Peptides 31 (2010) 541–547

Contents lists available at ScienceDirect

Peptides

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

Characterization of the novel antifungal protein PgAFP and the encoding gene of *Penicillium chrysogenum*

Andrea Rodríguez-Martín [a], Raquel Acosta [a], Susan Liddell [b], Félix Núñez [a], Mª José Benito [c], Miguel A. Asensio [a,*]

[a] Higiene y Seguridad Alimentaria, Facultad de Veterinaria, Universidad de Extremadura, Avda. de la Universidad, s/n. 10071 - Cáceres, Spain
[b] Division of Animal Sciences, University of Nottingham, Sutton Bonington Campus, Loughborough, Leicestershire LE12 5RD, United Kingdom
[c] Nutrición y Bromatología, Escuela de Ingenierías Agrarias, Universidad de Extremadura, Carretera de Cáceres s/n. 06071 - Badajoz, Spain

ARTICLE INFO

Article history:
Received 29 September 2009
Received in revised form 4 November 2009
Accepted 4 November 2009
Available online 13 November 2009

Keywords:
Antifungal protein
Penicillium chrysogenum
Glycosylation
Toxigenic molds
RACE-PCR
Preproprotein

ABSTRACT

The strain RP42C from *Penicillium chrysogenum* produces a small protein PgAFP that inhibits the growth of some toxigenic molds. The molecular mass of the protein determined by electrospray ionization mass spectrometry (ESI-MS) was 6 494 Da. PgAFP showed a cationic character with an estimated pI value of 9.22. Upon chemical and enzymatic treatments of PgAFP, no evidence for N- or O-glycosylations was obtained. Five partial sequences of PgAFP were obtained by Edman degradation and by ESI-MS/MS after trypsin and chymotrypsin digestions. Using degenerate primers from these peptide sequences, a segment of 70 bp was amplified by PCR from *pgafp* gene. 5′- and 3′-ends of *pgafp* were obtained by RACE-PCR with gene-specific primers designed from the 70 bp segment. The complete *pgafp* sequence of 404 bp was obtained using primers designed from 5′- and 3′-ends. Comparison of genomic and cDNA sequences revealed a 279 bp coding region interrupted by two introns of 63 and 62 bp. The precursor of the antifungal protein consists of 92 amino acids and appears to be processed to the mature 58 amino acids PgAFP. The deduced amino acid sequence of the mature protein shares 79% identity to the antifungal protein Anafp from *Aspergillus niger*. PgAFP is a new protein that belongs to the group of small, cysteine-rich, and basic proteins with antifungal activity produced by ascomycetes. Given that *P. chrysogenum* is regarded as safe mold commonly found in foods, PgAFP may be useful to prevent growth of toxigenic molds in food and agricultural products.

© 2009 Elsevier Inc. All rights reserved.

1. Introduction

Uncontrolled fungal growth in agricultural and food commodities causes various problems, from spoilage to health hazards due to mycotoxins. Unwanted molds can be controlled by effective antifungal methods, including both chemical and physical treatments. However, most of these methods are not suitable for some applications, particularly on mold-ripened foods, where molds play a key role in developing the main characteristics. The use of antifungal compounds selective against the unwanted molds can be a means of control in such foods. Recently, a group of cysteine-rich, highly basic, and low molecular weight (5.8–6.6 kDa) antifungal proteins from ascomycetes was described [20]. Only four proteins from this group have been purified: AFP from *Aspergillus giganteus* [14,23], PAF from *Penicillium chrysogenum* [19], Anafp from *Aspergillus niger* [16], and AcAFP from *Aspergillus clavatus* [27]. In addition, *Penicillium nalgiovense* showed antifungal activity attributed to NAF, a non-demonstrated protein, encoded by a gene homologous to gene encoding PAF [10]. The available genetic characterization of the genes encoding AFP [32], PAF [19], AcAFP [28], and NAF [10] revealed open reading frames interrupted by two introns with conserved splice sites. They encode products of 92 or 94 amino acids, which are long precursors with a peptide signal and a prosequence that are removed to obtain mature proteins.

The strain RP42C, characterized as *P. chrysogenum* by classical methods, displayed a strong antifungal activity against selected toxigenic molds from meat products [1]. This activity was related to an extracellular protein with a molecular mass of 9.1 kDa estimated by SDS-PAGE that may belong to the group of small antifungal proteins produced by Ascomycetes [1]. The difference in molecular weight with other antifungal proteins from molds can be related to post-translational modifications, such as the glycosylations present in antifungal proteins from plants [8,33]. However, no post-translational modification has been reported for this group of antifungal proteins from molds.

Since *P. chrysogenum* RP42C was isolated from dry-cured ham, the potent 9.1 kDa protein could be of interest to control unwanted

* Corresponding author. Tel.: +34 927 257 125; fax: +34 927 257 110.
E-mail address: masensio@unex.es (Miguel A. Asensio).

0196-9781/$ – see front matter © 2009 Elsevier Inc. All rights reserved.
doi:10.1016/j.peptides.2009.11.002

fungi in foods and agriculture. Therefore, further characterization of this antifungal protein from *P. chrysogenum* (PgAFP) is necessary to develop potential applications.

The aim of the present work was to characterize PgAFP, including determining the amino acid sequence and the presence of glycan chains, as well as to obtain the sequence of the gene encoding this protein.

