The Structure of the Elicitor Cerato-platanin (CP), the FirstMember of the CP Fungal Protein Family, Reveals a Double-Barrel Fold and

THE JOURNAL OF BIOLOGICAL CHEMISTRY VOL. 286, NO. 20, pp. 17560–17568, May 20, 2011

The Structure of the Elicitor Cerato-platanin (CP), the First Member of the CP Fungal Protein Family, Reveals a Double ψβ-Barrel Fold and Carbohydrate Binding*[S]

Received for publication, January 20, 2011, and in revised form, March 14, 2011 Published, JBC Papers in Press, March 30, 2011, DOI 10.1074/jbc.M111.223644

Aline L. de Oliveira[‡1,2], Mariana Gallo[§1,3], Luigia Pazzagli[¶], Celso E. Benedetti[‡4], Gianni Cappugi[¶], Aniello Scala[‖], Barbara Pantera[¶], Alberto Spisni[5], Thelma A. Pertinhez[**], and Daniel O. Cicero[§6]**

*From the [‡]National Laboratory of Biosciences, 13083-100 Campinas, Brazil, the [§]Department of Chemical Science and Technology, University of Rome "Tor Vergata," 00133 Rome, Italy, the Departments of [¶]Biochemical Sciences and [‖]Agricultural Biotechnologies, University of Florence, 50134 Florence, Italy, and the **Department of Experimental Medicine, University of Parma, 43125 Parma, Italy*

Cerato-platanin (CP) is a secretion protein produced by the fungal pathogen *Ceratocystis platani*, the causal agent of the plane canker disease and the first member of the CP family. CP is considered a pathogen-associated molecular pattern because it induces various defense responses in the host, including production of phytoalexins and cell death. Although much is known about the properties of CP and related proteins as elicitors of plant defense mechanisms, its biochemical activity and host target(s) remain elusive. Here, we present the three-dimensional structure of CP. The protein, which exhibits a remarkable pH and thermal stability, has a double ψβ-barrel fold quite similar to those found in expansins, endoglucanases, and the plant defense protein barwin. Interestingly, although CP lacks lytic activity against a variety of carbohydrates, it binds oligosaccharides. We identified the CP region responsible for binding as a shallow surface located at one side of the β-barrel. Chemical shift perturbation of the protein amide protons, induced by oligo-*N*-acetylglucosamines of various size, showed that all the residues involved in oligosaccharide binding are conserved among the members of the CP family. Overall, the results suggest that CP might be involved in polysaccharide recognition and that the double ψβ-barrel fold is widespread in distantly related organisms, where it is often involved in host-microbe interactions.

Plants are often infected by microbes with diverse lifestyles, and to defend themselves, they have developed a two-layered innate immune system (1). The first line provides a basal defense against all potential pathogens and relies on the recognition either of conserved pathogen-associated molecular patterns (PAMPs),[7] including flagellin, peptidoglycans, chitin, and β-glucan, some of the proteins relevant for the life cycle of the pathogen (2), or of an array of cell wall-degrading enzymes not necessarily involved in pathogenicity (3). This defense line leads to cell wall alterations, deposition of callose, induction of reactive oxygen species synthesis, activation of MAPK cascades, and accumulation of defense-related proteins (4).

Nevertheless, some pathogens may succeed in suppressing this primary defense mechanism. Thus, during evolution, plants have developed a second recognition system based on the perception of the effectors by R proteins, with subsequent activation of an effector-triggered immunity, which often ends up in a localized programmed cell death response (5–7).

Phytopathogenic fungi and oomycetes secrete several non-catalytic proteins (generally characterized by low molecular weight and by the presence of two to four disulfide bridges) that belong to specific protein families: elicitin, hydrophobin, NIP-1, and cerato-platanin (CP). Some members of these families have been proposed to have PAMP activity, *e.g.* cryptogein, INF1 elicitin, and the Pep1 and Stp1 hydrophobins (8–10).

The members of the CP protein family are secreted by a number of phytopathogenic fungi, animal pathogens, and non-pathogenic fungal species (supplemental Fig. S1). These proteins have been shown to act either as virulence factors and effectors, such as Snodprot1 from *Phaeosphaeria nodorum*, Sp1 from *Leptosphaeria maculans*, and MpCP1 from *Moniliophthora perniciosa* (11–13), or either as elicitors and PAMPs, such as

* This work was supported in part by grants from the Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) (NMR Project 99/11030-9, TAP 00/02026-7, Smolbnet 00/10266-8) and from the Ministero dell'Istruzione dell'Università e della Ricerca (MIUR) (PRIN 2005 20050717.5-002). Some of the CD and NMR experiments were carried out using equipment at the Centro Interfacoltà Misure, University of Parma.

[S] The on-line version of this article (available at http://www.jbc.org) contains supplemental Figs. S1–S5 and additional references.

The atomic coordinates and structure factors (code 2KQA) have been deposited in the Protein Data Bank, Research Collaboratory for Structural Bioinformatics, Rutgers University, New Brunswick, NJ (http://www.rcsb.org/).

[1] Both authors contributed equally to this work.

[2] Supported by Fundação de Amparo à Pesquisa do Estado de São Paulo (FAPESP) Fellowship 03/12418-8 and Coordenação de Aperfeiçoamento de Pessoal de Nível Superior (CAPES) Fellowship BEX3666/05-0. Present address: Inst. de Química, Universidade Federal de Goiás, 74001-970 Goiânia, Brazil.

[3] Present address: Fundación Instituto Leloir, C1405BWE Buenos Aires, Argentina.

[4] Supported by a fellowship from the Conselho Nacional de Pesquisa e Tecnologia (CNPq).

[5] To whom correspondence may be addressed: Università degli Studi di Parma, Via Volturno 39, 43125 Parma, Italy. Tel.: 39-052-103-3807; Fax: 39-052-103-3802; E-mail: alberto.spisni@unipr.it.

[6] To whom correspondence may be addressed: Dipt. di Scienze e Tecnologie Chimiche, Università degli Studi di Roma "Tor Vergata," Via della Ricerca Scientifica 1, 00133 Rome, Italy. Tel.: 39-067-259-4835; E-mail: cicero@scienze.uniroma2.it.

[7] The abbreviations used are: PAMP, pathogen-associated molecular pattern; CP, cerato-platanin; RDC, residual dipolar constant; LT, lytic transglycosylase.

the proteins Sm1 and Epl1 from the soil-borne *Trichoderma virens* and *Hypocrea atroviridis*, respectively (14, 15). In particular, Sm1 elicits the production of reactive oxygen species and triggers the expression of several defense-related genes in plants, confirming the capacity of this biocontrol fungus to induce resistance, whereas Epl1 is the major protein found in the secretome of the biocontrol fungal strain *H. atroviridis* (14–16). Moreover, *Magnaporthe grisea* expresses a MgSM1-CP-like protein that is able to induce a hypersensitive response in leaves and enhanced disease resistance against *Botrytis cinerea* and *Pseudomonas syringae* when transiently expressed in *Arabidopsis* (17).

In addition, members of the CP family play a major role in animal-fungus interaction, where they exhibit allergenic properties and induce strong immunological reactions. This is the case of the *Coccidioides*-specific antigen from *Coccidioides immitis* and the Aspf13 antigen from *Aspergillum fumigatus*, both causative agents of human lung diseases (18, 19), and of the immunomodulatory protein Aca1 from the medicinal fungus *Antrodia camphorata* (20).

The first member of the CP family is the CP protein from *Ceratocystis platani*, the causal agent of the canker stain disease of *Platanus acerifolia* (21, 22). CP plays a role in host-fungus interaction because it induces cell death, phytoalexin synthesis, and overexpression of defense-related genes (23–25). For this reason, CP and the orthologous protein cerato-populin from *Ceratocystis populicola*, which triggers a CP-like reaction in host-microbe interaction (26, 27), have been considered PAMPs. CP can be found in the cell wall of ascospores, hyphae, and conidia, suggesting that the protein plays a role in forming the fungal cell wall (28). This last observation supports the notion that PAMPs are generally surface-associated molecules that cannot be easily altered without deeply interfering with the fungus fitness (10).

Until now, no information on the structure of CP or other CP-like proteins was available. In fact, only few examples of fungal PAMPs have been structurally characterized (10). Here, we report the NMR-derived solution structure of the first member of the CP family, the CP protein. CP has a globular fold containing two α-helices and six β-strands that form a six-stranded double $\psi\beta$-barrel, a motif remarkably similar to the those present in endoglucanases, domain I of expansins, and the defense protein barwin. Because of these similarities, the oligosaccharide-binding properties and the glycosidase activities of CP have been explored.

EXPERIMENTAL PROCEDURES

Expression and Purification of Cerato-platanin

The recombinant CP protein was expressed in *Pichia pastoris* and purified by reverse-phase liquid chromatography as described previously (29). The recombinant CP protein used in this work presents an additional 9-residue N-terminal stretch (EEGVSLEKR) because, under our experimental conditions, the signal peptide used for extracellular expression turned out not to be cut at the expected site. However, the protein shows the same biological activity and the same CD and ^{1}H NMR spectra of native CP (29).

For isotopic labeling, CP was expressed in minimal medium supplemented with [^{15}N]ammonium sulfate and, in the case of double-labeled proteins, with [^{13}C]glycerol and [^{13}C]methanol as described previously (30). The purified protein was estimated to be ~95% pure by SDS-PAGE, and MALDI-TOF mass spectrometry confirmed that it was 100% ^{13}C- and ^{15}N-isotopically labeled.

