content and causing the cell to shrink. These findings more dramatically corroborate with earlier evidence regarding the mechanism of action of plant antifungal TLPs (Roberts and Selitrennikoff, 1990; Woloshuk et al., 1991; Abad et al., 1996).

A causal relationship between *Saccharomyces cerevisiae* cell wall phosphomannoproteins, tobacco osmotin binding and cytotoxicity has been found (Ibeas et al., 2000). Moreover, several TLPs, including the antifungal barley protein BP-R and corn zeamatin, bind to fungal cell wall β-1,3-glucans *in vitro* (Trudel et al., 1998). Based on these observations, Ibeas et al. (2000) have suggested that the probable function of these cell wall components is to capture the antifungal TLP from the environment, increasing the rate of protein diffusion across the cell wall.

In the AFM image showed in Fig. 6D, the close association of CpOsm molecules around the *F. solani* spore, rather than being distributed all over the substrate surface, supports this capture hypothesis. The initial interaction of CpOsm with the fungal cell may occur between the overall positively charged surface of

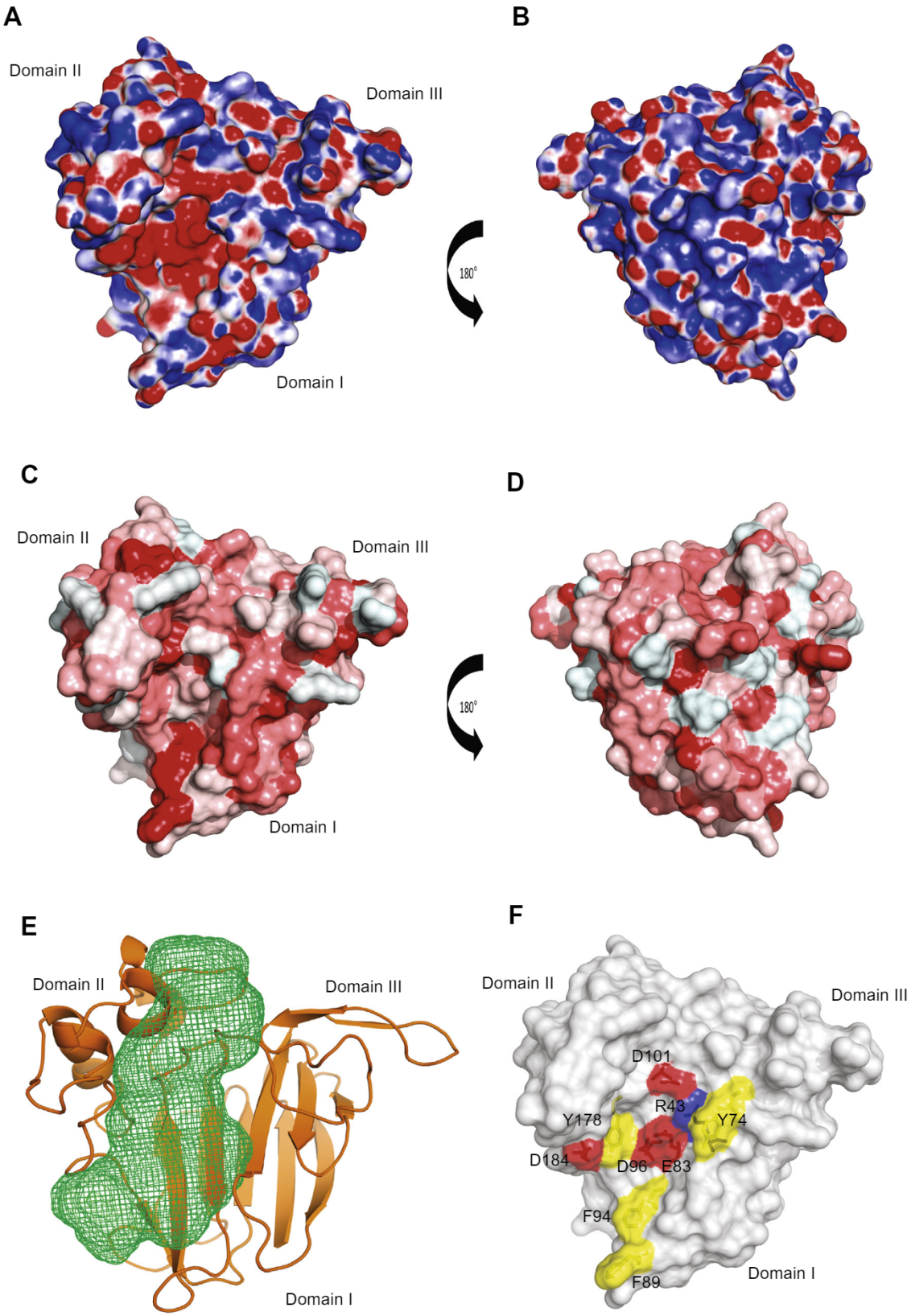


Fig. 5. Molecular surface of CpOsm. Mapping of the electrostatic potentials on the molecular surfaces of the front (A) and back (B) sides of CpOsm. The negative potentials (colored red) and the positive potentials (colored blue) are displayed at −10 kT/e level and +10 kT/e level, respectively. Neutral surfaces are white. Mapping of the surface hydrophobicity on the front (C) and back (D) faces of CpOsm. The protein surface is colored according to the Eisenberg hydrophobicity scale (Eisenberg et al., 1984), from red (hydrophobic) to white (hydrophilic). (E) CpOsm structure shown as a cartoon (colored in orange) depicting the location and shape of the main surface cleft, which is represented as a green colored triangulated mesh. The cleft was calculated using the SURFNET program (Laskowski, 1995), available at the PDBsum web server (http://www.ebi.ac.uk/pdbsum/). (F) Molecular surface of the front side of CpOsm. The acidic (red), basic (blue) and aromatic (yellow) residues located inside and around the central cleft are colored and labeled. (For interpretation of the references to color in this figure legend, the reader is referred to the web version of this article.)