2. Materials and methods

2.1. Microbial strain

P. chrysogenum RP42C isolated from dry-cured ham [1] was used in this study.

2.2. Genetic identification of strain RP42C

The identification of the producer mold was carried out by amplifying the fragments 18S–28S rRNA intergenic spacer (ITS) by polymerase chain reaction. RP42C was grown in malt extract broth (MEB) [25], pH 4.5, at 25 °C for 5 days. DNA was obtained by phenolic extraction according to the method of Benito et al. [3] and stored at −70 °C. The quantity and quality of the purified DNA was determined using a biophotometer (Eppendorf AG, Hamburg, Germany). DNA fragments for sequence analysis were obtained by PCR using primers ITS1and ITS4 [30]. For each 50 μl reaction, a mixture was prepared containing 50 ng of genomic DNA, 0.1 vol of 10× PCR buffer (10 mM Tris–HCl, pH 8.8, 50 mM KCl, 0.1% Triton X-100), 4 mM $MgCl_2$, 1 mM each of dNTPs (Roche, Mannheim, Germany), 0.4 μM each of primer, and 2 U DNAzyme I DNA Polymerase (Finnzymes, Espoo, Finland). Reactions were run on a programmable thermal cycler Mastercycler epgradient (Eppendorf AG) with an initial denaturation of 3 min at 96 °C, followed by 30 cycles at 94 °C for 30 s, 60 °C for 30 s, and 72 °C for 40 s, with a final extension at 72 °C for 10 min. Aliquots of the PCR were electrophoresed on 1% agarose gels and the fragments were gel-purified. DNA fragments were sequenced at the Institute of Biomedicine (CSIC, Valencia, Spain).

2.3. Protein characterization

PgAFP was obtained according to the method of Acosta et al. [1] by fast protein liquid chromatography with an ion exchange column and further purified with a gel filtration column. Finally, the fraction containing PgAFP was desalted and concentrated with YM-3 Microcon centrifugal filter units (Millipore Iberica S.A.U., Madrid, Spain) before storage at −20 °C. SDS-PAGE was performed as described by Laemmli [15]. PgAFP was separated on 4% polyacrylamide stacking and 12% polyacrylamide resolving gels, and stained with Imperial Protein Stain (Pierce Biotechnology, Rockford, IL). Two molecular weight marker sets were used: Precision Plus Protein Standards with ten proteins between 10 and 250 kDa (Bio-Rad) and Sigma Marker Low Range with proteins between 6.5 and 66 kDa (Sigma–Aldrich Química S.A., Madrid, Spain).

2.3.1. Determination of isoelectric point (pI)

Capillary isoelectric focusing electrophoresis was performed in an automated P/ACE 2200 (Beckman Instruments, Palo Alto, CA), using a 50 μm diameter by 27 cm total length (20 cm until window detector) and neutral uncoated fused silica capillary (Supelco, Bellefonte, PA). The capillary column was conditioned by flushing at high pressure (20 psi) with 10 mM phosphoric acid for 1 min and with deionized water for 3 min. The catholyte was 20 mM sodium hydroxide aqueous solution and the anolyte was 91 mM phosphoric acid in cIEF gel (Beckman Coulter, London, UK). Sample injection was at low pressure (0.5 psi) for 1 min. Electrophoretic separation was performed at constant potential of 500 V/cm (13.5 kV). The absorbance was recorded at 280 nm. The pI was estimated by comparing the electrophoretic mobility of the protein and the cIEF 3–10 pI marker proteins from Beckman Coulter.

2.3.2. Molecular mass determination

The molecular mass of PgAFP was determined by electrospray mass spectrometry on a Q-TOF2 mass spectrometer (Waters, Manchester, UK). The protein was concentrated and dissolved in ammonium acetate. PgAFP (0.1 μg/μl) with 1% formic acid was desalted using C_{18} reversed-phase pipette Zip-Tips (Millipore) according to the manufacturer's instructions. Samples were eluted into 50% methanol, 0.1% formic acid, and were loaded into nanospray needles (Waters) for analysis. The mass spectrometer was operated with a capillary voltage of 900–1200 V and the sampling cone at 45–50 V. Data were acquired between 400 and 3000 *m*/*z* with a cycle time of 1 s. Spectra were deconvoluted using the MaxEnt1 maximum entropy software (Waters, UK) with an output range of 4000 and 11 000 Da, at a resolution of 1 Da and a peak width of 0.75 Da.

2.3.3. Analysis for glycosylations

Both chemical and enzymatic treatments of deglycosylation were performed after reduction and alkylation of the protein in acrylamide gel pieces as well as in-solution. For in-gel deglycosylations, the PgAFP band was excised from the SDS-PAGE gel (25 μg of protein per well) and washed three times in 50 mM ammonium bicarbonate and 50% acetonitrile for 15 min at 40 °C with intermittent vortexing. Gel pieces were incubated at 56 °C in 100 μl of reducing solution (10 mM dithiothreitol in 50 mM ammonium bicarbonate) that was removed after 45 min. One hundred microliter of alkylation solution (55 mM iodoacetamide in 50 mM ammonium bicarbonate) was added and incubation continued at room temperature in the dark for 45 min. The gel slice was then washed in 100 μl of 50 mM ammonium bicarbonate for 5 min at 40 °C and the supernatant was discarded. Gel pieces were washed twice in 100 μl of 50 mM ammonium bicarbonate in 50% acetonitrile for 1 min, and centrifuged. Sample was dehydrated in 100 μl of acetonitrile and incubated for 1–2 min. Finally, the supernatant was discarded and the gel pieces were dried at 40 °C for a few min.

For PgAFP in-solution deglycosylations, different quantities of the protein (12.5, 25, and 62.5 μg) were added to tubes with 50 mM ammonium bicarbonate up to a volume of 5 μl. Five microliters of 5 mM dithiothreitol in 50 mM ammonium bicarbonate were added and incubated at 50 °C for 30 min. After adding 10 μl of 20 mM iodoacetamide in 50 mM ammonium bicarbonate, sample was incubated at room temperature for 30 min in the dark before the digestion with each deglycosidase.

For chemical deglycosylation, PgAFP in solution was treated with two different methods: acid hydrolysis with HCl [17] and β-elimination with NaOH [26]. For acid hydrolysis, 1 μg of PgAFP was incubated in 20 mM HCl at 80 °C for 3 h. β-Elimination was performed incubating 1 μg of PgAFP with 0.2 M NaOH at 45 °C for 16 h. Finally, both samples were analyzed by ESI-MS.