NMR Measurements

NMR samples contained 0.7 mM double-labeled [^{13}C,^{15}N]CP in 10 mM phosphate buffer (pH 5.8), 0.05% NaN_3, and 5 or 100% D_2O (Cambridge Isotope Laboratories, Inc.). Two- and three-dimensional NMR experiments were carried out at 20 °C on 500- and 600-MHz Varian Inova AS spectrometers (National Laboratory of Biosciences), a 600-MHz Varian Inova AS spectrometer (Centro Interfacoltà Misure, University of Parma), and 700- and 400-MHz Bruker AVANCE spectrometers (University of Rome "Tor Vergata"). All NMR spectrometers were equipped with a *z* pulsed field gradient unit and a triple resonance probe. Sequence schemes employing pulsed field gradients were used to achieve suppression of the solvent signal and spectral artifacts. Selective pulses to cover aliphatic, aromatic, or carbonyl ^{13}C nuclear spectral zones were obtained by adiabatic pulse modulation. Quadrature detection in the indirectly detected dimensions was obtained by the hybrid States-TPPI (times proportional phase increment) method or the sensitivity-enhanced Echo-Antiecho combination. Linear prediction was applied to extend the indirect ^{13}C-detected dimension. Direct and indirect dimensions were normally apodized using 90°-shifted squared sine-bell functions (for ^{13}C- and ^{15}N-edited dimensions) or Lorentzian-to-Gaussian functions (for the ^{1}H dimension), followed by zero filling and Fourier transform. The NMR data were processed on Silicon Graphics workstations with NMRView/NMRPipe software (31, 32).

NOESY Experiments—NOE-derived distance restraints were obtained from three-dimensional ^{1}H-^{15}N NOESY (33) and three-dimensional ^{1}H-^{13}C NOESY, separately optimized for aliphatic and aromatic side chains (34) collected with 100-ms mixing times. Two-dimensional NOESY was performed to assign the peak at 11.40 ppm, which turned out to be the hydroxyl proton of Tyr-9, which is hydrogen-bonded to the side chain carboxyl of Asp-77.

Residual Dipolar Constant (RDC) Measurements—Five distinct sets of RDCs were measured: $^1J_{HN\text{-}N}$, $^3J_{H\alpha\text{-}C\alpha}$, $^1J_{C\alpha\text{-}C'}$, $^1J_{N\text{-}C'}$, and $^2J_{HN\text{-}C'}$. Anisotropic data were collected by dissolving the protein directly in 14 mg/ml liquid crystalline Pf1 (Asla Lab) as the alignment medium. To decrease the signal width, we added NaCl up to 150 mM. $^3J_{H\alpha\text{-}C\alpha}$ couplings were obtained from the ^{1}H-coupled (F1) version of the HACACO experiment (35). $^1J_{N\text{-}C'}$ and $^2J_{HN\text{-}C'}$ were measured from ^{1}H-^{15}N heteronuclear single quantum coherence spectra with only Cα decoupling (36). $^1J_{HN\text{-}N}$ RDCs were measured by an in-phase/anti-phase heteronuclear single quantum coherence (HSQC) experiment (37). $^1J_{C\alpha\text{-}C'}$ RDCs were measured in a modified version of the J-module HN(COCA) experiment (38, 39).

Diffusion Experiments—Diffusion-ordered spectroscopy experiments were carried out using the BPPSTE (bipolar pulse pairs stimulated echo) method (40). The duration of the total

diffusion phase-encoding gradient pulse was 2 ms, the diffusion delay was 0.05 s, and the minimum gradient strength was set to 0.3 gauss/cm. Diffusion coefficients were measured using 0.2 mM CP in 10 mM phosphate buffer (pH 5.8), and 0.02% dioxane was used as a radius standard (41).

Structure Calculation

Structure calculation was performed with XPLOR-NIH (42) using a simulated annealing protocol. Calculations were carried out over residues 1–120, excluding the first 9 N-terminal residues that do not belong to the native protein (see "Expression and Purification of Cerato-platanin" above) and that turned out to be unstructured (see supplemental Fig. S2). The simulated annealing was driven by NMR-derived constraints during the entire calculation, with force constraints and weighting factors optimized to increase the score of low energy structures. A set of 3203 interproton distance restraints was used, subdivided into three groups: strong (1.8–2.4 Å), medium (1.8–4.0 Å), and weak (1.8–5.0 Å). Because of very weak cross-peaks detected in the ^{15}N-edited NOESY spectrum, we included also a restraint comprising 6.0–3.0 Å. Backbone dihedral restrains (ϕ and ψ angles) were derived from chemical shift analyses using the program TALOS (43). Backbone hydrogen bonds were recognized by evaluating the spatial relationship of amide protons with potential acceptors in the initial structures produced without the use of hydrogen bond constraints. In addition, specific side chain hydrogen bonds were clearly identified using the WHAT IF server (44) with structures generated in later stages of the structural determination and were employed as additional restraints in the last cycle of minimization. Five distinct sets of backbone RDCs (84 $^{1}J_{\mathrm{HN\text{-}N}}$, 71 $^{3}J_{\mathrm{H\alpha\text{-}C\alpha}}$, 60 $^{1}J_{\mathrm{C\alpha\text{-}C'}}$, 57 $^{1}J_{\mathrm{N\text{-}C'}}$, and 52 $^{2}J_{\mathrm{HN\text{-}C'}}$) were included. Fitting the observed RDCs to a preliminary structure, using the singular value decomposition method (45) with the PALES program (46), provided an initial guess of the magnitude and orientation of the molecular alignment tensor. Subsequent optimization during the simulated annealing process led to the following final values for the N-H normalized magnitude of the RDC tensor and rhombicity, respectively: $Da^{\mathrm{NH}} = 5.55$ Hz and $R = 0.625$. A complete cross-validation of RDCs was conducted by carrying out a series of simulated annealing calculations, each one lacking 10% of RDCs randomly chosen from the whole data set. The missing RDCs were back-calculated to evaluate how well each RDC could be predicted; an average R_{free} of 33.6% was obtained (47). An ensemble of the 20 lower energy target function structures (from a total of 1000), in which the number of residues in disallowed regions of the Ramachandran plot is 0.0%, was chosen to represent the protein solution structure and, together with the constraint list, was deposited in the Protein Data Bank (2KQA). The programs AQUA and PROCHECK (48) were used to analyze the structures. Table 1 provides the structural statistics.

Glycosidase Assay

Endo-1,3-β-glucanase, polygalacturonase, and cellulase activities were determined using laminarin, polygalacturonic acid, and carboxymethylcellulose as substrates, respectively, according to previously published methods (49–51). The released reducing sugars were detected using the PAHBAH procedure (52) with glucose as a standard. Enzymes (laminarinase from *Trichoderma* sp., pectinase from *Aspergillus niger*, and cellulase from *Trichoderma reseei*) and substrates blanks were also included. Chitinase activity was detected at 575 nm using chitin azure as a substrate. Chitinase from *Trichoderma viride* was used as a positive control.

TABLE 1

Experimental restraints and structural statistics for the 20 lower energy structures of the CP protein

As indicated under "Experimental Procedures," the data refer to residues 1–120 of the protein and do not consider the extra 9 residues at N terminus. r.m.s.d., root mean square deviation.

No. of experimental restraints	3740
Distance restraints from NOEs	3203
Intraresidue	1094
Sequential	712
Medium-range	406
Long-range	991
Hydrogen bond distance restraints[a]	116
Dihedral angle restraints	97
Residual dipolar coupling constants	324
H-N	84
Hα-Cα	71
Cα-CO	60
H-CO	52
N-CO	57
Average No. of restraints per residue	31
XPLOR energies (kcal/mol)	
E_{total}	2569.5 ± 25.7
E_{bond}	25.7 ± 1.0
E_{angle}	309.5 ± 7.3
E_{improper}	51.3 ± 2.1
E_{vdw}	88.9 ± 8.7
E_{cdih}	0.18 ± 0.22
E_{rama}	1716.9 ± 17.6
E_{noe}	111.4 ± 9.4
E_{sani}	283.6 ± 6.3
r.m.s.d. from experimental restraints	
Average distance restraint violation (Å)	0.041 ± 0.002
Average dihedral angle restraint violation	0.11 ± 0.02°
Average H-N RDC violation (Hz)	0.94 ± 0.03
Average Hα-Cα RDC constant violation (Hz)	0.75 ± 0.02[b]
Average Cα-CO RDC violation (Hz)	2.18 ± 0.07[b]
Average H-CO RDC violation (Hz)	2.66 ± 0.17[b]
Average N-CO RDC violation (Hz)	3.25 ± 0.68[b]
R_{free} (%)	33.6
r.m.s.d. from idealized covalent geometry	
Bond (Å)	0.00372 ± 0.00007
Angle	0.753 ± 0.009°
Improper	0.586 ± 0.012°
Ramachandran analysis[c]	
Residues in favored regions (%)	85.5 ± 1.2
Residues in additionally allowed regions (%)	11.7 ± 1.3
Residues in generously allowed regions (%)	2.8 ± 0.6
Residues in disallowed regions (%)	0.0 ± 0.0
Ramachandran *Z*-score	−1.636 ± 0.331
Coordinate precision[d]	
Backbone (Å)	0.40 ± 0.07
All heavy atoms (Å)	0.72 ± 0.08

[a] HN-O and N-O distances were constrained to 2.1 ± 0.5 and 3.0 ± 0.5 Å, respectively.
[b] Values were normalized with respect to the H-N value.
[c] Residues 1–117.
[d] Values represent r.m.s.d. with respect to the non-minimized average structure (residues 1–117).