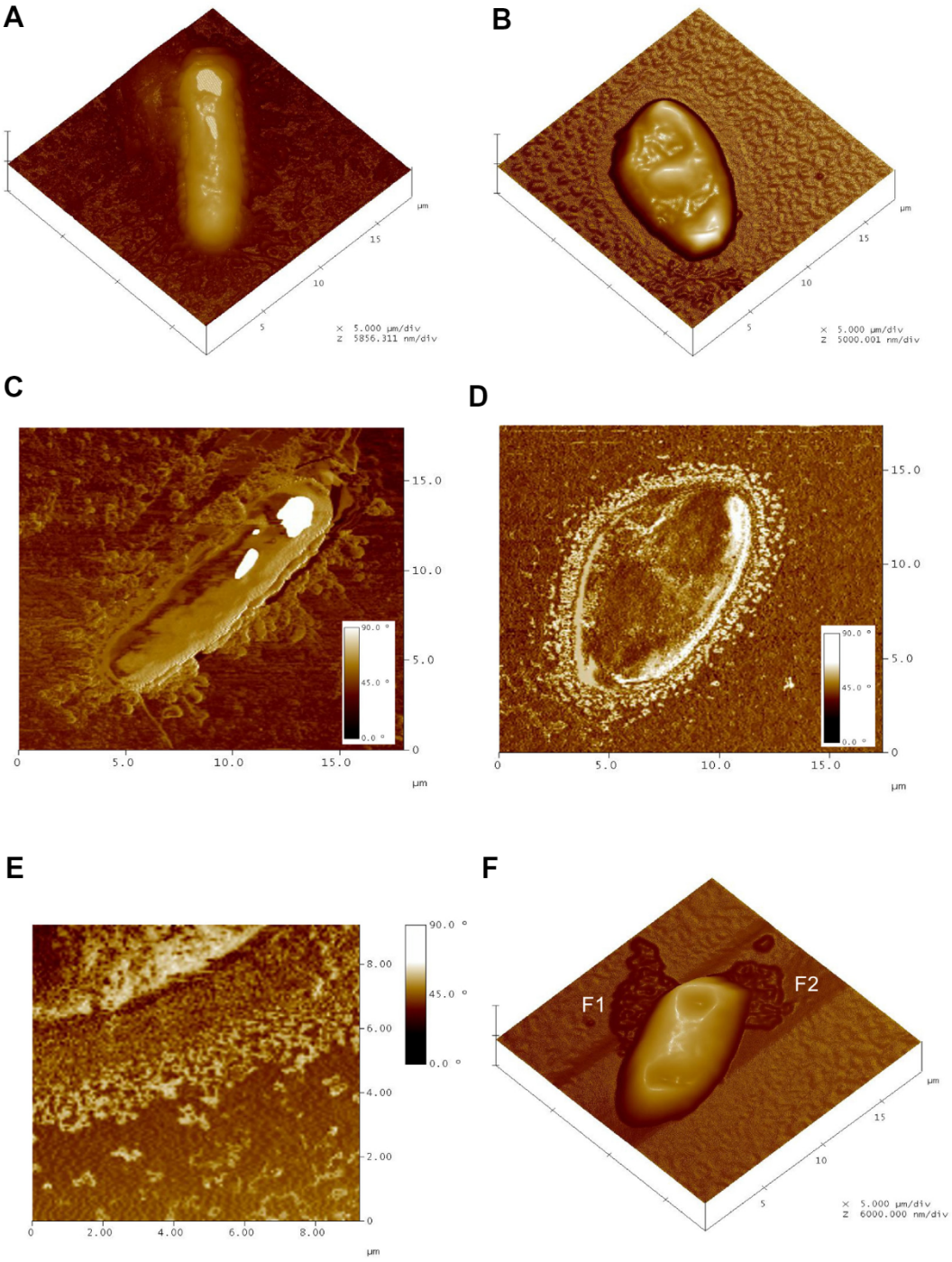


Fig. 6. Atomic force microcopy images of *Fusarium solani* spores treated with CpOsm. Three-dimensional images of untreated spores (A) and spores treated with CpOsm (B). Phase-contrast angle images of the same specimens as in panels (A) (C) and (B) (D). The brighter region in image (C) corresponds to the *z*-limit of the equipment, and salt crystals are observed surrounding the spore. These salt crystals are found not only around the spore but also in other regions of the substrate. In the treated spore (D), however, a lighter halo is visible around the cell. (E) Zoom of the light halo observed in image (D). Ring-like structures, compatible with protein oligomers or aggregates, are observed. (F) Three-dimensional image of a wrinkled spore treated with CpOsm. In addition to the evident changes in the cell morphology, the leakage of cellular content in the vicinity of the spore is also observed (areas labeled as F1 and F2).

CpOsm and the negatively charged fungal cell surface (Beezer and Sharma, 1981). Indeed, CpOsm interacts with large unilamellar vesicles of the negatively charged 1-palmitoyl-2-oleoyl-sn-glycero-3-phospho-rac-(1-glycerol) (de Freitas et al., 2011b). Therefore, non-specific ionic interactions would direct CpOsm to the fungal cell surface, thereby increasing its concentration around the spore and its diffusion towards the cell membrane.

