Identification and characterization of GEO1, a new class II hydrophobin from *Geosmithia* spp.

Priscilla P. Bettini, Arcangela Frascella, Cecilia Comparini, Lara Carresi, Alessia L. Pepori, Luigia Pazzagli, Gianni Cappugi, Felice Scala, and Aniello Scala

Abstract: In the present paper we describe a new noncatalytic protein belonging to the hydrophobin family, designated GEO1, purified from the culture filtrate of *Geosmithia pallida* (Ascomycota: Hypocreales), and the corresponding gene sequence. In the fungal genome, GEO1 was encoded by a single-copy gene with a 450 bp open reading frame interrupted by 2 small introns whose primary translation product was 109 amino acids long and included a 23 amino acids signal peptide. The mature protein had a molecular mass of 8111.75 Da and a theoretical pI of 4.33. The deduced amino acid sequence showed similarity to class II hydrophobins and contained 8 conserved cysteine residues, present in all hydrophobins isolated so far. Biochemical properties, such as foam-forming ability and trapezoid-like shape of a GEO1 drop, also resembled the typical features of the class II hydrophobins. Expression of the *geo1* gene was assessed after 2, 4, 7, 9, and 11 days of culture and showed that the *geo1* transcript appeared after 7 days and increased up to 11 days.

Key words: Ascomycetes, *Geosmithia pallida*, class II hydrophobin.

Résumé : Nous décrivons dans cet article une nouvelle protéine non-catalytique appartenant à la famille des hydrophobines et appelée GEO1, purifiée à partir d'une filtration de culture de *Geosmithia pallida* (Ascomycota: Hypocreales), ainsi que la séquence du gène correspondant. GEO1 était codée dans le génome du champignon par un gène simple copie comportant un CLO de 450 pb interrompu par deux petits introns, dont le principal produit de la traduction avait une longueur de 109 acides aminés et comprenait un peptide signal de 23 acides aminés. La protéine mature avait un poids moléculaire de 8111,75 Da et un pI théorique de 4,33. La séquence déduite en acides aminés montrait des similarités avec les hydrophobines de classe II et contenait huit résidus cystéine, présents dans toutes les hydrophobines isolées jusqu'à présent. Les propriétés biochimiques, comme la capacité de mousser et la forme trapézoïde d'une goutte de GEO1, ressemblaient aussi aux caractéristiques typiques des hydrophobines de classe II. Le décours temporel d'expression de *geo1* a été évalué après 2, 4, 7, 9 et 11 jours en culture et a montré que le transcrit de *geo1* apparaissait au jour 7 et augmentait jusqu'au 11ième jour.

Mots-clés : Ascomycètes, *Geosmithia pallida*, hydrophobine de classe II.

[Traduit par la Rédaction]

Introduction

Hydrophobins are low molecular mass, ubiquitous proteins produced and secreted by filamentous fungi. Based on their hydropathy patterns and differences in their solubility and amino acid sequences, hydrophobins have been divided in 2 classes. Class I hydrophobins are produced by Ascomycetes and Basidiomycetes; class II hydrophobins are produced by Ascomycetes only (Whiteford and Spanu 2002; Linder et al. 2005). Class I and class II are now considered as 2 separate protein families (Pfam PF01185 and PF06766, respectively), as sequence homology between proteins of the 2 classes is limited except for 8 conserved cysteine residues forming 4 intramolecular disulfide bonds involved in protein folding and stabilization (Linder et al. 2005). However, a genome-wide analysis of *Trichoderma* spp. has recently identified new hydrophobins related to class I but which form a distinct clade,