N-Acetylglucosamine Binding

The binding of the β-(1,4)-tetramer of *N*-acetyl-D-glucosamine (Sigma) was studied after the addition of 0, 0.1, 2, 4, and 8 mM oligosaccharide to a 0.04 mM solution of ^{15}N-labeled CP. To evaluate the binding of larger oligosaccharides, a chito-oligosaccharide mixture containing oligomers from the dimer to

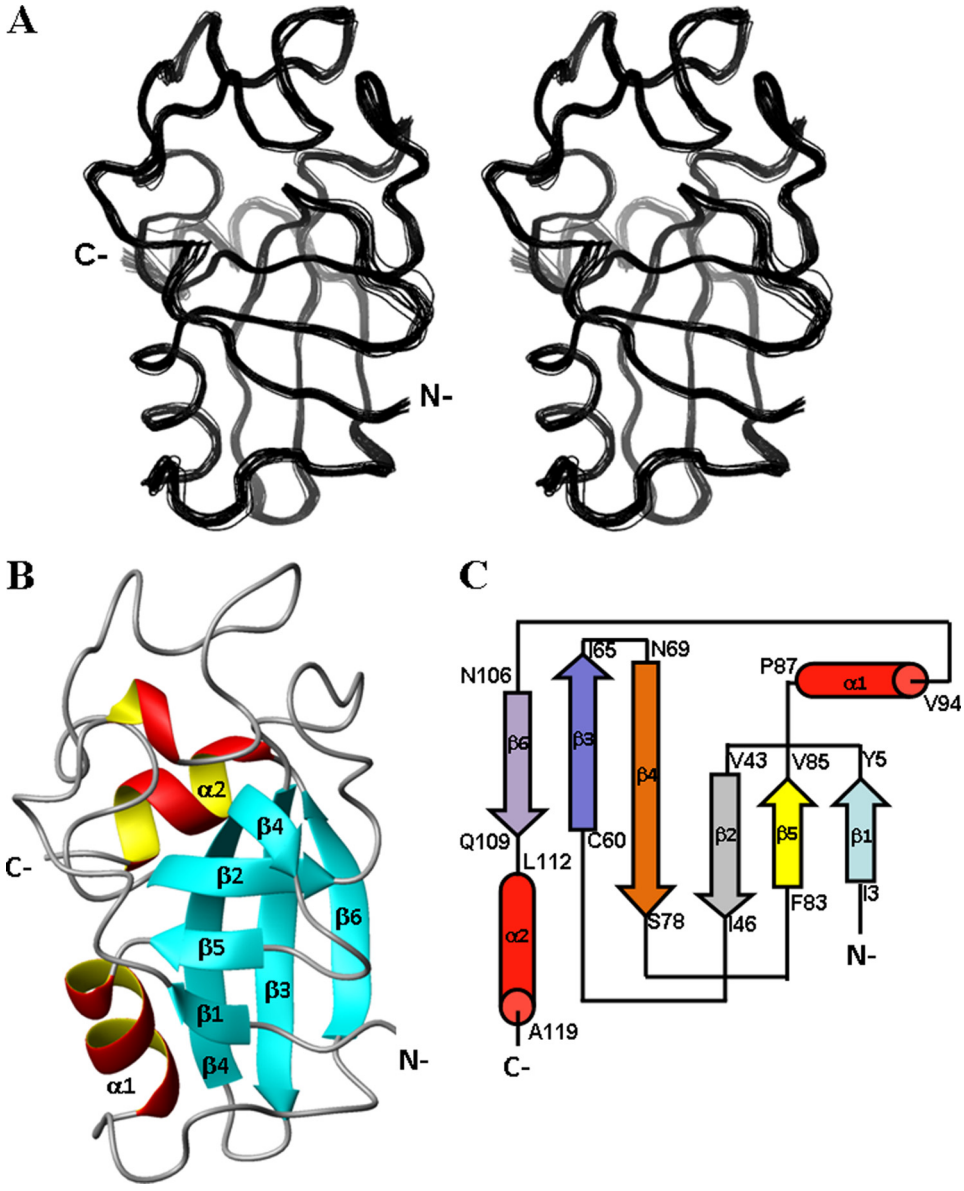

FIGURE 1. **Three-dimensional structure of CP.** *A*, stereo view of the superimposition of the backbone atoms of the 20 lower energy models. *B*, ribbon representation of the structure. *C*, topological representation of CP.

hexamer of β-(1,4)-linked *N*-acetyl-D-glucosamine (Seikagaku Biobusiness Corp.) was used. The mixture (containing 5 mg of each oligosaccharide) was dissolved in 200 μl of H_2O, and aliquots of the solution were added to a 0.1 mM solution of ^{15}N-labeled CP to obtain final concentrations of the hexamer of 0, 0.5, 1.0, 1.5, and 2.0 mM. The chemical shift perturbation in the ^{1}H-^{15}N HSQC spectrum upon oligosaccharide addition was evaluated by the weighted average chemical shift change (53): $\Delta\delta_{av} = ((\Delta\delta^2_{NH} + \Delta\delta^2_{N}/25)/2)^{1/2}$ (ppm).

RESULTS

CP Presents a Double ψβ-Barrel Fold—The three-dimensional structure of CP was solved by triple resonance NMR methods using >3400 distance and dihedral angle restraints together with 324 RDCs in iterative cycles of structure calculations. The quality of the CP structure is highlighted by the structural statistics (Table 1). The protein region Val-1–Asn-117 is defined with high precision as reflected by the low root mean square deviations from the mean structure for the backbone atoms (±0.40 Å) and for all heavy atoms (±0.72 Å) (Fig. 1*A* and Table 1). A quality check using PROCHECK-NMR (48) indicated that 85.5% of the residues display dihedral angles in the allowed regions of the Ramachandran plot (Table 1).

The solution structure of CP reveals a globular fold containing two α-helices and six β-strands forming a double ψβ-barrel (Fig. 1, *B* and *C*). The six-stranded β-barrel has a shear number of 10. The structure is consistent with CD measurements that indicate a predominant β-sheet structure and some α-helical elements (supplemental Fig. S3). In addition, the CP hydrodynamic radius, estimated by diffusion measurements, is ~19.5 Å, in agreement with the value expected for a monomeric and globular protein of that size (54). The long and superficial ψ-loop between β1 and β2 is rigid. In fact, ^{15}N relaxation data show quite homogeneous T_1 and T_2 values and an average

TABLE 2

Results of the pairwise superposition of CP with protein representatives of families bearing a double ψβ-barrel fold using the Dali server

A *Z*-score is a measure of the quality of the alignment. The higher the *Z*-score, the more homologous are the structures. A *Z*-score below 2 is not significant. PDB, Protein Data Bank; r.m.s.d., root mean square deviation; EGV, endoglucanase V.

Protein	Source	PDB entry	*Z*-score[a]	r.m.s.d.	No. of aligned residues	Identity
				Å		%
Expansins						
Phlp1	Timothy grass pollen	1N10	9.5	2.1	91	15
EXPB1	Maize pollen	2HCZ	9.1	3.1	97	20
BsEXLX1	Grass bacillus	2BH0	8.6	2.7	90	16
Plant defense						
Barwin	Barley seed	1BW3	8.1	2.5	99	28
Endoglucanases						
MeCel45A	Blue mussel	1WC2	7.2	2.8	98	17
EGV	*Humicola insolens*	2ENG	5.3	3.2	95	13
maEG	*Melanocarpus albomyces*	1OA9	5.2	3.1	94	10
LTs						
AtMltA	*Agrobacterium tumefaciens*	2PNW	3.6	3.1	74	8
NgMltA	*Neisseria gonorrhoeae*	2G5D	3.3	3.1	75	13
EcMltA	*Escherichia coli*	2AE0	3.2	2.9	71	8

^{1}H-^{15}N NOE of 0.85, indicating reduced internal motions (supplemental Fig. S2). The only slightly flexible region of the protein comprises the sequence Ser-96–Arg-101, which exhibits relatively low ^{1}H-^{15}N NOE and long T_2 values (supplemental Fig. S2). The last 3 C-terminal residues are unstructured and highly mobile (supplemental Fig. S2). Interestingly, CD experiments showed that the protein is quite stable in a wide pH range and exhibits a somewhat high unfolding temperature of ~76 °C (supplemental Fig. S4).

Structures Related to CP—Protein sequence alignments show that CP shares significant similarities only with fungi proteins that are currently classified as members of the CP family (supplemental Fig. S1). However, a search for structurally related proteins using the Dali server (55) identified various proteins involved in polysaccharide recognition and modification (Table 2), including plant (56) and bacterial (57) expansins, endoglucanases (56, 58–60), and the plant defense protein barwin (61). Although these proteins share a moderately low sequence homology with CP, they display relatively high Dali *Z*-scores ranging from 9.5 to 5.2 (Table 2). Superposition of CP with these proteins highlights the significant structural match especially in the β-region (Fig. 2). Other less structurally related proteins, with *Z*-scores in the range of 3.6 to 3.2, include the lytic transglycosylases (LTs) (62).

CP Does Not Present Glycosidase Activity—Because CP is structurally related to endoglucanase and LT proteins, we tested whether it has glycoside hydrolase activity in *in vitro* assays. No cellulase, endo-1,3-β-glucanase, polygalacturonase, or chitinase activity could be detected even when CP was used at a high protein concentration relative to the positive controls (data not shown).

CP Binds Oligosaccharides—As shown in Table 2, the proteins most structurally related to CP display carbohydrate-binding properties; are involved in polysaccharide recognition; and in some cases, exhibit hydrolytic activity (56, 61). Therefore, we tested the oligosaccharide-binding capability of CP by measuring the chemical shift perturbation of the amide protons induced by oligosaccharides. Knowing that small oligosaccharides are able to bind barwin, we assayed the interaction of the tetramer of *N*-acetylglucosamine (GlcNAc-4) with CP. After

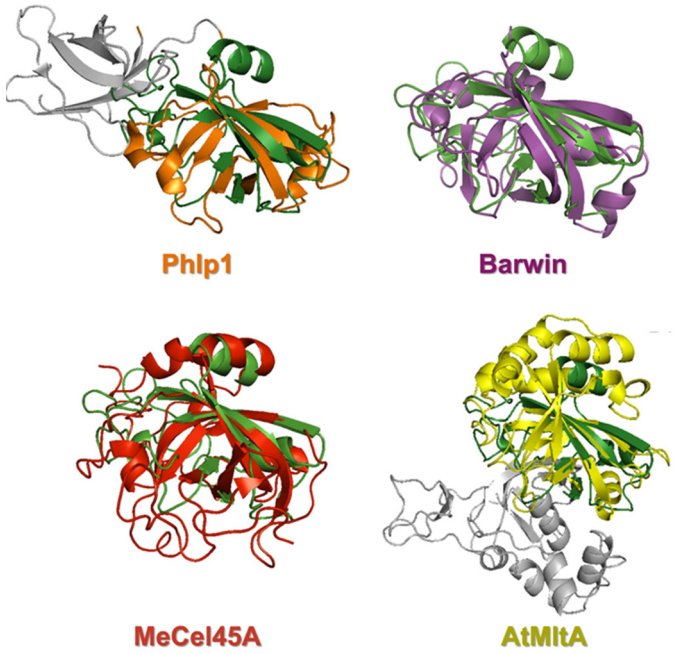

FIGURE 2. **Comparison of the CP structure with other proteins containing a double ψβ-barrel fold.** Shown is a diagram representation of the superposition of CP (*green*) with proteins that exhibit high *Z*-score values: the expansin Phlp1 (*orange*), the plant defense protein barwin (*magenta*), the endoglucanase MeCel45A (*red*), and the LT AtMltA (*yellow*). The protein domains of Phlp1 and AtMltA not structurally related to CP are depicted in *gray*.