Once CpOsm has bound to the cell surface and diffused across the cell wall, it would reach the plasma membrane, where it could affect its permeability, causing the leakage of cytoplasmic content. Although the precise mechanism by which CpOsm and other antifungal TLPs alter the cell membrane permeability still remains unknown, it has been speculated that negatively charged residues in the acidic cleft would enable these proteins to bind positively

charged membrane proteins such as ion or water channels or osmoreceptors, which would eventually cause a dramatic increase in the water flow across the membrane (Batalia et al., 1996). The side chains of Phe[89] and Phe[94] of CpOsm, which form a hydrophobic patch at the bottom edge of the acidic cleft (Fig. 5F), could also contribute to the interaction between CpOsm and fungal cell surface receptors. These hydrophobic residues could be involved in stacking interactions with protein aromatic side-chains or non-protein ligands containing themselves aromatic groups. These Phe residues of CpOsm are conserved in antifungal PR-5 proteins, but not in thaumatin, which lacks antifungal activity, and contains Tyr instead of Phe residues (Koiwa et al., 1999). The aromatic side chain of Tyr differs from that of Phe only in the presence of a hydroxyl group in place of a hydrogen atom in the ortho position on the benzene ring. Therefore, the two Tyr residues of thaumatin would play essentially the same role hypothesized for the two Phe residues of CpOsm. This suggests that, although the hydrophobic patch could play a role in the interaction between CpOsm and the fungal cell receptor(s), these hydrophobic residues are not the key structural determinants that could explain the dichotomic effect of thaumatin and CpOsm on fungal cells.

4. Conclusions

In conclusion, the data reported in the present study and the experimental evidence in previous reports supports a two-step mechanism of action for the antifungal CpOsm, involving cell wall binding and membrane perturbation. However, more work is needed to unravel the exact details of this proposed mechanism.

5. Experimental

5.1. Plasmid, cells and reagents

The plasmid pGEM-T Easy was purchased from Promega (Madison, WI, USA). The general *Escherichia coli* cloning strain TOP10F′ was purchased from Invitrogen (Carlsbad, CA, USA). All other reagents were of analytical grade.

5.2. RNA extraction

Total RNA was isolated from *C. procera* leaves using the Concert Plant RNA Reagent (Invitrogen) according to the manufacturer's instructions. The integrity of the RNA samples was assessed by 1% agarose gel electrophoresis, and the yield was estimated by measuring the absorbance at 260 nm (Sambrook et al., 1989).

5.3. cDNA synthesis, amplification and cloning

Prior to cDNA synthesis, total RNA was treated with RQ1 RNase-free DNase I (Promega) at 37 °C for 30 min (1 U of DNase I per μg of RNA) and cleaned up using the RNeasy Mini kit (Qiagen, Hilden, Germany). Treated RNA was recovered in nuclease-free H_2O (35 μL) and used for cDNA synthesis with oligo$(dT)_{18}$ (Fermentas Life Sciences, Ontario, Canada) and the ImProm-II Reverse Transcription System (Promega) according to the protocol supplied by the manufacturer. The first-strand cDNA products were then amplified by polymerase chain reaction (PCR) using two oligonucleotide primers, which were designed to amplify a cDNA segment encoding the mature sequence of CpOsm. The sequences of these primers were as follows: 5′-GCCACNTTYACNATHCGNAACAAYTGYCC-3′ (forward) and 5′-GGGRCARAANAYAACYCTRTARTTDGT-3′ (reverse). The design of the forward primer was based on the first 10 N-terminal amino acid residues of the mature CpOsm as previously reported (de Freitas et al., 2011a). To design the reverse primer, the sequence of a tryptic CpOsm peptide (DDPTSTFTCPGGTNYR), which was identified previously by LC–ESI–MS/MS (de Freitas et al., 2011a), was used as the starting information. This peptide was aligned to a multiple alignment containing the complete primary structures of 10 representative TLPs from plant species belonging to different families. The sequences were from *Piper colubrinum* (Piperaceae; GenBank accession number ABX71220), *Helianthus annuus* (Asteraceae; AAM21199), *Actinidia deliciosa* (Actinidiaceae; AGC39175 and ABQ42566), *Fragaria* × *ananassa* (Rosaceae; AAF13707), *Populus trichocarpa* (Salicaceae; XP_002299521 and XP_002299547), *Populus deltoides* (ACZ67185), *Populus tomentosa* (ADP69173), and *Sambucus nigra* (Adoxaceae; AAK59275). Based on these alignments, the last 9 C-terminal residues of CpOsm were presumed to be TNYRVV/IFCP, which were then used to design the reverse primer. In this sequence, the first 4 residues (underlined) were derived from the CpOsm tryptic peptide, whereas the other 5 residues were inferred from multiple alignments of the primary structures of the representative TLPs. Moreover, the cDNA sequences encoding the 10 representative TLPs, as described above, were also retrieved from the GenBank sequence database (accession numbers EU271754, AF364864, JX905282, EF417825, AF199508, XM_002299485, XM_002299511, GU129189, HM589200, and AF378571). These nucleotide sequences were aligned, the codon preferences for each N- and C-terminal amino acid residue were recorded, and these data were also considered during primer design. Amplifications were carried out in a final volume of 10 μL containing first-strand cDNA (570 ng), 1× GoTaq reaction buffer (Promega), 200 μM each dNTP, 0.5 μM each primer, and 1.25 U GoTaq DNA Polymerase (Promega). Reactions were performed in the Mastercycler gradient thermocycler (Eppendorf, Hamburg, Germany) using the following cycling parameters: an initial denaturation step (3 min at 95 °C) followed by 34 cycles of 1 min at 95 °C, 1 min at 55–65.5 °C and 1 min at 72 °C. After the last cycle, the reactions were further incubated for 5 min at 72 °C. The amplification of a DNA band of the expected size was confirmed by analyzing a 5 μL aliquot of the PCR reactions by 1% agarose gel electrophoresis. An aliquot of the remaining reaction (2 μL) was ligated into the pGEM-T Easy vector using T4 DNA Ligase. Products from the ligation reaction were introduced in *E. coli* TOP10F′ by electroporation, and the transformants were selected on LB agar containing 100 μg mL^{-1} carbenicillin, 30 μg mL^{-1} streptomycin, 0.5 mM IPTG and 80 μg mL^{-1} X-Gal. Plasmid DNA was isolated from antibiotic-resistant colonies using the PureLink Quick Plasmid Miniprep kit (Invitrogen) according to the supplier's protocol, and the presence of the insert was confirmed by restriction digestion with *Eco*RI.