For enzymatic deglycosylations of PgAFP three different treatments were used. PgAFP protein (200 μg) was treated with 4 μl of endoglycosidase F1 (Northstar BioProducts, Liverpool, UK) in 250 mM NaH_2PO_4, pH 4.5 at 37 °C for 1 h, and analyzed by ESI-MS. PNGase F (New England Biolabs, Ipswich, UK) was used following the manufacturer's instructions. Two additional assays were performed without detergent with PgAFP in-gel and in-solution, one with a prior reduction and alkylation as described above in this subsection and another one with previous heating of PgAFP at 100 °C for 5 min. All samples were analyzed by SDS-PAGE

and ESI-MS. The E-DEGLY deglycosylation kit (Sigma, Madrid, Spain) containing PNGase F, α-2(3,6,8,9) neuraminidase, O-glycosidase, (1–4)-galactosidase, and β-N-acetylglucosaminidase was also used. Reactions were carried out under native and denaturing conditions according to the manufacturer's instructions and analyzed by SDS-PAGE.

2.3.4. Amino acid sequencing

The N-terminal sequence of PgAFP was carried out by automated Edman degradation in the Proteomics Platform (University of Barcelona, Spain).

The *de novo* sequencing of PgAFP was performed by electrospray ionization mass spectrometry (ESI-MS/MS) after trypsin and chymotrypsin digestion. This process was carried out with and without previous reduction and alkylation of PgAFP as described in Section 2.3.3.

For trypsin digestion, gel pieces were incubated with 10 ng/μl Trypsin Gold (Promega, Hampshire, UK) in 50 mM ammonium bicarbonate at 40 °C or 60 °C for 1.5, 3, and 16 h. PgAFP in-solution was incubated at 37 °C for 16 h in 50 mM ammonium bicarbonate, 5% acetonitrile, and 10 ng/μl of Trypsin Gold.

For chymotrypsin digestion, gel pieces were digested at different concentrations (1, 5, 10, and 15 ng/μl) of chymotrypsin (Roche, Welwyn Garden City, UK) in buffer (100 mM Tris–HCl, 10 mM $CaCl_2$, pH 7.8) at 25 and 30 °C for 16 h. PgAFP in-solution was digested by chymotrypsin at 0.1, 1, 5, or 10 ng/μl with 5% acetonitrile and the same buffers detailed above. Temperatures and incubation times were identical to the in-gel digested samples.

For ESI-MS/MS analysis, peptide samples were made to 0.1% formic acid, then desalted using Zip-Tip C_{18} reversed-phase pipette tips (Millipore) according to manufacturer's instructions. Peptides were eluted with 50% acetonitrile, 0.1% formic acid. Samples were loaded into individual borosilicate nanospray needles (Waters, UK) using 1–10 μl GELoader pipette tips (Eppendorf) and analyzed in a Q-TOF2 mass spectrometer (Waters, UK), fitted with a nanoflow ESI source. The mass spectrometer was operated with a capillary voltage of 900–1200 V in positive ion mode, using argon as the collision gas. Survey scans of peptide mass spectra were performed with the sampling cone set at 45–50 V and data were acquired from 400–1600 m/z with a cycle time of 2.4 s. Individual peptide ions were selected manually for tandem MS analysis. Tandem MS fragmentation spectra were collected typically from 50 to 1600 m/z. Tandem MS spectra were deconvoluted into singly charged, mono-isotopic masses using the MaxENT3 maximum entropy software, and peptide sequences were determined by manual interpretation using the PepSeq software within the MassLynx V 4.0 package (Waters, UK).

2.4. pgafp gene characterization

2.4.1. Partial amplification of pgafp gene with degenerate primers

DNA was isolated as described for the genetic identification of the strain in Section 2.2. PCRs were performed with genomic DNA (100 ng) in 50 μl reaction mixtures containing 50 pmol each of forward and reverse primers, 0.5 mM each of the dNTPs, 0.1 vol of 10× PCR buffer (22.5 mM $MgCl_2$, 500 mM Tris–HCl pH 9.2, 140 mM ammonium sulfate), 1.8 mM $MgCl_2$, and 2.5 units pfu turbo polymerase (Stratagene, La Jolla, CA). The degenerate primers *pgafp*-DPF1 and *pgafp*-DPR1A (Table 1) were designed from amino acid sequences obtained by ESI-MS/MS. Three reactions were done: one with both forward and reverse primers, another only with forward, and a third one only with reverse primer. DNA amplification was carried out in a Px2 thermocycler (Thermo Scientific, Cramlington, UK). The PCR program consisted of initial denaturation (94 °C for 5 min), 35 cycles of denaturation (94 °C for 5 s), annealing (45–57 °C range for 5 s, gradient 12), and extension (72 °C for 1 s) steps, followed by a final extension at 72 °C for 4 min. PCR products were electrophoresed on 2% agarose gels with MetaPhor agarose (BMA, Rockland, ME). Two different markers were used for agarose gels: DNA marker from Gibco (0.15–12 kb bands) (Gibco, Grand Island, NY) and DNA size standard from Bio-Rad (0.05–2 kb) (Bio-Rad Laboratories, Madrid, Spain). PCR products detected only in the reactions with both reverse and forward primers were gel-purified and cloned into the pCR2.1-TOPO vector (Invitrogen, Barcelona, Spain). One Shot TOP10 Chemically Competent *E. coli* (Invitrogen) were transformed with that vector. To confirm the insert in the vector, digestions with restriction enzyme EcoRI (Roche, Madrid, Spain) or PCR with M13 forward and reverse primers (Invitrogen) were carried out. Finally, vectors were sequenced in the Institute of Biomedicine (CSIC, Valencia, Spain).

Table 1
Degenerate primers used to amplify the *pgafp* gene.

Primer names	Amino acid sequences and designed primers[a]						
pgafp-DPF1	L 5′-TN	K AAR	H CAY	N AAY	T ACN	C TGN	T AC-3′
pgafp-DPR1A	C 3′-CR	G CCN	S AGN	A CGN	A CGN	N TT-5′	

[a] R for purine, Y for pyrimidine, N for A, C, T, or G.