the addition of increasing amounts of GlcNAc-4, we observed changes in the chemical shifts of a limited number of residues. The alterations were proportional to the quantity of added ligand, thus indicating a fast exchange between the apo- and holo-forms of the protein (see examples in supplemental Fig. S5*A*). Although the affinity of CP binding to GlcNAc-4 is weak, with an estimated K_d in the range of 10–100 μM, the weighted average chemical shift changes after the addition of >25 eq of GlcNAc-4 highlight three main regions involved in the oligosaccharide tetramer binding (Fig. 3*A*). These regions preferentially encompass portions of the loops: Ser-17–Ala-29 located in the β1-β2 loop, Gly-51–Ser-54 in the β2-β3 loop, and Asp-77–Gly-81 comprising the last part of the β4-strand and the β4-β5 loop. To further map the boundary of the CP oligosac-

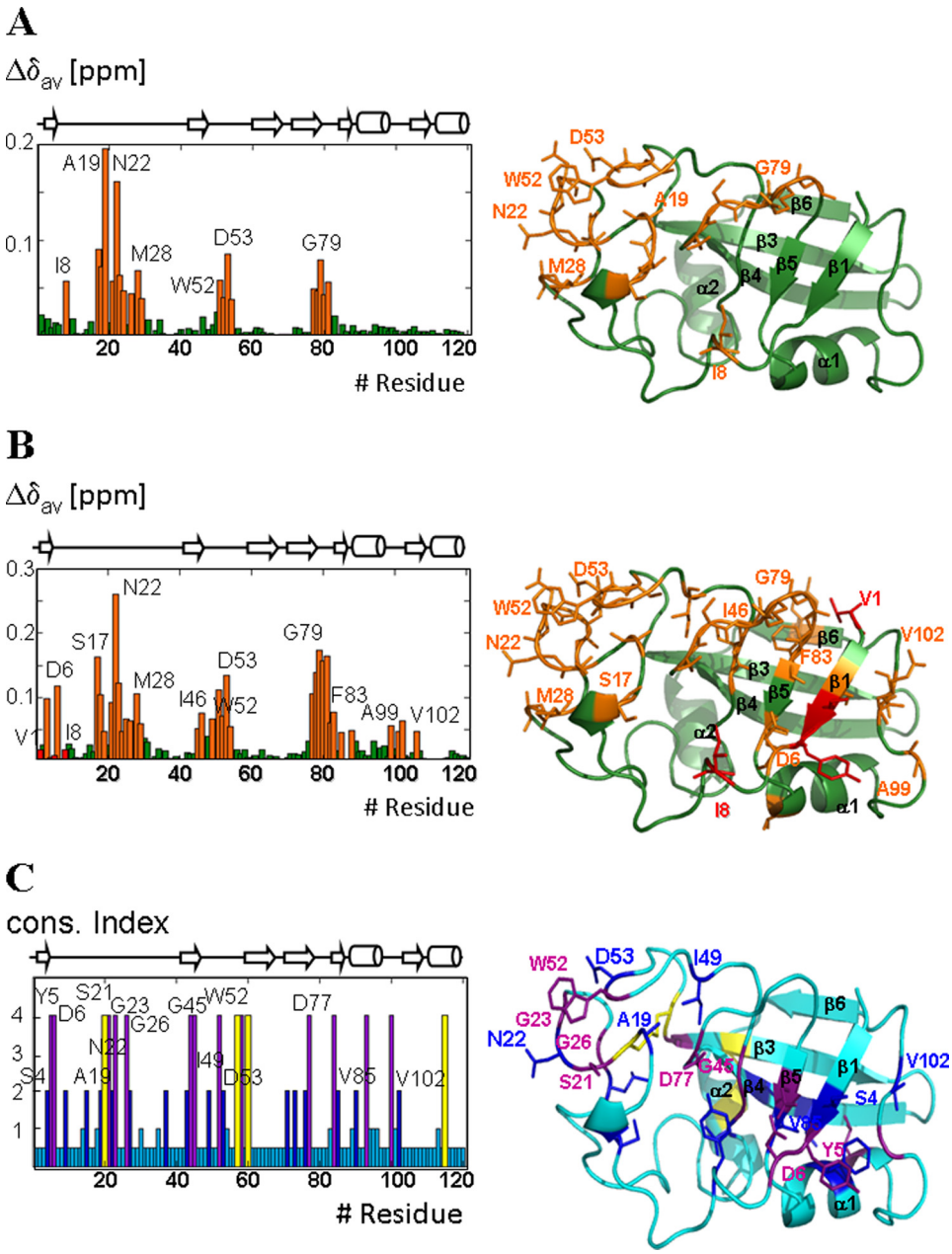

FIGURE 3. **Interaction of CP with polysaccharides.** *A* and *B*, average chemical shift perturbation upon interaction with GlcNAc-4 and with a mixture of longer GlcNAcs, respectively. The results are plotted *versus* residue number (*left panels*) and mapped on the structure of CP (*right panels*). Some of the residues are indicated in the structure and in the graph. *Green*, unperturbed residues; *orange*, residues whose amide chemical shift is perturbed; *red*, residues experiencing conformational effects due to the carbohydrate interaction. *C*, residue conservation among the CP family members from different fungal species. Invariable residues are indicated in *magenta*, except Cys residues (*yellow*), highly conserved residues (*blue*), and low and no conserved residues (*cyan*). This is shown along the CP sequence (*left panel*) and mapped on the CP structure (*right panel*). Residues that are invariable or highly conserved and that interact with oligosaccharides are indicated.

charide-binding region, ^{15}N-isotopically labeled CP was titrated with a mixture of oligosaccharides containing dimers, trimers, tetramers, pentamers, and hexamers of *N*-acetylglucosamine. Fig. 3*B* shows that, in addition to the oligosaccharide-binding regions identified with GlcNAc-4, the α1-β6 loop, Ala-99–Val-102, and the β1-strand, carrying the conserved SYD motif, were also affected. In particular, the signal belonging to Ser-4 in the β1-strand disappeared after the addition of the oligosaccharide mixture, whereas it remained unaffected in the presence of GlcNAc-4 (supplemental Fig. S5, *A* and *B*, *left panels*). We interpret these results as an indication of the selective interaction of these regions with the longer GlcNAcs.

Overall, the region involved in GlcNAc binding forms a flat and shallow groove on one face of the β-barrel, which is rich in polar (*e.g.* Asp-6, Asn-22, and Asp-53) and aromatic (*e.g.* Tyr-5 and Trp-52) residues suitable for sugar carbohydrate binding (Fig. 3*B*). Furthermore, all the residues involved in GlcNAc binding identified in this study are among the most conserved ones in the CP family (Fig. 3*C*).

DISCUSSION

Despite recent advances in the characterization of Avr plant pathogens and elicitor proteins, the limited number of structures available for these proteins has hampered the understand-

ing of the molecular mechanism associated with the interaction of these effectors with their targets or cognate pattern recognition receptors. Indeed, to date, only a few pairs of fungal effectors/targets and PAMPs/pattern recognition receptors have been identified and structurally characterized (2).

CP Structure—Thermal unfolding measurements show that the protein is particularly stable: the unfolding temperature is up to 76 °C (supplemental Fig. S4*B*), although it is irreversible. On the other hand, the protein secondary structure is preserved in a wide pH range, 3–9 (supplemental Fig. S4*A*). This last feature, perhaps the result of the presence of two disulfide bonds resistant to pH but not to temperature, is expected to favor the efficient secretion and/or translocation of CP into the host cell.

The solution structure of the CP protein that we report here is the first one for members of the CP family. The protein presents a double $\psi\beta$-barrel fold remarkably similar to those found in plant and bacterial expansins, LTs, endoglucanases, formate dehydrogenase H, dimethyl-sulfoxide reductase, and aspartic proteinases (63) and in the plant defense protein barwin (61).

Expansins are found primarily in plants, where they play a critical role in cell enlargement and other developmental processes requiring cell wall loosening (56). On the other hand, endoglucanases are ubiquitous enzymes; catalyze the hydrolysis of cellulose to smaller oligosaccharides in numerous plants, bacteria, and fungi; and are present in 13 of 80 families of glycoside hydrolase enzymes (64). LTs cleave the β-(1,4)-glycosidic bond between *N*-acetylmuramic acid and GlcNAc in peptidoglycans, with the concomitant formation of a 1,6-anhydro-ring at the *N*-acetylmuramic acid residue.

Except for barwin, which is also the most similar protein in primary sequence (28% identity), the other proteins are larger than CP (Fig. 2). Expansins present two domains: the N-terminal domain, which constitutes the double $\psi\beta$-barrel, and the C-terminal domain, which forms an Ig-like β-sandwich. Endoglucanases have insertions in the β3-β4 and β4-β5 loops and extensions at the N and C termini. In addition, they frequently present a second domain, called the cellulose-binding domain, attached with a linker to the C terminus (65). This is not the case, however, in MeCel45A, which presents a single catalytic domain. LTs are generally multidomain proteins, showing a long insertion (~140 residues) in the β1-β2 loop, which constitutes a second domain, and an N-terminal extension. It is worth noting a significant fold similarity between CP and EXLX1, a two-domain expansin from *Bacillus subtilis* that binds to plant cell walls and to carbohydrates but lacks lytic activity against a variety of cell wall polysaccharides and peptidoglycan (57), functional features shared, at least in part, also by CP.

Rationale for CP Lack of Glycoside Hydrolase Activity and for Oligosaccharide-binding Capability—The absence of glycoside hydrolase activity can be rationalized by structurally comparing CP with endoglucanases and LTs. The superposition of the endoglucanase MeCel45A catalytic site with the corresponding region of CP is depicted in Fig. 4*A*. The enzymatic activity of MeCel45A depends on 2 residues, Asp-24 and Asp-132, which are located in the β1- and β5-strands and act as a catalytic base and acid, respectively (58). The corresponding residues in CP are Ile-8 and Asn-84, which are conserved among the members of the CP family and exhibit different chemical properties.