5.4. DNA sequencing and assembly

PCR products cloned into pGEM-T Easy vector were sequenced using the DYEnamic ET Dye terminator cycle sequencing kit (GE Healthcare, Buckinghamshire, UK), following the protocol supplied by the manufacturer. Both strands were sequenced using the universal primers M13 (-40) forward (5′-GTTTTCCCAGTCACGACGTTGTA-3′) and M13 (-46) reverse (5′-GAGCGGATAACAATTTCACACAGG-3′). Prior to capillary electrophoresis, agarose (10 μL) was added (0.06% final concentration) to the sequencing products resuspended in 70% formamide/1 mM EDTA (10 μL), as previously described (Vatcher et al., 2002; Almira et al., 2003). Sequencing reactions were analyzed with a MegaBACE 1000 automatic sequencer (GE Healthcare) using the following parameters: injection at 3 kV for 50 s and electrophoresis at 6 kV for 200 min. Automated base-calling was performed using Cimarron 3.12 software, and complete sequences were assembled using the Phred/Phrap/Consed package (Ewing et al., 1998; Ewing and Green, 1998; Gordon et al., 1998).

5.5. Sequence analysis

The Translate tool on the ExPASy Proteomics Server was used to translate DNA sequences to protein sequences (Gasteiger et al., 2003). Multiple alignments of DNA and amino acid sequences were usually performed using the program ClustalW (Thompson et al., 1994) implemented with the BioEdit 7.2.5 software package (Hall, 1999). Searches for homologous proteins in public sequence databases were performed using BLASTp (Altschul et al., 1990). The presence and delimitation of protein domains was performed by searching the NCBI Conserved Domain Database (CDD) through the CD-Search web service (Marchler-Bauer and Bryant, 2004).

5.6. Phylogenetic analysis

Phylogenetic analysis was performed using Molecular Evolutionary Genetics Analysis (MEGA) software, version 6.0 (Tamura et al., 2013). Amino acid sequences were aligned using the program Clustal Omega (Sievers et al., 2011) at the web server www.ebi.ac.uk/Tools/msa/clustalo/. The alignment was generated using the default parameters, except that the number of combined iterations, the maximum number of guide tree iterations, and the maximum number of HMM iterations were all set to five. Sites containing gaps at the N- and C-terminal ends of the aligned sequences were trimmed off, and the pairwise evolutionary distances were then computed using the Poisson correction method (Zuckerkandl and Pauling, 1965). The trees were generated using the neighbor-joining method (Saitou and Nei, 1987), and the stability of the clades was assessed using the bootstrap method (Felsenstein, 1985).

5.7. Protein purification, crystallization, structure resolution and refinement

CpOsm was purified to homogeneity from the latex of *C. procera*, as previously described (de Freitas et al., 2011a). Freeze-dried samples of pure CpOsm were dissolved (8.5 mg mL^{-1}) in 10 mM sodium phosphate buffer (pH 6.0), centrifuged at 12,000*g* for 10 min at 25 °C, and the supernatant was used in crystallization trials, as described elsewhere (Bruno-Moreno et al., 2013). The best crystals were obtained at 35% 2-methyl-2,4-pentanediol (MPD) containing 0.7 M $(NH_4)_2SO_4$. X-ray diffraction data were collected at 100 K at beam line I24 of the Diamond Light Source (DLS, Oxfordshire, United Kingdom) using $\lambda = 0.9686$ Å and a Pilatus detector (Pilatus3 6 M). The data were indexed, integrated and scaled using XDS (Kabsch, 2010) and SCALA (Evans, 2006) software and the last program implemented in the CCP4 program suite 6.2.0 (Collaborative Computational Project, Number 4, 1994).

The CpOsm structure was solved by molecular replacement, employing the Phaser (McCoy, 2007) program and using the *N. tabacum* osmotin (PDB ID: 1PCV) as a search model (the amino acid sequence identity between CpOsm and the tobacco osmotin is *ca* 72%), which was modified using the CHAINSAW program (Stein, 2008). The refinement was conducted using the Phenix (Adams et al., 2010) and Coot (Emsley and Cowtan, 2004) programs. The stereochemical quality of the model was evaluated with PROCHECK (Laskowski et al., 1993) and MolProbity (Chen et al., 2010). The data collection and refinement statistics are summarized in Table 1. The structure has been deposited in the Protein Data Bank under the ID 4L2J. Three-dimensional molecular images were produced and rendered using the PyMOL Molecular Graphics System, version 1.7.4 (Schrödinger, LLC). Electrostatic surface calculations were conducted using the PDB2PQR (Dolinsky et al., 2004) and APBS (Baker et al., 2001) tools implemented in PyMOL.