2.4.2. Rapid amplification of cDNA ends (RACE-PCR)

RNA was obtained after growing *P. chrysogenum* RP42C in MEB at 25 °C for 5 days. Mycelium was quickly frozen with liquid nitrogen and broken with mortar and pestle. RNA extraction was performed with RNeasy Plant Mini Kit (Qiagen, Crawley, UK) following the manufacturer's instructions. The final extracted RNA was resuspended in 40 μl of RNase-free deionized water.

Amplification of 5′- and 3′-ends of *pgafp* gene was carried out using the SMART RACE cDNA Amplification Kit (Clontech Laboratories, Saint Germain en Laye, France). Two separate populations of cDNA (5′- and 3′-RACE-Ready cDNAs) were synthesized according to the amplification kit instructions.

A gene-specific primer was designed for amplifying the 3′-end (3′-RACE GSP) from the genomic DNA sequence previously obtained by PCR with degenerate primers *pgafp*-DPF1 and *pgafp*-DPR1A. The resulting 3′-end sequence was used to design a gene-specific primer (5′-RACE GSP) to amplify 5′-end.

PCRs were performed with 5′- or 3′-RACE-Ready cDNAs (50 ng) in 50 μl reaction mixtures containing 1 mM each of dNTPs (Roche), 0.5 μM of the 5′- or 3′-RACE GSP primers, 5 μl of Universal Primer A Mix (10×), 0.1 vol of 10× PCR buffer (10 mM Tris–HCl, pH 8.8, 50 mM KCl, 0.1% Triton X-100), 4 mM $MgCl_2$, and 2.5 units pfu turbo polymerase. The PCR program for obtaining the 3′-end consisted of 35 cycles of denaturation (94 °C for 30 s), annealing (68 °C for 30 s), and extension (72 °C for 3 min). The PCR program for 5′-end amplification was: 5 cycles of 2 steps (94 °C for 30 s and 72 °C for 3 min), 5 cycles of 3 steps (94 °C for 30 s, 70 °C for 30 s, and 72 °C for 3 min), and 27 cycles of 3 steps (94 °C for 30 s, 68 °C for 30 s, and 72 °C for 3 min). Finally, 5′- and 3′-RACE-PCR products were gel-purified, cloned, and sequenced as described in Section 2.4.1.

2.4.3. Amplification of the whole pgafp gene

The gene sequence was amplified by PCR using Fwd-full *pgafp* and Rev-full *pgafp* primers. PCR mixtures were composed by 100 ng of genomic DNA, 0.5 μM each of forward and reverse primers, 0.5 mM each of the dNTPs, 0.1 vol of 10× PCR buffer (10 mM Tris–HCl, pH 8.8, 50 mM KCl, 0.1% Triton X-100), 4 mM $MgCl_2$, and 2.5 units pfu turbo polymerase and deionized water in a final volume of 50 μl. The PCR program consisted of 30 cycles of

denaturation (94 °C for 30 s), annealing (65 °C for 30 s), and extension (72 °C for 3 min). PCR products were gel-purified, cloned, and sequenced as described in Section 2.4.1.

2.5. Sequence analysis

The NCBI database (http://blast.ncbi.nlm.nih.gov/Blast.cgi) was used to compare the obtained data against other genomic sequences or proteins. Intron sequences were confirmed by GenScan software (http://genes.mit.edu/GENSCAN.html) [4,5,6]. To translate the open reading frame to protein and to estimate the molecular mass and the isoelectric point of the predicted protein, Translate tool and Compute pI/Mw tool from ExPASy proteomics server at the Swiss Institute of Bioinformatics (http://www.expasy.org) [9] were used. A search for the presence of signal peptides was performed with SignalP version 3.0 (http://www.cbs.dtu.dk/services/SignalP/) [2,24].

The tertiary structure of PgAFP was predicted by the computer program to extract 3D models CPHmodels 3.0 server (http://www.cbs.dtu.dk/services/CPHmodels/) [18]. To edit the pdb file the ICM-Browser 3.6-1d program (Molsoft, La Jolla, CA) was used.

2.6. Nucleotide sequence accession number

The nucleotide sequence of the *pgafp* gene has been deposited in the GenBank database under accession number GQ911150.

3. Results

3.1. Genetic identification of P. chrysogenum RP42C

The ITS sequence from RP42C strain showed 100% of identity and a maximum score with that of *P. chrysogenum* strain L1 (GenBank accession no. FJ467371). Thus, the strain RP42C was confirmed as *P. chrysogenum*.

3.2. PgAFP characterization

The pI value was estimated as 9.22 according to the electrophoretic mobility of marker proteins. PgAFP was resolved in ESI-MS as a short series of multiply charged peaks, with the dominant peak set of around 1 083 Da corresponding to the +6 charge state. After data deconvolution, the molecular mass profile shows a dominant peak of 6494 Da, corresponding to the mass of mature PgAFP.

Analysis of PgAFP for glycosylations did not show any evidence of N- or O-linked-oligosaccharides. No differences on gel-mobility of PgAFP were observed on SDS-PAGE after attempting deglycosylation by enzymatic means using endoglycosidase F1 (data not shown), PNGase F, or the E-DEGLY deglycosylation kit. Furthermore, no change in the mass spectrum profiles was obtained in ESI-MS when PgAFP was exposed either to endoglycosidase F1 or PNGase F or to chemical deglycosylation (data not shown).

The N-terminal sequence obtained from PgAFP by Edman degradation was LSKFGGECSLKH. In addition, five *de novo* peptide sequences were obtained by ESI-MS after treatments with trypsin and chymotrypsin (Table 2), with no major differences in-gel compared to in-solution digestions. The sequences 1a and 1b (Table 2) matched the N-terminal end. The calculated molecular mass of all the peptides (43 amino acids) was 4.8 kDa, 74% of the 6.5 kDa mass observed by ESI-MS.