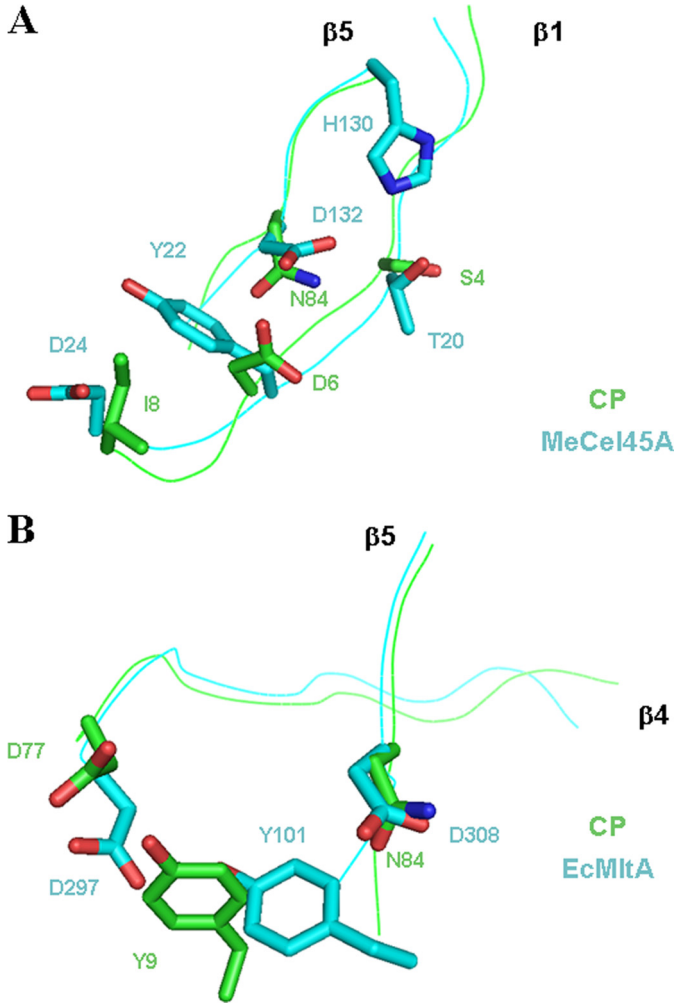

FIGURE 4. **Superposition of the active sites of the endoglucanase MeCel45A (*A*) and the LT EcMltA (*B*) with the corresponding region of CP.** Residues that are important for catalysis are displayed as *sticks*.

Other MeCel45A residues potentially involved in catalysis, as suggested by their being conserved among endoglucanases, are Thr-20 and Tyr-22 (Fig. 4*A*), replaced in CP with the conserved residues Ser-4 and Asp-6, respectively, again, at least in the case of Tyr-22/Asp-6, with diverse chemical properties (57).

Also the active site of LTs, presented here in the enzyme EcMltA, contains 2 aspartic acids as the main catalytic residues (Fig. 4*B*), which, as for the endoglucanase active sites, are located in β5 (Asp-308) and at the end of β4 (Asp-297). In the catalytic mechanism proposed for LTs (66), Asp-308 acts as an acid, and Asp-297, which is hydrogen-bonded to Tyr-101, acts as a base. In fact, Tyr-101 is supposed to shuttle a proton from Asp-297 to Asp-308.

This network of interactions observed in endoglucanases and LTs is only partially conserved in CP. In particular, in the CP structure, the substitution of an aspartic acid residue with Asn-84 by removing the carboxylic group might justify the disappearance of the hydrolase activity.

It is interesting to note, however, that Asp-77 and Tyr-9, both conserved in all CP family members, correspond to Asp-297 and Tyr-101 of EcMltA, respectively, whose side chains are expected to interact (Fig. 4*B*). The interaction between the

Asp-77 and Tyr-9 side chains is clearly visible by NMR because of the reduced solvent exchange of the hydroxyl of the Tyr-9 side chain, which appeared at 11.4 ppm. In a two-dimensional NOESY spectrum, clear NOE cross-peaks between this proton and the two β-protons of Asp-77 could be observed, suggesting the presence of a hydrogen bond between them (data not shown).

Fig. 3*C* highlights that the residues conserved among the members of the CP family are involved in polysaccharide binding. This fact strongly supports the hypothesis that the biological function of proteins belonging to the CP family may be related to polysaccharide binding, this feature being also consistent with the location of the CP protein in the fungal cell wall (28). Overall, site-specific mutations with respect to endoglucanase and LT proteins seem to explain why CP has lost glycoside hydrolase activity while maintaining the capability to bind oligosaccharides.

The data presented in this study contribute to deeper knowledge of the role of CP in both fungus fitness and host-microbe interaction. Because CP is located in the cell wall of hyphae, conidia, and ascospores of *C. platani* (28), its similarity to expansins could be a sign of the involvement of the protein in cell wall remodeling and enlargement. Moreover, the ability to bind GlcNAc might imply a role of CP in host interaction.

Recognizing that chitin oligomers can act as PAMPs, eliciting basal immune responses in plants (2, 67), it has been recently reported that the virulence effector Ecp6 from the fungal pathogen *Cladosporium fulvum* sequesters β-(1,4)-poly-*N*-acetyl-D-glucosamine, released from the hyphal cell wall during infection, to prevent elicitation of host immunity (68). All considered, our data point to the possibility that, during plant-microbe interaction, CP is overexpressed and interacts with the host, acting as a typical PAMP (24, 25). The elicited plant immune responses would include also overexpression of plant cell wall-degrading enzymes, which will produce chitin oligomers (69). At this stage, we may envisage that the expected enhancement of the plant primary defense response induced by those oligosaccharides can be blocked by the CP ability to bind GlcNAc, albeit a bit more weakly compared with Ecp6.

In conclusion, we trust that the solution of CP three-dimensional structure has opened the way to unravel the multifaceted functions not only of this protein but also of the other members of the CP family. 1) Because CP oligosaccharide binding involves residues conserved throughout the CP family, it is reasonable to expect that the other family members are able to bind oligosaccharides. 2) The observation that CP and Ecp6 bind GlcNAcs with different affinities might explain why members of the same family can act either as elicitors or effectors and provide critical hints to understand the incompatibility/compatibility of some plant-pathogen interactions. In other words, it appears that we are disclosing a more widespread strategy of host immune response suppression by fungal pathogens and elicitors (68).

REFERENCES

1. de Wit, P. J. G. M. (2007) *Cell. Mol. Life Sci.* **64,** 2726–2732
2. Zipfel, C. (2009) *Curr. Opin. Plant Biol.* **12,** 414–420
3. Juge, N. (2006) *Trends Plant Sci.* **11,** 359–367
4. Pitzschke, A., Schikora, A., and Hirt, H. (2009) *Curr. Opin. Plant Biol.* **12,** 421–426
5. Bent, A. F., and Mackey, D. (2007) *Annu. Rev. Phytopathol.* **45,** 399–436
6. Boller, T., and He, S. Y. (2009) *Science* **324,** 742–744
7. Dodds, P. N., and Rathjen, J. P. (2010) *Nat. Rev. Genet.* **11,** 539–548
8. Svozilová, Z., Kasparovský, T., Skládal, P., and Lochman, J. (2009) *Anal. Biochem.* **390,** 115–120
9. Kim, Y. T., Oh, J., Kim, K. H., Uhm, J. Y., and Lee, B. M. (2010) *Mol. Biol. Rep.* **37,** 717–727
10. De Wit, P. J., Mehrabi, R., Van den Burg, H. A., and Stergiopoulos, I. (2009) *Mol. Plant Pathol.* **10,** 735–747
11. Wilson, L. M., Idnurm, A., and Howlett, B. J. (2002) *Mol. Plant Pathol.* **3,** 487–493
12. Jeong, J. S., Mitchell, T. K., and Dean, R. A. (2007) *FEMS Microbiol. Lett.* **273,** 157–165
13. Zaparoli, G., Cabrera, O. G., Medrano, F. J., Tiburcio, R., Lacerda, G., and Pereira, G. G. (2009) *Mycol. Res.* **113,** 61–72
14. Djonovic, S., Vargas, W. A., Kolomiets, M. V., Horndeski, M., Wiest, A., and Kenerley, C. M. (2007) *Plant Physiol.* **145,** 875–889
15. Seidl, V., Marchetti, M., Schandl, R., Allmaier, G., and Kubicek, C. P. (2006) *FEBS J.* **273,** 4346–4359
16. Vargas, W. A., Djonović, S., Sukno, S. A., and Kenerley, C. M. (2008) *J. Biol. Chem.* **283,** 19804–19815
17. Yang, Y., Zhang, H., Li, G., Li, W., Wang, X., and Song, F. (2009) *Plant Biotechnol. J.* **7,** 763–777
18. Pan, S., and Cole, G. T. (1995) *Infect. Immun.* **63,** 3994–4002
19. Rementeria, A., López-Molina, N., Ludwig, A., Vivanco, A. B., Bikandi, J., Pontón, J., and Garaizar, J. (2005) *Rev. Iberoam. Micol.* **22,** 1–23
20. Hseu, Y. C., Wu, F. Y., Wu, J. J., Chen, J. Y., Chang, W. H., Lu, F. J., Lai, Y. C., and Yang, H. L. (2005) *Int. Immunopharmacol.* **5,** 1914–1925
21. Pazzagli, L., Cappugi, G., Manao, G., Camici, G., Santini, A., and Scala, A. (1999) *J. Biol. Chem.* **274,** 24959–24964
22. Pazzagli, L., Pantera, B., Carresi, L., Zoppi, C., Pertinhez, T. A., Spisni, A., Tegli, S., Scala, A., and Cappugi, G. (2006) *Cell Biochem. Biophys.* **44,** 512–521
23. Scala, A., Pazzagli, L., Comparini, C., Santini, A., Tegli, S., and Cappugi, G. (2004) *J. Plant Pathol.* **86,** 23–29
24. Fontana, F., Santini, A., Salvini, M., Pazzagli, L., Cappugi, G., Scala, A., Durante, M., and Bernardi, R. (2008) *J. Plant Pathol.* **90,** 295–306
25. Bernardi, R., Baccelli, I, Carresi, L, Comparini, C, Pazzagli, L., and Scala, A. (2011) *Forest Pathol.* **41,** in press
26. Pazzagli, L., Zoppi, C., Carresi, L., Tiribilli, B., Sbrana, F., Schiff, S., Pertinhez, T. A., Scala, A., and Cappugi, G. (2009) *Biochim. Biophys. Acta* **1790,** 1334–1344
27. Comparini, C., Carresi, L., Pagni, E., Sbrana, F., Sebastiani, F., Luchi, N., Santini, A., Capretti, P., Tiribilli, B., Pazzagli, L., Cappugi, G., and Scala, A. (2009) *Appl. Microbiol. Biotechnol.* **84,** 309–322
28. Boddi, S., Comparini, C., Calamassi, R., Pazzagli, L., Cappugi, G., and Scala, A. (2004) *FEMS Microbiol. Lett.* **233,** 341–346
29. Carresi, L., Pantera, B., Zoppi, C., Cappugi, G., Oliveira, A. L., Pertinhez, T. A., Spisni, A., Scala, A., and Pazzagli, L. (2006) *Protein Expr. Purif.* **49,** 159–167
30. Oliveira, A. L., Pazzagli, L., Pantera, B., Cappugi, G., Benedetti, C. E., Spisni, A., and Pertinhez, T. A. (2006) *J. Biomol. NMR* **36,** 50
31. Johnson, B. A. (2004) *Methods Mol. Biol.* **278,** 313–352
32. Delaglio, F., Grzesiek, S., Vuister, G. W., Zhu, G., Pfeifer, J., and Bax, A. (1995) *J. Biomol. NMR* **6,** 277–293
33. Zhang, O., Kay, L. E., Olivier, J. P., and Forman-Kay, J. D. (1994) *J. Biomol. NMR* **4,** 845–858
34. Muhandiram, D. R., Farrow, N. A., Xu, G., Smallcombe, S. H., and Kay, L. E. (1993) *J. Magn. Reson.* **102,** 314–321
35. Cicero, D. O., Contessa, G. M., Paci, M., and Bazzo, R. (2006) *J. Magn. Reson.* **180,** 222–228
36. Wang, Y. X., Marquardt, J. L., Wingfield, P., Stahl, S. J., Lee-Huang, S., Torchia, D., and Bax, A. (1998) *J. Am. Chem. Soc.* **120,** 7385–7386
37. Ottiger, M., Delaglio, F., and Bax, A. (1998) *J. Magn. Reson.* **131,** 373–378
38. McFeeters, R. L., Fowler, C. A., Gaponenko, V. V., and Byrd, R. A. (2005) *J. Biomol. NMR* **31,** 35–47
39. Eliseo, T., Ragona, L., Catalano, M., Assfalg, M., Paci, M., Zetta, L., Moli-