5.8. Capillary liquid chromatography/nanoelectrospray ionization tandem mass spectrometry (LC–ESI–MS/MS)

Protein samples were dissolved (1 mg mL^{-1}) in deionized H_2O, reduced with 5 mM DTT at 60 °C for 30 min, treated with 15 mM iodoacetamide at room temperature for 30 min in the dark, and digested with sequencing-grade trypsin (Promega) at 37 °C for 16 h. The tryptic peptides were analyzed by LC–ESI–MS/MS using a Synapt G1 HDMS Q-ToF mass spectrometer (Waters Co., Milford, MA, USA) coupled to a Waters ultra-high-performance liquid chromatography (UPLC) unit. The digested samples were injected using the nanoACQUITY UPLC sample manager, and the chromatography was performed using a UPLC C18 nanoACQUITY column (75 μm × 10 cm, 1.7 μm particle size) at a flow rate of 0.35 μL/min. The mass spectra were acquired using the data-dependent acquisition (DDA) mode, in which the top three peaks were subjected to MS/MS. Mobile phases A and B consisted of 0.1% HCO_2H in H_2O and 0.1% HCO_2H in CH_3CN, respectively. The peptides were eluted using the following step gradient: 3–40% B for 30 min and 40–85% B for 5 min. The data were processed using Protein Lynx Global Server software (Waters Co.) and subjected to a database search using the Mascot search engine (Perkins et al., 1999). The searches were performed with the assumptions that there was a maximum of one missed trypsin cleavage and that the experimental masses of the peptides were monoisotopic. Furthermore, carbamidomethylation of cysteine was included as a fixed modification, whereas the oxidation of methionine was included as a possible variable modification. MS/MS ions searches were performed against the NCBI non-redundant database (last accessed on February 11, 2015) using a significance threshold of $p < 0.05$. The peptide mass tolerance and fragment mass tolerance were both initially set to ±0.1 Da for MS/MS ion searching. However, candidate peptide IDs were only accepted if the m/z values were within 0.1 Da (typically less than 0.05 Da) of the theoretical mass of the candidate ID, as determined when manually reviewing the MASCOT search results.

5.9. Atomic force microscopy (AFM) of F. solani spores treated with CpOsm

Pure CpOsm was dissolved (0.12 μg mL^{-1}) in distilled H_2O and centrifuged at 10,000*g* for 10 min at 25 °C. An aliquot (50 μL) of the supernatant was added to an equal volume of a suspension of *F. solani* spores (2×10^5 spores mL^{-1}), which was prepared as previously described (de Freitas et al., 2011b), and the mixture was incubated for 30 min at 25 °C. AFM images were acquired by a Multimode Nanoscope IIIa equipment (Bruker, Santa Barbara, CA, USA) using the intermittent (or tapping) mode. In this scanning mode, the AFM tip touches the sample surface with beats of similar amplitude to the resonance amplitude of the AFM cantilever. This minimizes the action of lateral forces between the tip and the sample, which prevents the specimen from being dragged or damaged. Rectangular probes of the model TESP (Bruker) with a nominal spring constant of $k = 40$ N/m, a resonance frequency of $f = 320$ kHz, and a nominal tip radius of 8 nm were used. The scan parameters for the acquisition of the nanoparticle images were as follows: 0.5 Hz scan rate and 512 × 512 lines of resolution.

Acknowledgments

This work was supported by research grants from Conselho Nacional de Desenvolvimento Científico e Tecnológico (CNPq), Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) and Fundação Cearense de Apoio ao Desenvolvimento Científico e Tecnológico (FUNCAP). M.V.R. and T.B.G. are senior investigators of CNPq.

References

Abad, L.R., DUrzo, M.P., Liu, D., Narasimhan, M.L., Reuveni, M., Zhu, J.K., Niu, X.M., Singh, N.K., Hasegawa, P.M., Bressan, R.A., 1996. Antifungal activity of tobacco osmotin has specificity and involves plasma membrane permeabilization. Plant Sci. 118, 11–23.

Adams, P.D., Afonine, P.V., Bunkóczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.-W., Kapral, G.J., Grosse-Kunstleve, R.W., McCoy, A.J., Moriarty, N.W., Oeffner, R., Read, R.J., Richardson, D.C., Richardson, J.S., Terwilliger, T.C., Zwart, P.H., 2010. PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. D Biol. Crystallogr. 66, 213–221.

Almira, E.C., Panayotova, N., Farmerie, W.G., 2003. Capillary DNA sequencing: maximizing the sequence output. J. Biomol. Tech. 14, 270–277.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., Lipman, D.J., 1990. Basic local alignment search tool. J. Mol. Biol. 215, 403–410.

Baker, N.A., Sept, D., Joseph, S., Holst, M.J., McCammon, J.A., 2001. Electrostatics of nanosystems: application to microtubules and the ribosome. Proc. Natl. Acad. Sci. U.S.A. 98, 10037–10041.

Batalia, M.A., Monzingo, A.F., Ernst, S., Roberts, W., Robertus, J.D., 1996. The crystal structure of the antifungal protein zeamatin, a member of the thaumatin-like, PR-5 protein family. Nat. Struct. Biol. 3, 19–23.

Beezer, A.E., Sharma, P.B., 1981. On the uptake of nystatin by *Saccharomyces cerevisiae*. 3. Electrochemistry of the yeast cell surface. Microbios 31, 71–82.