P.P. Bettini. Dipartimento di Biologia Evoluzionistica "Leo Pardi", University of Florence, via Romana 17-19, 50125 Florence, Italy.
A. Frascella. Dipartimento di Biologia Evoluzionistica "Leo Pardi", University of Florence, via Romana 17-19, 50125 Florence, Italy; Dipartimento di Biotecnologie Agrarie, Sezione di Protezione delle Piante, University of Florence, via della Lastruccia 10, 50019 Sesto Fiorentino, Florence, Italy.
C. Comparini, L. Carresi, and A. Scala. Dipartimento di Biotecnologie Agrarie, Sezione di Protezione delle Piante, University of Florence, via della Lastruccia 10, 50019 Sesto Fiorentino, Florence, Italy.
A.L. Pepori. Istituto per la Protezione delle Piante, National Research Council, via Madonna del Piano 10, 50019 Sesto Fiorentino, Florence, Italy.
L. Pazzagli and G. Cappugi. Dipartimento di Scienze Biochimiche, University of Florence, Viale Morgagni 50, 50134 Florence, Italy.
F. Scala. Dipartimento di Arboricoltura, Botanica e Patologia Vegetale, Sezione di Patologia Vegetale, University of Naples "Federico II" via Università 100, 80055 Portici, Naples, Italy.

Corresponding author: Priscilla P. Bettini (e-mail: p.bettini@unifi.it).

thus raising the possibility of a revision of the present classification (Seidl-Seiboth et al. 2011). Hydrophobins can fulfill multiple roles in fungal growth and developmental processes, as well as in the interaction between fungi and their hosts, by forming an amphipathic membrane at the hydrophobic–hydrophilic interface, with the hydrophobic side facing outwards (Sunde et al. 2008). Hydrophobins take part in the formation of fungal aerial structures (Ng et al. 2000; Karlsson et al. 2007), in the development of microsclerotia and microconidial chains (Fuchs et al. 2004; Klimes et al. 2008), in the dispersion of conidia by wind or water (Wessels 2000), and in the maturation of the fungal cell wall (van Wetter et al. 2000). They are also required for the establishment of the lichen symbiosis (Dyer 2002) and in the interaction between ectomycorrhizal fungi and their host trees (Mankel et al. 2002). Moreover, since they render surfaces hydrophobic, hydrophobins favour the attachment of fungi to the hydrophobic surfaces of the hosts (Temple and Horgen 2000; Izumitsu et al. 2010) and, in some cases, are involved as the primary determinants of pathogenesis (Talbot 2003; Kim et al. 2005; Aimanianda et al. 2009). The polyphyletic genus *Geosmithia* Pitt (Ascomycota: Hypocreales) actually includes 22 published and at least 20 unpublished species (Kolařík et al. 2011). *Geosmithia* species occur mostly in galleries built by phloem-feeding bark beetles, such as scolytids and bostrichids; these galleries are their primary habitat (Kolařík et al. 2007, 2008), but they can also be found on wood, soil, and food. Inside host trees *Geosmithia* spp. can coexist with other fungal species, as shown by Scala et al. (2007) who isolated *Geosmithia pallida* from an elm tree showing symptoms of Dutch elm disease. Interestingly, this isolate also harboured a copy of the cerato–ulmin gene probably derived from a horizontal gene transfer event between *Geosmithia* and the Dutch elm disease fungus *Ophiostoma novo-ulmi*. The *Geosmithia* species analyzed so far are nonpathogenic endophytes, with the notable exception of *Geosmithia morbida*, a new species associated with the walnut twig beetle (*Pityophthorus juglandis*), which is responsible for the development of the so-called thousand canker disease in black walnut (Kolařík et al. 2011). These fungi show a strong association with beetle species living on angiosperms and conifers that have a fundamental role in ensuring the dispersal of conidia. In the present paper we report on the isolation of a previously unknown class II hydrophobin, which we have designated as GEO1, the first to be described from the genus *Geosmithia*, from the culture filtrate of the species *G. pallida*. The corresponding gene sequence was also cloned and characterized.

Materials and methods

Fungal strains and culture conditions

Conditions for growth and maintenance of the *G. pallida* (G. Sm.) Kolařík, Kubátová & Pažoutová, comb. nov., isolate IVV7, used in this study were as previously described (Scala et al. 2007).

Protein extraction

For protein extraction the fungus was grown in 100 mL flasks containing 10–20 mL of modified Takai medium (Scala et al. 2007) for 14 days at 25 °C, in agitation. Culture filtrate, obtained by removing mycelium and spores by filtra-

tion on a 0.45 μm membrane (Millipore, Billerica, Massachusetts, USA), was subjected to precipitation with 10% trichloroacetic acid at –20 °C for 30 min. Samples were centrifuged at 11 000g for 10 min at 4 °C, and pellets were dried under vacuum (SpeedVac SC110, Savant Instruments Inc.). Finally, dry pellets were extracted with 200 μL of 60% ethanol.