nari, H., and Cicero, D. O. (2007) *Biochemistry* **46,** 12557–12567
40. Waldeck, A. R., Kuchel, P. W., Lennon, A. J., and Chapman, B. E. (1997) *Prog. Nucl. Magn. Reson. Spectrosc.* **30,** 39–68
41. Jones, J. A., Wilkins, D. K., Smith, L. J., and Dobson, C. M. (1997) *J. Biomol. NMR* **10,** 199–203
42. Schwieters, C. D., Kuszewski, J. J., Tjandra, N., and Clore, G. M. (2003) *J. Magn. Reson.* **160,** 65–73
43. Cornilescu, G., Delaglio, F., and Bax, A. (1999) *J. Biomol. NMR* **13,** 289–302
44. Vriend, G. (1990) *J. Mol. Graph.* **8,** 52–56
45. Losonczi, J. A., Andrec, M., Fischer, M. W., and Prestegard, J. H. (1999) *J. Magn. Reson.* **138,** 334–342
46. Zweckstetter, M., and Bax, A. (2000) *J. Am. Chem. Soc.* **122,** 3791–3792
47. Clore, G. M., and Garrett, D. S. (1999) *J. Am. Chem. Soc.* **121,** 9008–9012
48. Laskowski, R. A., Rullmann, J. A., MacArthur, M. W., Kaptein, R., and Thornton, J. M. (1996) *J. Biomol. NMR* **8,** 477–486
49. Gueguen, Y., Voorhorst, W. G., van der Oost, J., and de Vos, W. M. (1997) *J. Biol. Chem.* **272,** 31258–31264
50. Cervone, F., Hahn, M. G., De Lorenzo, G., Darvill, A., and Albersheim, P. (1989) *Plant Physiol.* **90,** 542–548
51. Wood, T. M., and Bhat, K. M. (1998) *Methods Enzymol.* **160,** 87–112
52. York, W. S., Darvill, A. G., McNeil, M., Stevenson, T. I., and Albersheim, P. (1985) *Methods Enzymol.* **118,** 3–40
53. Grzesiek, S., Bax, A., Clore, G. M., Gronenborn, A. M., Hu, J. S., Kaufman, J., Palmer, I., Stahl, S. J., and Wingfield, P. T. (1996) *Nat. Struct. Biol.* **3,** 340–345
54. Wilkins, D. K., Grimshaw, S. B., Receveur, V., Dobson, C. M., Jones, J. A., and Smith, L. J. (1999) *Biochemistry* **38,** 16424–16431
55. Holm, L., and Park, J. (2000) *Bioinformatics* **16,** 566–567
56. Sampedro, J., and Cosgrove, D. J. (2005) *Genome Biol.* **6,** 242
57. Kerff, F., Amoroso, A., Herman, R., Sauvage, E., Petrella, S., Filée, P., Charlier, P., Joris, B., Tabuchi, A., Nikolaidis, N., and Cosgrove, D. J. (2008) *Proc. Natl. Acad. Sci. U.S.A.* **105,** 16876–16881
58. Davies, G. J., Tolley, S. P., Henrissat, B., Hjort, C., and Schülein, M. (1995) *Biochemistry* **34,** 16210–16220
59. Hirvonen, M., and Papageorgiou, A. C. (2003) *J. Mol. Biol.* **329,** 403–410
60. Eriksson, J., Malmsten, M., Tiberg, F., Callisen, T. H., Damhus, T., and Johansen, K. S. (2005) *J. Colloid Interface Sci.* **285,** 94–99
61. Ludvigsen, S., and Poulsen, F. M. (1992) *Biochemistry* **31,** 8783–8789
62. Blackburn, N. T., and Clarke, A. J. (2001) *J. Mol. Evol.* **52,** 78–84
63. Castillo, R. M., Mizuguchi, K., Dhanaraj, V., Albert, A., Blundell, T. L., and Murzin, A. G. (1999) *Structure* **7,** 227–236
64. Henrissat, B., and Bairoch, A. (1996) *Biochem. J.* **316,** 695–696
65. Gilkes, N. R., Henrissat, B., Kilburn, D. G., Miller, R. C., Jr., and Warren, R. A. (1991) *Microbiol. Rev.* **55,** 303–315
66. Powell, A. J., Liu, Z. J., Nicholas, R. A., and Davies, C. (2006) *J. Mol. Biol.* **359,** 122–136
67. Lee, C. G., Da Silva, C. A., Lee, J. Y., Hartl, D., and Elias, J. A. (2008) *Curr. Opin. Immunol.* **20,** 684–689
68. de Jonge, R., van Esse, H. P., Kombrink, A., Shinya, T., Desaki, Y., Bours, R., van der Krol, S., Shibuya, N., Joosten, M. H., and Thomma, B. P. (2010) *Science* **329,** 953–955
69. Hématy, K., Cherk, C., and Somerville, S. (2009) *Curr. Opin. Plant Biol.* **12,** 406–413

SUPPLEMENTARY MATERIAL

Sequence alignment of CP with other members of the CP family

BLAST search (http://blast.ncbi.nlm.nih.gov/Blast.cgi) revealed that CP shares significant similarities to fungi proteins only. Interestingly, all the identified proteins are currently classified as members of the CP family. The Figure reports the sequence alignments for the selected proteins.