Bruno-Moreno, F., Basílio, Sombra., de Oliveira, R., de Azevedo Moreira, R., Pinto Lobo, M.D., Teixeira de Freitas, C.D., Viana Ramos, M., Barbosa Grangeiro, T., Oliveira Monteiro-Moreira, A.C., 2013. Crystallization and X-ray diffraction analysis of an antifungal laticifer protein. Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun. 69, 646–649.

Cantino, P.D., Doyle, J.A., Graham, S.W., Judd, W.S., Olmstead, R.G., Soltis, D.E., Soltis, P.S., Donoghue, M.J., 2007. Towards a phylogenetic nomenclature of Tracheophyta. Taxon 56, 822–846.

Chen, V.B., Arendall, W.B., Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S., Richardson, D.C., 2010. MolProbity: all-atom structure validation for macromolecular crystallography. Acta Crystallogr. D Biol. Crystallogr. 66, 12–21.

Collaborative Computational Project, Number 4, 1994. The CCP4 suite: programs for protein crystallography. Acta Crystallogr. D Biol. Crystallogr. 50, 760–763.

Cornelissen, B.J., Hooft van Huijsduijnen, R.A., Bol, J.F., 1986. A tobacco mosaic virus-induced tobacco protein is homologous to the sweet-tasting protein thaumatin. Nature 321, 531–532.

Dall'Antonia, Y., Pavkov, T., Fuchs, H., Breiteneder, H., Keller, W., 2005. Crystallization and preliminary structure determination of the plant food allergen Pru av 2. Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun. 61, 186–188.

de Freitas, C.D.T., Nogueira, F.C.S., Vasconcelos, I.M., Oliveira, J.T.A., Domont, G.B., Ramos, M.V., 2011a. Osmotin purified from the latex of *Calotropis procera*: biochemical characterization, biological activity and role in plant defense. Plant Physiol. Biochem. 49, 738–743.

de Freitas, C.D.T., de S. Lopes, J.L., Beltramini, L.M., de Oliveira, R.S.B., Oliveira, J.T.A., Ramos, M.V., 2011b. Osmotin from *Calotropis procera* latex: new insights into structure and antifungal properties. Biochim. Biophys. Acta 1808, 2501–2507.

Dolinsky, T.J., Nielsen, J.E., McCammon, J.A., Baker, N.A., 2004. PDB2PQR: an automated pipeline for the setup of Poisson–Boltzmann electrostatics calculations. Nucleic Acids Res. 32, W665–W667.

Drinnan, A.N., Crane, P.R., Hoot, S.B., 1994. Patterns of floral evolution in the early diversification of non-magnoliid dicotyledons (eudicots). In: Early Evolution of Flowers, Plant Systematics and Evolution, Suppl. 8. Springer-Verlag, New Delhi, pp. 93–122.

Eisenberg, D., Schwarz, E., Komaromy, M., Wall, R., 1984. Analysis of membrane and surface protein sequences with the hydrophobic moment plot. J. Mol. Biol. 179, 125–142.

Emsley, P., Cowtan, K., 2004. Coot: model-building tools for molecular graphics. Acta Crystallogr. D Biol. Crystallogr. 60, 2126–2132.

Evans, P., 2006. Scaling and assessment of data quality. Acta Crystallogr. D Biol. Crystallogr. 62, 72–82.

Ewing, B., Green, P., 1998. Base-calling of automated sequencer traces using phred. II. Error probabilities. Genome Res. 8, 186–194.

Ewing, B., Hillier, L., Wendl, M.C., Green, P., 1998. Base-calling of automated sequencer traces using phred. I. Accuracy assessment. Genome Res. 8, 175–185.

Felsenstein, J., 1985. Confidence-limits on phylogenies – an approach using the bootstrap. Evolution 39, 783–791.

Fierens, E., Rombouts, S., Gebruers, K., Goesaert, H., Brijs, K., Beaugrand, J., Volckaert, G., Van Campenhout, S., Proost, P., Courtin, C.M., Delcour, J.A., 2007. TLXI, a novel type of xylanase inhibitor from wheat (*Triticum aestivum*) belonging to the thaumatin family. Biochem. J. 403, 583–591.

Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R.D., Bairoch, A., 2003. ExPASy: the proteomics server for in-depth protein knowledge and analysis. Nucleic Acids Res. 31, 3784–3788.

Ghosh, R., Chakrabarti, C., 2008. Crystal structure analysis of NP24-I: a thaumatin-like protein. Planta 228, 883–890.

Gordon, D., Abajian, C., Green, P., 1998. Consed: a graphical tool for sequence finishing. Genome Res. 8, 195–202.

Guo, Z.-J., Chen, X.-J., Wu, X.-L., Ling, J.-Q., Xu, P., 2004. Overexpression of the AP2/EREBP transcription factor OPBP1 enhances disease resistance and salt tolerance in tobacco. Plant Mol. Biol. 55, 607–618.

Hall, T.A., 1999. BioEdit: a user-friendly biological sequence alignment editor and analysis program for Windows 95/98/NT. Nucleic Acids Symp. Ser. 41, 95–98.

Ibeas, J.I., Lee, H., Damsz, B., Prasad, D.T., Pardo, J.M., Hasegawa, P.M., Bressan, R.A., Narasimhan, M.L., 2000. Fungal cell wall phosphomannans facilitate the toxic activity of a plant PR-5 protein. Plant J. 23, 375–383.

Kabsch, W., 2010. XDS. Acta Crystallogr. D Biol. Crystallogr. 66, 125–132.