Reversed-phase high-performance liquid chromatography

The *G. pallida* protein was extracted from 500 mL of culture filtrate as described. The amount of protein obtained was determined with the bicinchoninic acid assay. Reversed-phase – high-performance liquid chromatography (RP–HPLC) was carried out with a C4 reversed-phase column with 5 μm resin beads (Vydac, 4.6 mm × 250 mm). Samples were vacuum-dried (Univapo Vacuum Concentrator, Montreal Biotech Inc., Dorval, Quebec, Canada), dissolved in 1 mL of 30% acetonitrile containing 10 mmol/L trifluoroacetic acid (TFA), and applied to the column previously equilibrated with H_2O–TFA. The elution was carried out with a gradient of H_2O–TFA (solvent A) and acetonitrile–TFA (solvent B), with a flux of 0.8 mL/min. The acetonitrile–TFA gradient was as follows: 30% for 5 min, from 30% to 50% in 40 min, from 50% to 100% in 10 min. The column was connected to a photometer set to a 214 nm wavelength.

Matrix-assisted laser desorption ionization-time of flight mass spectrometry

The GEO1 protein was analyzed by matrix-assisted laser desorption ionization time-of-flight (MALDI-TOF) (Bruker Omniflex, Bruker Daltonics Inc., Billerica, Massachusetts, USA) mass spectrometry to precisely determine its molecular mass. A sample amount of 40-60 pmol was dissolved in 2 μL of 50% acetonitrile and 0.1% TFA and was diluted 1:1 in a sinapinic acid matrix. Before the analysis, the mass spectrometer was calibrated with a protein standard in the range of 6.6 to 18.0 kDa.

Protein sequencing

Prior to protein sequencing, the free thiol groups of the protein were carboamidomethylated to avoid the formation of disulfide bonds. To this aim, 3 nmol of the mature *G. pallida* protein was vacuum-dried and dissolved in 90 μL of 25 mmol/L ammonium acetate containing 10 mmol/L dithiothreitol and 6 mol/L guanidinium chloride. The solution was then exposed to N_2 flux and incubated at 56 °C for 45 min. Once returned to room temperature, 10 μL of iodoacetamide in 25 mmol/L ammonium acetate was added to the sample, and the solution was incubated in the dark for 30 min at room temperature. The sample was then purified by RP–HPLC on a C4 column, with the following acetonitrile–TFA gradient: from 30% to 60% in 30 min and from 60% to 100% in 20 min. Finally, the sample was vacuum-dried again, redissolved in H_2O, and analyzed by MALDI-TOF as a control for the reaction. An aliquot (0.5 nmol) of the carboamidomethylated protein was vacuum-dried and resuspended in 80 μL of 100 mmol/L ammonium bicarbonate, pH 8.5. Six microlitres of a trypsin solution (1 mg/mL in 1 mmol/L HCl) was added to the protein, and the sample was incubated at 37 °C for 18 h. The sample was finally purified by RP–HPLC with the following acetonitrile–TFA gradient: from 0% to 20% in

20 min, from 20% to 50% in 40 min, from 50% to 70% in 15 min, from 70% to 100% in 5 min. Protein sequencing was performed by automated Edman sequencing using a Procise Protein Sequencing System (Applied Biosystems Life Technologies, Carlsbad, California, USA), connected to a reversed-phase column HPLC.

Protein characterization

Turbidity of the protein solution after agitation, foaming capability, and film dissolution by ethanol were evaluated. To assess foaming capability 3.5 mL of an aqueous solution of GEO1 (200 µg/mL) in a 25 mL vial were vigorously shaken by vortexing for 5 min. A 200 µg/mL solution of the class II hydrophobin cerato–ulmin was used as a control. To verify film dissolution in ethanol, the same GEO1 solution was poured in a 60-mm-diameter Petri dish. Upon formation of a white superficial film, drops of 95% ethanol were added with a Pasteur pipette. Finally, the shape change of a 40 µL drop of GEO1 solution on a hydrophobic polytetrafluoroethylene surface was monitored for 1 h in ambient conditions and compared with a drop of distilled water.