3.3. Characterization of the gene encoding PgAFP

Partial amplification of the *pgafp* gene was done using the two degenerate primers, *pgafp*-DPF1 and *pgafp*-DPR1A (Table 1). After different tests, the optimal annealing temperature was set at 45.1 °C. PCR products run in agarose gel showed one small band of 70 bp that was amplified only in reactions with both forward and reverse primers. The sequence of the amplified product was 5′-TGAAGCATAATACGTGCACATACCTAAAGGGTGGAAAGAACCATGTAGTCAATTGCGGTTCTGCCGCTAA-3′. The deduced peptide (KHNTCTYLKGGKNHVVNCGSAA), with a predicted mass of 2303 Da, included both amino acid sequences used to design the degenerate primers, but neither N- nor C-end of PgAFP.

Both 3′- and 5′-ends of the gene *pgafp* were obtained by RACE-PCR from cDNA. For amplifying the 3′-end, the gene-specific primer 3′-RACE GSP (5′-GGGTGGAAAGAACCATGTAGTCAATTGCG-3′) was designed from the above described *pgafp* partial sequence of genomic DNA described above. The product of 450 bp included 121 bp from the 3′-end of *pgafp* that codes for 39 amino acids and ends with a stop codon. Amplification of the 5′-end was carried out with a new gene-specific primer, 5′-RACE GSP (5′-AACTGGGGTCTGGCAGTCAACCCTC-3′), designed from 3′-end sequence. The product of approximately 440 bp included a segment of 279 bp corresponding to the entire cDNA sequence. The predicted amino acid sequence confirmed that the obtained segment encoded all fragments from PgAFP described in Section 3.2 (Table 2).

Amplification from genomic DNA was performed using two new primers Fwd-full *pgafp* (5′-ATGCAGATCACCAGCATTGCC-3′) and Rev-full *pgafp* (5′-TCAAACTGGGGTCTGGCAGTC-3′) designed from the 279 bp segment. The amplification product from genomic DNA was 404 bp. The comparison of genomic and cDNA sequences revealed two introns of 63 and 62 bp (Fig. 1).

3.4. Analysis of the deduced protein

The predicted protein sequence comprised 92 amino acids (Fig. 1). The analysis of this protein predicted a presequence consisting of the first 18 amino acids (MQITSIAIVFFAAMGAVA). Amino acids 35–53 matched the N-terminal sequence of PgAFP. Amino acids 19–34 (NPIARESDDLDARDVQ), putatively correspond to a prosequence not present in the mature protein. The predicted PgAFP protein had a calculated mass of 6500 Da and a deduced pI of 8.8.

Table 2
De novo obtained peptide sequences obtained by ESI-MS/MS from trypsin and chymotrypsin digestions from PgAFP[a].

Sequence Number	*De novo* sequence	Peptide Mass	Charge state	Digestion enzyme
1a	LSKFGGECSL	549.26	+2	Chymotrypsin
1b	FGGECSLK	449.24	+2	Trypsin
2	HNTCTYLK	490.26	+2	Trypsin
3a	NHVVNCGSAANKK	466.88	+3	Trypsin
3b	NHVVNCGSAANK	607.308	+2	Trypsin
4	-yDEHHK	468.19	+2	Trypsin
5	-cqTPV	540.92	+3	Chymotrypsin

- indicates the presence of additional unassigned amino acids.
Lower case letters indicate amino acids which were tentative assignments.
[a] L and I are interchangeable here since these two residues are isobaric.

```
  1 atgcagatcaccagcattgccattgtcttcttcgccgcaatgggtgcggttgctaaccccatcgcgagggagtcggacgatcttgatgcccgagacgtaca 101
    M  Q  I  T  S  I  A  I  V  F  F  A  A  M  G  A  V  A  N  P  I  A  R  E  S  D  D  L  D  A  R  D  V  Q
102 gcttagtaaattcggaggagtaagttcttcttacaagacgtctatatagaaatagcactaacctttctgaaccactttacaggaatgcagcttgaaacaca 202
     L  S  K  F  G  G <<                       Intron 1                          >>E  C  S  L  K  H  N
203 acacgtgcacatacctaaagggtggaaagaaccatgtagtcaattgcggttcggccgccaacaagaaggtaggttccgattcgattcggggccaattgatt 303
      T  C  T  Y  L  K  G  G  K  N  H  V  V  N  C  G  S  A  A  N  K  K <<                     Intron 2
304 tgttcttatcatttaatcttcatctacagtgcaagtctgatcgccaccactgtgaatacgatgagcaccacaagagggttgactgccagaccccagtttga 404
                               >>C  K  S  D  R  H  H  C  E  Y  D  E  H  H  K  R  V  D  C  Q  T  P  V  -
```

Fig. 1. Full-length DNA sequence of *pgafp* gene. The open reading frame has been translated to amino acids. Consensus splice sequences and internal marker of fungal introns are underlined.

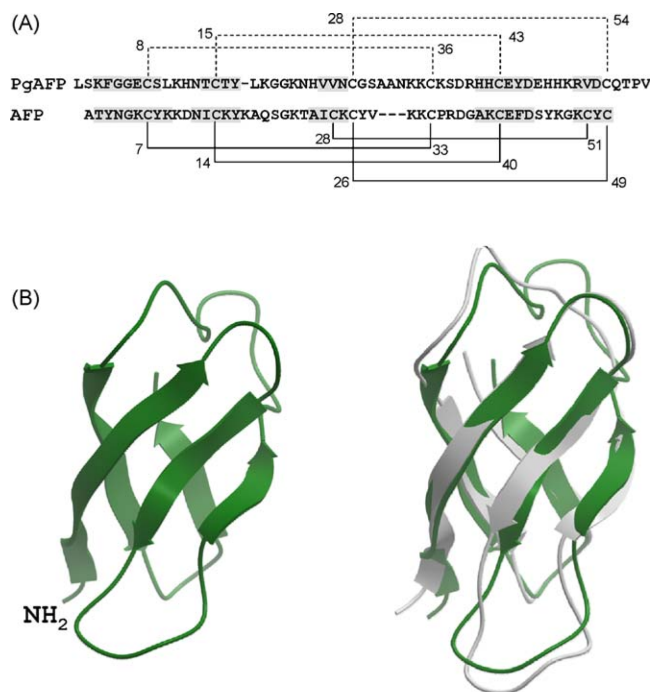

Fig. 2. (A) Amino acid sequence alignment of PgAFP and AFP. Disulfide bridges of AFP as described Campos-Olivas et al. [7] are indicated with solid lines. Putative disulfide bridges of PgAFP are indicated with dashed lines. (B) Predicted tertiary structure of PgAFP (left) and superposition of PgAFP and AFP structures (right).