Kabsch, W., Sander, C., 1983. Dictionary of protein secondary structure: pattern recognition of hydrogen-bonded and geometrical features. Biopolymers 22, 2577–2637.

Koiwa, H., Kato, H., Nakatsu, T., Oda, J., Yamada, Y., Sato, F., 1999. Crystal structure of tobacco PR-5d protein at 1.8 A resolution reveals a conserved acidic cleft structure in antifungal thaumatin-like proteins. J. Mol. Biol. 286, 1137–1145.

Laskowski, R.A., 1995. SURFNET: a program for visualizing molecular surfaces, cavities, and intermolecular interactions. J. Mol. Graph. 13 (323–330), 307–308.

Laskowski, R.A., MacArthur, M.W., Moss, D.S., Thornton, J.M., 1993. PROCHECK: a program to check the stereochemical quality of protein structures. J. Appl. Crystallogr. 26, 283–291.

Lee, E.K., Cibrian-Jaramillo, A., Kolokotronis, S.-O., Katari, M.S., Stamatakis, A., Ott, M., Chiu, J.C., Little, D.P., Stevenson, D.W., McCombie, W.R., Martienssen, R.A., Coruzzi, G., Desalle, R., 2011. A functional phylogenomic view of the seed plants. PLoS Genet. 7, e1002411.

Leone, P., Menu-Bouaouiche, L., Peumans, W.J., Payan, F., Barre, A., Roussel, A., Van Damme, E.J.M., Rougé, P., 2006. Resolution of the structure of the allergenic and antifungal banana fruit thaumatin-like protein at 1.7-Å. Biochimie 88, 45–52.

Lin, K.C., Bushnell, W.R., Szabo, L.J., Smith, A.G., 1996. Isolation and expression of a host response gene family encoding thaumatin-like proteins in incompatible oat–stem rust fungus interactions. Mol. Plant Microbe Interact. 9, 511–522.

Liu, D., Raghothama, K.G., Hasegawa, P.M., Bressan, R.A., 1994. Osmotin overexpression in potato delays development of disease symptoms. Proc. Natl. Acad. Sci. U.S.A. 91, 1888–1892.

Marangon, M., Van Sluyter, S.C., Waters, E.J., Menz, R.I., 2014. Structure of haze forming proteins in white wines: *Vitis vinifera* thaumatin-like proteins. PLoS One 9, e113757.

Marchler-Bauer, A., Bryant, S.H., 2004. CD-Search: protein domain annotations on the fly. Nucleic Acids Res. 32, W327–W331.

Masuda, T., Ohta, K., Mikami, B., Kitabatake, N., 2011a. High-resolution structure of the recombinant sweet-tasting protein thaumatin I. Acta Crystallogr. Sect. F Struct. Biol. Cryst. Commun. 67, 652–658.

Masuda, T., Ohta, K., Tani, F., Mikami, B., Kitabatake, N., 2011b. Crystal structure of the sweet-tasting protein thaumatin II at 1.27 Å. Biochem. Biophys. Res. Commun. 410, 457–460.

McCoy, A.J., 2007. Solving structures of protein complexes by molecular replacement with Phaser. Acta Crystallogr. D Biol. Crystallogr. 63, 32–41.

Min, K., Ha, S.C., Hasegawa, P.M., Bressan, R.A., Yun, D.-J., Kim, K.K., 2004. Crystal structure of osmotin, a plant antifungal protein. Proteins 54, 170–173.

Perkins, D.N., Pappin, D.J., Creasy, D.M., Cottrell, J.S., 1999. Probability-based protein identification by searching sequence databases using mass spectrometry data. Electrophoresis 20, 3551–3567.

Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C., Ferrin, T.E., 2004. UCSF Chimera – a visualization system for exploratory research and analysis. J. Comput. Chem. 25, 1605–1612.

Pierpoint, W., Tatham, A., Pappin, D., 1987. Identification of the virus-induced protein of tobacco-leaves that resembles the sweet-protein thaumatin. Physiol. Mol. Plant Pathol. 31, 291–298.

Roberts, W.K., Selitrennikoff, C.P., 1990. Zeamatin, an antifungal protein from maize with membrane-permeabilizing activity. J. Gen. Microbiol. 136, 1771–1778.

Ruperti, B., Cattivelli, L., Pagni, S., Ramina, A., 2002. Ethylene-responsive genes are differentially regulated during abscission, organ senescence and wounding in peach (*Prunus persica*). J. Exp. Bot. 53, 429–437.

Saitou, N., Nei, M., 1987. The neighbor-joining method: a new method for reconstructing phylogenetic trees. Mol. Biol. Evol. 4, 406–425.

Sakamoto, Y., Watanabe, H., Nagai, M., Nakade, K., Takahashi, M., Sato, T., 2006. *Lentinula edodes* tlg1 encodes a thaumatin-like protein that is involved in lentinan degradation and fruiting body senescence. Plant Physiol. 141, 793–801.

Sambrook, J., Fritsch, E., Maniatis, T., 1989. Molecular Cloning: A Laboratory Manual, 2nd ed. Cold Spring Harbor Laboratory Press, Cold Spring Harbor.

Sievers, F., Wilm, A., Dineen, D., Gibson, T.J., Karplus, K., Li, W., Lopez, R., McWilliam, H., Remmert, M., Söding, J., Thompson, J.D., Higgins, D.G., 2011. Fast, scalable generation of high-quality protein multiple sequence alignments using Clustal Omega. Mol. Syst. Biol. 7, 539.