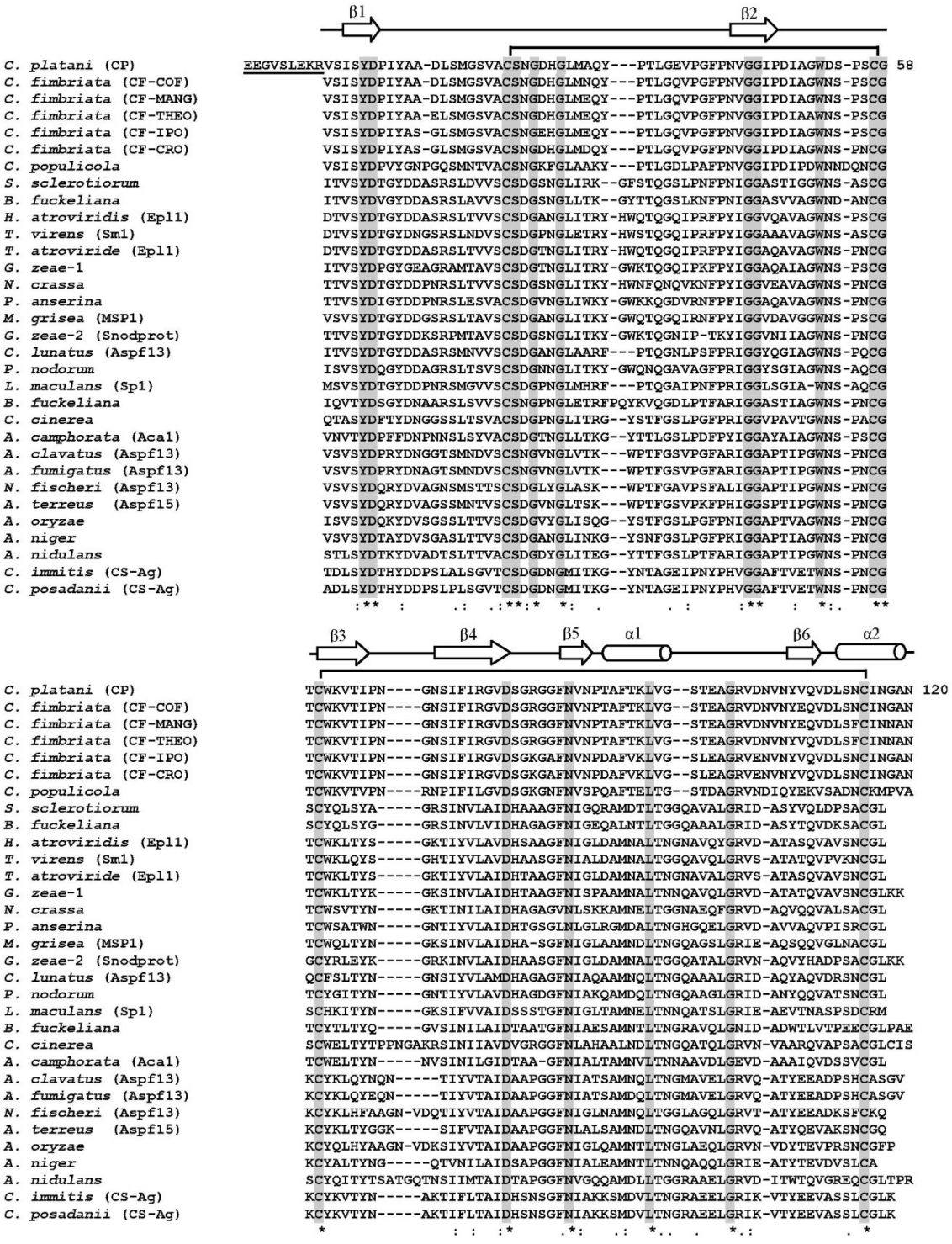

Fig. S1. Sequence alignment of CPF members from different fungal species, generated by ClustalW (http://www.ebi.ac.uk/clustalW). Invariable residues are shaded and marked with asterisks. The cysteines involved in disulfide bridges are linked. The N-terminal secretion signal sequences of all proteins were removed to optimize the alignment of CP with other CP variants from *Ceratocystis fimbriata*, CF-COF (ABM63509), CF-MANG (ABM63508), CF-THEO (ABM63511), CF-IPO (ABM63510), CF-CRO (ABM63507), and with related proteins from *Ceratocystis populicola* (ABM63506), *Sclerotinia sclerotiorum* (XP_001588549), *Botryotinia fuckeliana* (XP_001559499), *Hypocrea atroviridis* (ABE73692), *Thrichoderma virens* (AAZ80388), *Trichoderma antroviride* (CAL80754), *Gibberella zeae-1* (XP_391381), *Neurospora crassa* (XP_958708), *Podospora anserine* (XP_001911154), *Magnaporthe grisea* (XP_359969), *Gibberella zeae-2* (AAV83791), *Cochliobolus lunatus* (AAQ87930), *Phaeosphaeria nodorum* (XP_001803929), *Leptosphaeria maculans* (AAM33130), *Botryotinia fuckeliana* (XP_001552257), *Coprinopsis cinerea* (XP_001839821), *Antrodia camphorata* (AAT11911), *Aspergillus clavatus* (XP_001275544), *Aspergillus fumigatus* (XP_755595), *Neosartorya fischeri* (XP_001260717), *Aspergillus terreus* (XP_001208607), *Aspergillus oryzae* (XP_001821825), *Aspergillus niger* (XP_001399295), *Aspergillus nidulans* (XP_663877), *Coccidioides immitis* (XP_001247410) and *Coccidioides osadanii* (Q00398). The secondary structure elements and loops determined in the 3D structure of CP are shown. The additional 9 residues N-terminal stretch are underlined.

Backbone ^{15}N-relaxation measurements and analysis

Longitudinal (T_1) and transverse (T_2) relaxation times as well as {^{1}H-^{15}N} NOEs of backbone amides were determined using experiments acquired at 500 MHz with a ^{15}N-labelled CP sample (S1). The {^{1}H-^{15}N} NOE values were taken as the ratio of the peak volumes of the experiments recorded with and without ^{1}H saturation using a saturation time of 3 s and a recycle delay of 2 s. Relaxation delays of 10, 210, 510, 710, 1010, 1310, 1510 and 1710 ms were employed for T_1 measurements, and 10, 50, 90 130, 170, 210, 250 and 290 ms for T_2 measurements. T_1 and T_2 values were extracted by fitting the peak intensities to a single exponential decay function in NMRView. (S2) The uncertainties of peak intensities were evaluated using the standard deviation of the spectral noise measured in a region free of cross peaks.

Please note that the experimental points before residue 1 refer to the extra 9 residues as explained in the "***Expression and purification of cerato-platanin"*** paragraph in the Experimental Section. The data clearly indicate that the region has a disordered conformation.

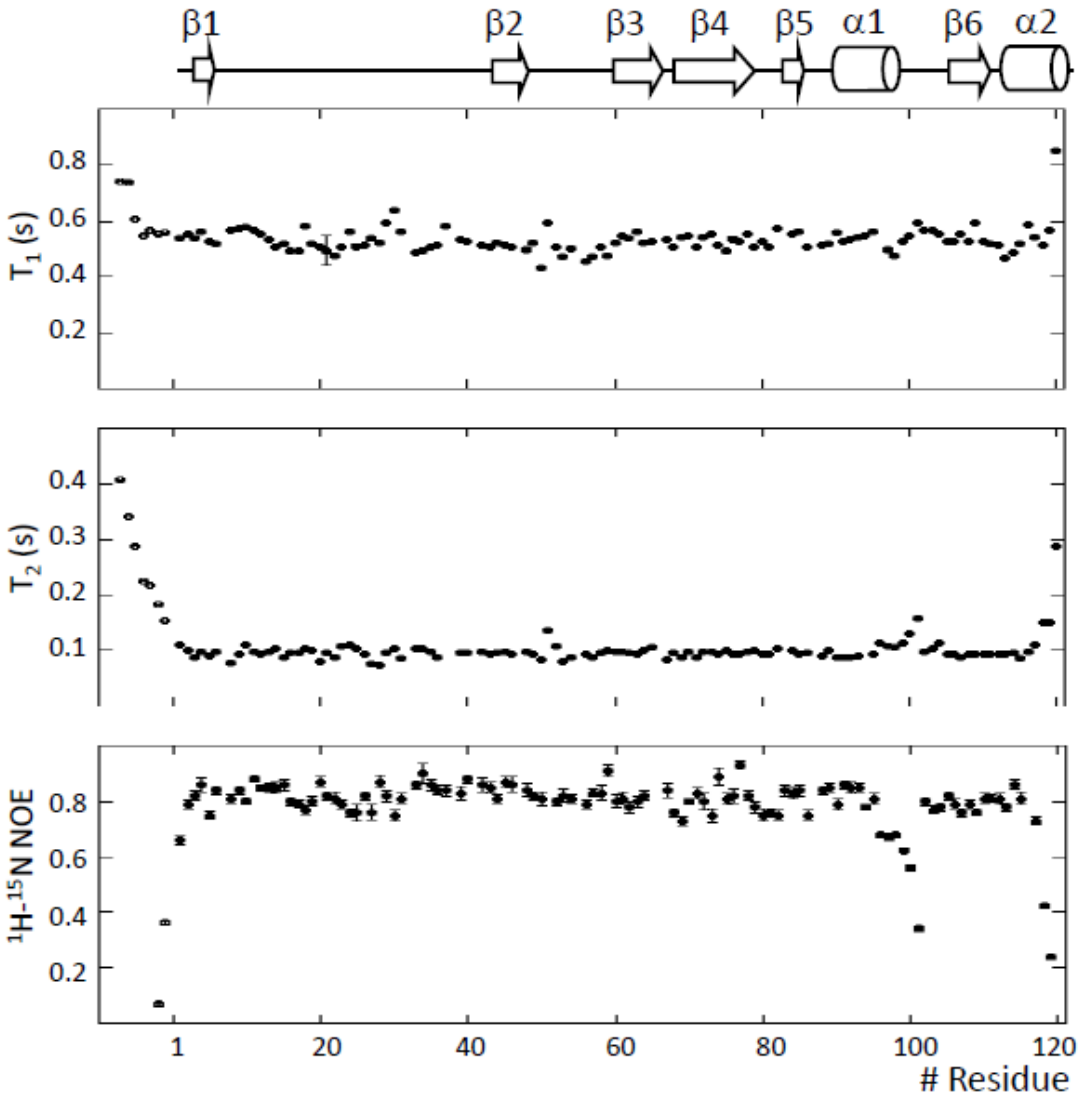

Fig. S2. Backbone relaxation analysis of CP. Plots of T_1 (upper panel), T_2 (middle panel) and ^{1}H-^{15}N NOE (lower panel) vs. the protein residues.

CD spectra of the recombinant CP protein

Circular dichroism (CD) measurements were performed on a JASCO J-810 spectropolarimeter equipped with a Peltier temperature control system. The far-UV spectra were recorded at 20 °C in a 1 mm optical path length cell in the wavelength range 190-260 nm with a bandwidth of 1 nm. The near-UV spectra were recorded in the wavelength range 250-350 nm, at 20 °C using a 10 mm optical path length cell. For each measurement, the average of four spectra was taken to improve the signal to noise ratio and each spectrum was corrected by subtracting its blank. Typically, 15 µM of purified CP solution in 10 mM sodium phosphate buffer, pH 5.8, was used for the CD measurements. The stability of the protein at different pH conditions was also assayed. The pH values were adjusted by adding few drops of concentrated HCl or NaOH.