The prediction of the tertiary structure of PgAFP with the server CPHmodels 3.0 used the NMR structure of AFP (pdb accession code 1AFP) as a template. The 3D model obtained (Fig. 2) suggested a structure consisting of five antiparallel β-strands defining a small and compact β-barrel.

4. Discussion

The ITS sequence confirmed that the strain RP42C belongs to the species *P. chrysogenum*, as it was proposed by the previous identification based on morphology and biochemical characteristics [1].

The molecular mass of PgAFP obtained by ESI-MS was close to 6.5 kDa which is lower than the 9.1 kDa previously reported for this protein by SDS-PAGE [1]. For large molecules, such as intact proteins, ESI-MS is capable of determining the mass to within 0.01% of the total molecular mass. For a mass of 6 494 Da, this is 0.65 Da. Therefore, it seems likely that PgAFP migrates anomalously in SDS-PAGE giving an apparent weight of 9.1 kDa. Many characteristics of proteins can cause them to migrate atypically on SDS-PAGE. Lipoproteins and glycoproteins migrate anomalously because SDS cannot bind with lipids or oligosaccharides [11]. No evidence of N- or O-linked oligosaccharides was found either by SDS-PAGE or MS analysis of PgAFP after deglycosylation treatments. The comparison between predicted and experimental masses of PgAFP, discussed later, confirmed that PgAFP is not a glycosylated protein. Atypical migration on SDS-PAGE can be due to other causes that do not involve post-translational modifications. For example, proteins with extreme isoelectric points can show a higher apparent molecular weight (around 25% higher) due to electrostatic repulsions causing less SDS to bind and a consequent decrease in their migration velocity [22]. Analysis of the protein by capillary isoelectric focusing electrophoresis confirmed the cationic character of the protein, with a pI of 9.22. This cationic characteristic, common to antifungal proteins from molds [20], has been directly related to a strong antimicrobial activity [13]. Given that the mass and isoelectric point are very close to those of the other known small antifungal proteins from molds [20,28], PgAFP seems to belong to this group of antifungal proteins from ascomycetes.

Amino acid sequences of PgAFP peptides 1a and 1b determined using tandem MS and *de novo* sequencing (Table 1) were contained in the N-terminal sequence obtained by Edman degradation. The peptide ending in TPV is likely to be the C-terminal sequence, given that all other tryptic fragments have arginine (R) or lysine (K) as C-terminal residues. For the initial amplification of *pgafp* gene, degenerate primers *pgafp*-DPF1 and *pgafp*-DPR1A were designed from the PgAFP amino acid sequences obtained (Table 1). To avoid the use of highly degenerate primers, the nine first amino acids at the N-terminal from the protein were not considered. Thus, primer *pgafp*-DPF1 was designed from sequences 1b and 2. Additionally, the reliable inferred sequence for the C-terminal was limited to TPV. For this, the reverse primer was designed from the inner sequence 3. Therefore, the 70 bp fragment obtained from genomic DNA would include neither 3′- nor 5′-ends of *pgafp* gene. The translated sequence included all expected amino acid, with a predicted molecular mass (2303 Da), less than half of PgAFP mass (6494 Da). To establish the entire coding sequence of *pgafp*, both ends were obtained by RACE-PCR using two gene-specific primers designed from the 70 bp fragment.

To obtain appropriate cDNA populations from *pgafp*, a 5 days-old culture of *P. chrysogenum* RP42C was used. mRNA should be actively produced at this time given that PgAFP was detected after 7 days of incubation (unpublished results).

The sequence of 121 bp obtained by 3′-RACE-PCR coded for GGKNHVVNCGSAANKKCKSDRHHCEYDEHHKRVDCQTPV which includes all of the amino acids expected from the *de novo* sequence data (Table 2). Therefore, this nucleotide fragment was confirmed as the 3′-end. Given the small size of the protein, a new primer 5′-RACE GSP was designed right from the 3′-end of *pgafp* to obtain the open reading frame of the gene, including the 5′-end. The segment of 279 bp obtained showed similar size to cDNA related to other antifungal proteins produced by molds [20,28]. The analysis of this segment revealed the whole primary structure of the mature PgAFP (Fig. 1). The primers Fwd-full *pgafp* and Rev-full *pgafp* designed from cDNA ends allowed the entire structural region of *pgafp* to be amplified from genomic DNA.

Comparison of the *pgafp* genomic (404 bp) and cDNA (279 bp) sequences revealed two introns of 63 and 62 bp (Fig. 1), similar to genes encoding other antifungal proteins from molds [20,28]. The

```
PgAFP   ---MQITSIAIVFFAAMGAVANPIARE---SDDLDARD--VQLSKFGGECSLKHNTCTYLK-GGKNHVVNCGSAANKKCKSDRHHCEYDEHHKRVDCQTPV
Anafp   ----------------------------------------LSKYGGECSLEHNTCTYRK-DGKNHVVSCPSAANLRCKTDRHHCRYDDHHKTVDCQTPV
PAF     ---MQITTVALFLFAAMGGVATPIESV---SNDLDARAEAGVLAKYTGKCTKSKNECKYKNDAGKDTFIKCPKFDNKKCTKDNNKCTVDTYNNAVDCD---
AFP     MKFVSLASLGFALVAALGAVATPVEADSLTAGGLDARDESAVLATYNGKCYKKDNICKYKAQSGKTAICKCYV---KKCPRDGAKCEFDSYKGKCYC----
AcAFP   MKFVSLASLGFALVAALGVVASPVDADSLAAGGLDARDESAVQATYDGKCYKKDNICKYKAQSGKTAICKCYV---KVCPRDGAKCEFDSYKGKCYC----
```

Fig. 3. Alignment of PgAFP sequence with the most similar antifungal proteins: Anafp from *A. niger* [16], PAF from *P. chrysogenum* (GenBank accession no. AAA92718), AFP from *A. giganteus* (P17737), and AcAFP from *A. clavatus* (EF600065).

small size of these introns compared to those of mammalian introns is a typical feature of fungal genes [12]. The 5′- and 3′-ends of the two introns are quite similar (Fig. 1) and coincide with the consensus splice sequences for fungal introns: 5′-splice donor site GTNNGT (N = A, G, T or C) and the consensus 3′-splice acceptor site DYAG (D = A, G or T; Y = C or T). In addition, the first intron has the internal CTAAC sequence, which is commonly found in fungal introns [31].