Sillitoe, I., Cuff, A.L., Dessailly, B.H., Dawson, N.L., Furnham, N., Lee, D., Lees, J.G., Lewis, T.E., Studer, R.A., Rentzsch, R., Yeats, C., Thornton, J.M., Orengo, C.A., 2013. New functional families (FunFams) in CATH to improve the mapping of conserved functional sites to 3D structures. Nucleic Acids Res. 41, D490–D498.

Singh, N.K., Bracker, C.A., Hasegawa, P.M., Handa, A.K., Buckel, S., Hermodson, M.A., Pfankoch, E., Regnier, F.E., Bressan, R.A., 1987. Characterization of osmotin: a thaumatin-like protein associated with osmotic adaptation in plant cells. Plant Physiol. 85, 529–536.

Stein, N., 2008. CHAINSAW: a program for mutating pdb files used as templates in molecular replacement. J. Appl. Crystallogr. 41, 641–643.

Stivala, A., Wybrow, M., Wirth, A., Whisstock, J.C., Stuckey, P.J., 2011. Automatic generation of protein structure cartoons with Pro-origami. Bioinformatics 27, 3315–3316.

Tamura, K., Stecher, G., Peterson, D., Filipski, A., Kumar, S., 2013. MEGA6: Molecular Evolutionary Genetics Analysis version 6.0. Mol. Biol. Evol. 30, 2725–2729.

Thompson, J.D., Higgins, D.G., Gibson, T.J., 1994. CLUSTAL W: improving the sensitivity of progressive multiple sequence alignment through sequence weighting, position-specific gap penalties and weight matrix choice. Nucleic Acids Res. 22, 4673–4680.

Trudel, J., Grenier, J., Potvin, C., Asselin, A., 1998. Several thaumatin-like proteins bind to beta-1,3-glucans. Plant Physiol. 118, 1431–1438.

Van der Wel, H., Loeve, K., 1972. Isolation and characterization of thaumatin I and II, the sweet-tasting proteins from *Thaumatococcus daniellii* Benth. Eur. J. Biochem. 31, 221–225.

Van Loon, L.C., Gerritsen, Y.A., Ritter, C.E., 1987. Identification, purification, and characterization of pathogenesis-related proteins from virus-infected Samsun NN tobacco leaves. Plant Mol. Biol. 9, 593–609.

Van Loon, L.C., Rep, M., Pieterse, C.M.J., 2006. Significance of inducible defense-related proteins in infected plants. Annu. Rev. Phytopathol. 44, 135–162.

Vandermarliere, E., Lammens, W., Schoepe, J., Rombouts, S., Fierens, E., Gebruers, K., Volckaert, G., Rabijns, A., Delcour, J.A., Strelkov, S.V., Courtin, C.M., 2010. Crystal structure of the noncompetitive xylanase inhibitor TLXI, member of the small thaumatin-like protein family. Proteins 78, 2391–2394.

Vatcher, G., Smailus, D., Krzywinski, M., Guin, R., Stott, J., Tsai, M., Chan, S., Pandoh, P., Yang, G., Asano, J., Olson, T., Prabhu, A.L., Coope, R., Marziali, A., Schein, J., Jones, S., Marra, M., 2002. Resuspension of DNA sequencing reaction products in agarose increases sequence quality on an automated sequencer. BioTechniques 33, 532–539.

Wickett, N.J., Mirarab, S., Nguyen, N., Warnow, T., Carpenter, E., Matasci, N., Ayyampalayam, S., Barker, M.S., Burleigh, J.G., Gitzendanner, M.A., Ruhfel, B.R., Wafula, E., Der, J.P., Graham, S.W., Mathews, S., Melkonian, M., Soltis, D.E., Soltis, P.S., Miles, N.W., Rothfels, C.J., Pokorny, L., Shaw, A.J., DeGironimo, L., Stevenson, D.W., Surek, B., Villarreal, J.C., Roure, B., Philippe, H., dePamphilis, C.W., Chen, T., Deyholos, M.K., Baucom, R.S., Kutchan, T.M., Augustin, M.M., Wang, J., Zhang, Y., Tian, Z., Yan, Z., Wu, X., Sun, X., Wong, G.K.-S., Leebens-Mack, J., 2014. Phylotranscriptomic analysis of the origin and early diversification of land plants. Proc. Natl. Acad. Sci. U.S.A. 111, E4859–E4868.

Woloshuk, C.P., Meulenhoff, J.S., Sela-Buurlage, M., van den Elzen, P.J., Cornelissen, B.J., 1991. Pathogen-induced proteins with inhibitory activity toward *Phytophthora infestans*. Plant Cell 3, 619–628.

Yun, D.J., Zhao, Y., Pardo, J.M., Narasimhan, M.L., Damsz, B., Lee, H., Abad, L.R., D'Urzo, M.P., Hasegawa, P.M., Bressan, R.A., 1997. Stress proteins on the yeast cell surface determine resistance to osmotin, a plant antifungal protein. Proc. Natl. Acad. Sci. U.S.A. 94, 7082–7087.

Zeng, L., Zhang, Q., Sun, R., Kong, H., Zhang, N., Ma, H., 2014. Resolution of deep angiosperm phylogeny using conserved nuclear genes and estimates of early divergence times. Nat. Commun. 5, 4956.

Zhu, B., Chen, T.H., Li, P.H., 1996. Analysis of late-blight disease resistance and freezing tolerance in transgenic potato plants expressing sense and antisense genes for an osmotin-like protein. Planta 198, 70–77.

Zuckerkandl, E., Pauling, L., 1965. Evolutionary divergence and convergence in proteins. In: Evolving Genes and Proteins. Academic Press, New York, pp. 97–166.