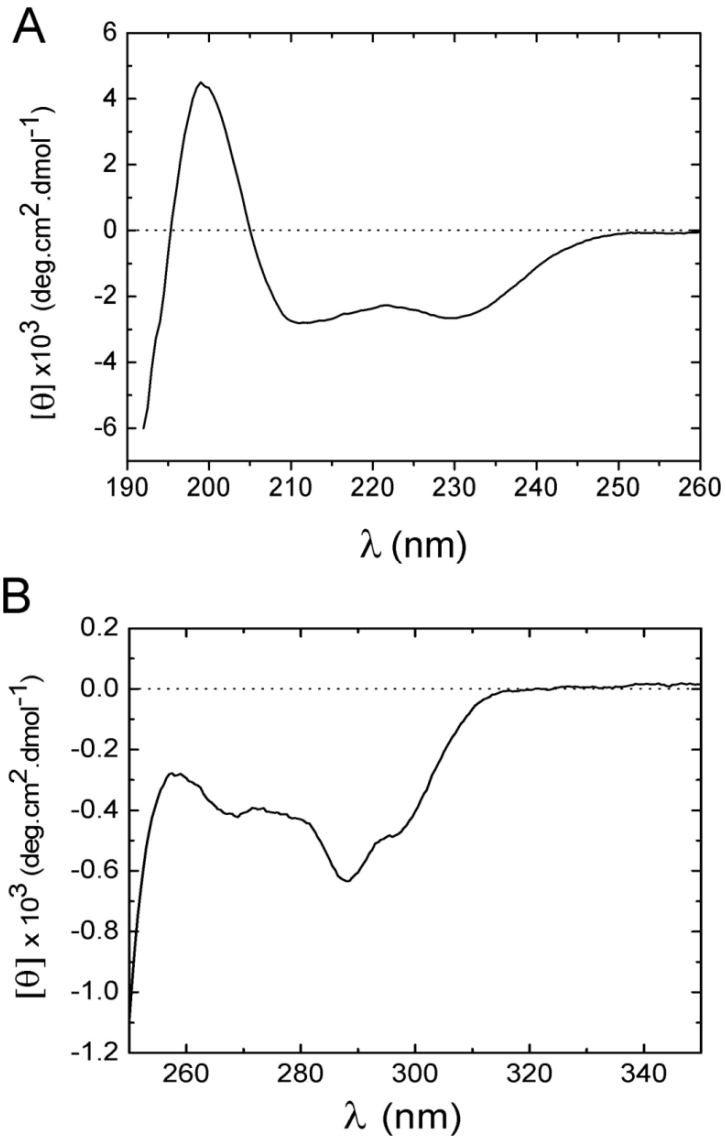

Fig. S3. Far-UV (A) and near-UV (B) circular dichroism spectra of recombinant CP at 15 µM in 10 mM sodium phosphate buffer, pH 5.8 at 20°C.

CP structure stability

Protein stability has been analyzed as a function of pH and temperature. CD measurements were performed as previously described.

The pH values were adjusted by adding few drops of concentrated HCl or NaOH.

Thermal denaturation curves were recorded over a 15-90 °C temperature interval, following the intensity of the CD signal at 212 nm. The temperature range was monitored at a temperature scan rate of 1 °C.min^{-1}, using 1 mm optical path length and 1 nm bandwidth. Reversibility of denaturation was tested by an inverse temperature scanning, immediately after the end point temperature (90 °C) was reached and using the same temperature scan rate.

The CD spectra of CP at different pHs show that the protein maintains its conformation in a wide pH range (Fig. S4A). The protein is resistant to heat denaturation up to 60 °C; CP unfolding is complete only at about 82 °C and the melting temperature turns out to be 76 °C (Fig. S4B). The unfolding is not reversible as only 61% of the secondary structure can be recovered (Fig. S4B).

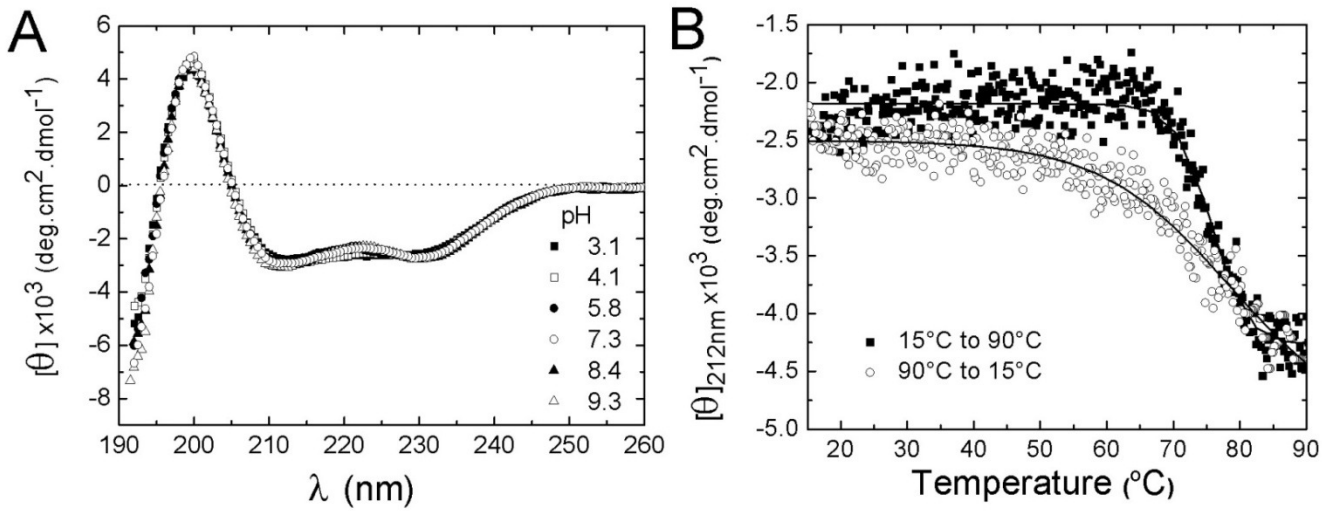

Fig. S4. Characterization of CP stability. (A) Far-UV CD spectra of CP at different pH conditions. (B) Temperature-induced denaturation (filled squares) and renaturation (opened circles) curves of CP. The curves were recorded over a 15–90°C temperature interval, following the CD signal at 212 nm.

Chemical Shift perturbation experiments

The experiments have been carried out as reported in Experimental Procedures.

In Fig. S5 are reported, as an example, the chemical shift perturbations of some amide crosspeaks upon interaction with NAG-4 (A) and a NAGs mix formed by tetramer, pentamer and examer (B). Some residues do not interact with any of the oligomers added (like T64, N67, and S70) and others are perturbed only by the addition of the NAGs mix (S4 and V102) thus indicating they interact only with NAGs longer than NAG-4.

In *section A* are reported, as an example, some of the residues affected (S17) or not affected (S4, T64, S70, V102) by the presence of NAG-4.

In *section B* it is possible to observe what happens to the same residues when the NAGs mix is added. In the case of S4 the interaction is so effective to produce a broadening of the line-width that precludes its detection. V102 turns out to be significantly shifted, S17 exhibits a diverse chemical shift perturbation with respect to the case in which only NAG-4 is present. Note that the shift for this peak move straight to lower chemical shift.

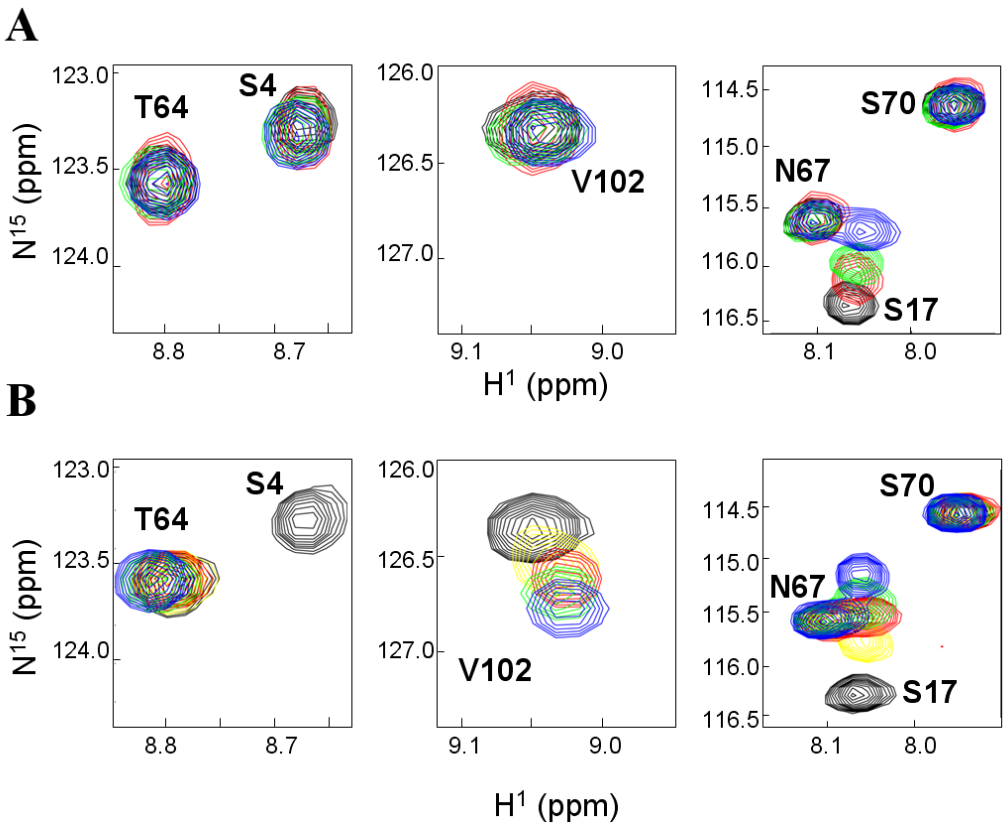

Fig. S5. Chemical shift perturbation upon interaction of NAG-4 (A) and NAGs-mix (B). Selected regions of ^{1}H-^{15}N HSQC spectra of 30 μM CP obtained in the absence (black) or after the addition of 2 mM (red), 4 mM (green) and 8 mM (blue) of NAG-4 (A). The same ^{1}H-^{15}N HSQC regions for 100 μM CP in the absence (black) or after the addition of 0.5 mM (yellow), 1 mM (red), 1.5 mM (green), 2 mM (blue) of NAGs-mix in a solution containing a mixture of mono, di, tri, tetra, penta and hexa chitooligosaccharides (see Experimental Procedures) (B).

References

S1 Ishima, R., and Torchia, D. A. (2000) *Nat. Struct. Biol.* **7**, 740-743

S2 Johnson, B. A. (2004) *Methods Mol. Biol.* **278**, 313-352