The *pgafp* open reading frame encodes a 92 amino acid-long precursor protein (preproPgAFP). The first 18 amino acids were not found in the *de novo* sequencing of the mature protein (Table 2), and correspond to a predicted signal sequence. Similarly, residues 19–34 were not found, and residues 35–46 correspond to the N-terminal sequence obtained from Edman degradation analysis (Table 2). Thus, the 16 amino acids from residues 19–34 constitute a prosequence that would be removed before or during the release of mature PgAFP [19]. Prosequences play an important role by preventing protein activity before secretion. As it has been proposed for PAF [21], the mature protein would adopt its active conformation after the prosequence is cleaved off. Antifungal proteins produced as preproproteins from Ascomycetes include PAF, AFP, and AcAFP [19,23,28].

The deduced mature PgAFP has 58 amino acid residues. The high percentage of basic amino acids in the deduced sequence (27.6%) is consistent with the pI of 9.2 estimated by isoelectrofocusing. Similarly, a pI of 8.8 was deduced from the 58 amino acid sequence using the Compute pI tool. Even though this prediction tool may not be precise for highly basic and small proteins [9], both pI values reveal a net positive charge under physiological conditions, which is a common characteristic among antifungal proteins from molds.

The comparison of the mature PgAFP sequence with other described proteins (Fig. 3) showed the highest identity (79%) with the antifungal protein Anafp from *A. niger*. Interestingly, the identity of PgAFP with PAF, also produced by *P. chrysogenum*, was only 34%. Although the identity with AFP was only 33%, this was the most similar protein with 3D structure available from public databases. The model obtained using AFP as template (Fig. 2) showed a structure with five antiparallel β-strands forming a compact β-barrel. As has been demonstrated for AFP [7,14], the PgAFP structure would be stabilized by internal disulfide bridges formed by cysteine residues. In fact, PgAFP contains six cysteines conserved in the group of antifungal proteins from molds (Fig. 3). In addition, the loss of 6 hydrogen atoms to form these three disulfide bridges would explain the mass difference between the predicted mature protein (6500 Da) and the experimental value of 6494 Da determined by mass spectrometry. Therefore, a pattern of disulfide bridges between cysteins pairs 8–36, 15–43, 28–54 is suggested (Fig. 2).

The recent report of the whole genome sequence of *P. chrysogenum* Wisconsin 54–1255 [29] revealed a nucleotide sequence (GeneBank accession no. CAP80456) similar to *pgafp*. A difference of 11 nucleotides, 7 in the coding region, was observed between both sequences. Consequently, the predicted protein from strain Wisconsin 54–1255 differs in only one amino acid in the signal peptide and two more in the prosequence. However, the production of the protein by the Wisonsin 54–1255 strain has not been demonstrated, and no characterization is available.

P. chrysogenum produces compounds classified as generally recognized as safe (GRAS) by the US Food and Drug Administration, and is commonly used in dry-cured foods. Thus, PgAFP may be useful to prevent surface growth of toxigenic molds by immersion or spraying treatments in mold-ripened foods. However, the safety of PgAFP should be established before any industrial application.

In conclusion, PgAFP from *P. chrysogenum* is a new non-glycosylated antifungal protein showing common characteristics to the group of small, basic, and cysteine-rich antifungal proteins from molds. The structural region of the gene *pgafp* encodes as a precursor with a peptide signal and a prosequence that are not present in the mature protein. The characterization of *pgafp* gene will permit the development of new strategies to promote the industrial application of PgAFP in different potential fields.

Acknowledgments

This work was supported by the Spanish Ministry of Education and Science (AGL2004-06546-ALI), INIA (RM2006-00013-00-00), and FEDER. Andrea Rodríguez-Martín was the recipient of a FPI grant from the Spanish Ministry of Education and Science. The authors would like to thank Prof. Ian Connerton (Division of Food Sciences) for his support during Andrea Rodríguez-Martín's stay at Nottingham University.

References

[1] Acosta R, Rodríguez-Martín A, Martín A, Núñez F, Asensio MA. Selection of antifungal protein-producing molds from dry-cured meat products. Int J Food Microbiol 2009;135:39–46.
[2] Bendtsen JD, Nielsen H, von Heijne G, Brunak S. Improved prediction of signal peptides: SignalP 3.0. J Mol Biol 2004;340:783–95.
[3] Benito MJ, Connerton IF, Córdoba JJ. Genetic characterization and expression of the novel fungal protease, EPg222 active in dry-cured meat products. Appl Microbiol Biotechnol 2006;73:356–65.
[4] Burge C, Karlin S. Prediction of complete gene structures in human genomic DNA. J Mol Biol 1997;268:78–94.
[5] Burge CB, Karlin S. Finding the genes in genomic DNA. Curr Opin Struct Biol 1998;8:346–54.
[6] Burge CB. Modeling dependencies in pre-mRNA splicing signals. In: Salzberg S, Searls D, Kasif S, editors. Computational methods in molecular biology. Amsterdam: Elsevier Science B.V.; 1998. p. 127–63.
[7] Campos-Olivas R, Bruix M, Santoro J, Lacadena L, Martínez del Pozo A, Gavilanes JG, et al. NMR solution structure of the antifungal protein from *Aspergillus giganteus*: evidence for cysteine pairing isomerism. Biochemistry 1995;34:3009–21.
[8] Deepak AV, Thippeswamy G, Shivakameshwari MN, Salimath BP. Isolation and characterization of a 29-kDa glycoprotein with antifungal activity from bulbs of *Urginea indica*. Biochem Biophys Res Commun 2003;311:735–42.
[9] Gasteiger E, Hoogland C, Gattiker A, Duvaud S, Wilkins MR, Appel RD, et al. Protein identification and analysis tools on the ExPASy Server. In: Walker JM, editor. The proteomics protocols handbook. New Jersey: Humana Press; 2005. p. 571–608.
[10] Geisen R. *P. nalgiovense* carries a gene which is homologous to the paf gene of *P. chrysogenum* which codes for an antifungal peptide. Int J Food Microbiol 2000;62:95–101.
[11] Grefrath SP, Reynolds JA. The molecular weight of the major glycoprotein from the human erythrocyte membrane. Proc Natl Acad Sci USA 1974;71:3913–6.
[12] Gurr SJ, Uncles SE, Kinghorn JR. The structure and organization of nuclear genes of filamentous fungi. In: Kinghorn JR, editor. Gene structure in eukaryotic microbes. Oxford: IRL Press; 1987. p. 93–193.
[13] Jiang Z, Vasil AI, Hale J, Hancock REW, Vasil ML, Hodges RS. Effects of net charge and the number of positively charged residues on the biological activity of amphipathic α-helical cationic antimicrobial peptides. Biopolymers (Peptide Science) 2008;90:369–83.
[14] Lacadena J, Martínez del Pozo A, Gasset M, Patiño B, Campos-Olivas R, Vázquez C, et al. Characterization of the antifungal protein secreted by the mold *Aspergillus giganteus*. Arch Biochem Biophys 1995;20:273–81.
[15] Laemmli UK. Cleavage of structural proteins during the assembly of the head of bacteriophage T4. Nature 1970;227:680–5.

[16] Lee DG, Shin SY, Maeng C-Y, Jin ZZ, Kim KL, Hahm K-S. Isolation and characterization of a novel antifungal peptide from *Aspergillus niger*. Biochem Biophys Res Commun 1999;263:646–51.
[17] Leis O, Madrid JF, Ballesta J, Hernández F. *N*- and *O*-linked oligosaccharides in the secretory granules of rat paneth cells: an ultrastructural cytochemical study. J Histochem Cytochem 1997;45:285–94.
[18] Lund O, Nielsen M, Lundegaard C, Worning P. CPHmodels 2.0: X3M a Computer Program to Extract 3D Models. Abstract CASP5 conference; 2002. p. A102.
[19] Marx F, Haas H, Reindl M, Stöffler G, Lottspeich F, Redl B. Cloning, structural organization and regulation of expression of the *Penicillium chrysogenum* paf gene encoding an abundantly secreted protein with antifungal activity. Gene 1995;167:167–71.
[20] Marx F. Small, basic antifungal proteins secreted from filamentous ascomycetes: a comparative study regarding expression, structure, function and potential application. Appl Microbiol Biotechnol 2004;65:133–42.
[21] Marx F, Binder U, Leiter E, Pocsi I. The *Penicillium chrysogenum* antifungal protein PAF, a promising tool for fungal cell biology studies and the development of new antifungal therapies. Cell Mol Life Sci 2008;65:445–54.
[22] Monstadt GM, Holldorf AW. Arginine deiminase from *Halobacterium salinarum*: purification and properties. Biochem J 1990;273:739–45.
[23] Nakaya K, Omata K, Okahashi I, Nakamura Y, Kolkenbrock H, Ulbrich N. Amino acid sequence and disulfide bridges of an antifungal protein isolated from *Aspergillus giganteus*. Eur J Biochem 1990;193:31–8.
[24] Nielsen H, Engelbrecht J, Brunak S, Von Heijne G. Identification of prokaryotic and eukaryotic signal peptides and prediction of their cleavage sites. Protein Eng 1997;10:1–6.
[25] Pitt JI. A laboratory guide to common Penicillium species. North Ryde: Commonwealth Scientific and Industrial Research Organization; 1986.
[26] Rademaker GJ, Thomas-Oates J. Analysis of glycoproteins and glycopeptides using fast-atom bombardment. Methods Mol Biol 1998;61:231–41.
[27] Skouri-Gargouri H, Gargouri A. First isolation of a novel thermostable antifungal peptide secreted by *Aspergillus clavatus*. Peptides 2008;29: 1871–7.
[28] Skouri-Gargouri H, Ali MB, Gargouri A. Molecular cloning, structural analysis and modelling of the antifungal peptide from *Aspergillus clavatus*. Peptides 2009;30:1798–804.
[29] Van den Berg MA, Albang R, Albermann K, Badger JH, Daran JM, Driessen AJ, et al. Genome sequencing and analysis of the filamentous fungus *Penicillium chrysogenum*. Nat Biotechnol 2008;26:1161–8.
[30] White TJ, Bruns T, Lee S, Taylor J. Amplification and direct sequencing of fungal ribosomal RNA genes for phylogenetics. In: Innis MA, Gelfand DH, Sninsky JJ, White TJ, editors. PCR protocols: a guide to methods and applications. San Diego: Academic Press; 1990. p. 315–22.
[31] Wiesner P, Beck J, Beck KF, Ripka S, Müller G, Lücke S, et al. Isolation and sequence analysis of the fatty acid synthetase FAS2 gene from *Penicillium patulum*. Eur J Biochem 1988;177:69–79.
[32] Wnendt S, Ulbrich N, Stahl U. Molecular cloning, sequence analysis and expression of the gene encoding an antifungal-protein from *Aspergillus giganteus*. Curr Genet 1994;25:519–23.
[33] Yamunarani K, Jaganathan R, Bhaskaran R, Govindaraju P, Velazhahan R. In vitro antifungal activity of a 29-kDa glycoprotein purified from the galls of *Quercus infectoria*. Acta Phytopathol Entomol Hung 2005;40:43–54.