DNA extraction and polymerase chain reaction

For the extraction of DNA from *G. pallida* isolate IVV7 mycelium, the NucleoSpin Plant II kit (Macherey-Nagel GmbH & Co. KG, Düren, Germany) was used following the manufacturer's instructions. Amplifications were carried out on 50 ng of genomic DNA with 0.2 mmol/L dNTPs, 0.5 µmol/L primers, and 1 unit of *Taq* polymerase (DreamTaq, Fermentas Inc., Burlington, Ontario, Canada) in a PTC-200 thermal cycler (Bio-Rad Laboratories Inc., Hercules, California, USA).

Genome walking

The complete sequence of the *G. pallida* isolate IVV7 hydrophobin gene was obtained by means of the GenomeWalker Universal kit (Clontech Laboratories Inc., Mountain View, California, USA). Four aliquots of fungal genomic DNA (2.5 µg each) were digested with the blunt-end enzymes *Dra*I, *Eco*RV, *Pvu*II, and *Stu*I, and the resulting fragments were ligated to a GenomeWalker Adaptor, provided in the kit. Genome walking in the regions of interest was then performed by 2 rounds of polymerase chain reaction (PCR) with gene-specific primers. For genome walking in the 5′ direction, the primer for the first PCR was 5′-AAACT-CACGGGGCGCGCAGTTGATCGA-3′, and the primer for the nested PCR was 5′-GACATCGACATCGCAGCACT-CAGGGTT-3′. For genome walking in the 3′ direction, the primer for the first PCR was 5′-TATCCATCG-GATCCCCTTGTCCCAGTC-3′, and the primer for the nested PCR was 5′-TTTGCCATCACTGCCGTTCTCTTTGCT-3′. The Adaptor Primer 1 and Nested Adaptor Primer 2, provided in the kit, annealed to the adaptor sequence. The resulting fragments were cloned (TA Cloning kit, Invitrogen Life Technologies, Carlsbad, California, USA) and sequenced as described.

DNA sequencing and bioinformatic analysis

DNA sequencing was performed by Eurofins MWG Operon (Ebersberg, Germany) on either purified (Illustra GFX PCR DNA and Gel Band Purification kit, GE Healthcare Europe GmbH, Freiburg, Germany) or cloned PCR products. The nucleotide sequence of the *geo1* gene was submitted to GenBank under accession No. JQ042234. GenBank homology searches were performed with the BLAST algorithm (Altschul et al. 1990), and amino acid sequences were aligned with MUSCLE (Edgar 2004). Tools at the ExPASy Proteomics Server (Gasteiger et al. 2003) were used for the characterization of the predicted GEO1 protein. The search for putative *cis*-regulatory elements in the 5′-upstream region of the gene was carried out with Patch 1.0 (http://www.gene-regulation.com/pub/programs.html) based on the TRANS-FAC database release 7.0, and MatInspector Release 8.0.4 (Genomatix Software Suite). Only binding sites with a high matrix similarity (≥0.85) were retained.

Southern hybridization

For Southern hybridization, 2.5 µg of total *G. pallida* isolate IVV7 DNA were digested overnight at 37 °C with the restriction endonucleases *Eco*RI and *Hin*dIII, which do not cut the target sequence, were separated by agarose gel electrophoresis, and were transferred to a positively charged nylon membrane (Roche Applied Science, Indianapolis, Indiana, USA). Hybridization was carried out overnight in a Techne HB-2D Hybridizer (Bibby Scientific Ltd., Stone, UK), using as a probe a *geo1* fragment labelled with the PCR DIG Probe Synthesis kit (Roche Applied Science). Procedures for blot hybridization, washing, and chemiluminescent detection with CDP-Star reagent were performed according to the DIG Applications Manual for Filter Hybridization (Roche Applied Science).

RNA extraction and semiquantitative RT–PCR

geo1 expression was evaluated on total RNA extracted from the mycelium of *G. pallida* isolate IVV7 after 2, 4, 7, 9, and 11 days of growth in axenic liquid culture in 50 mL flasks containing 15 mL of Takai culture medium. At each time point, mycelium was collected by filtration and stored at –80 °C. Total RNA was isolated with the RNeasy Plant Mini kit (Qiagen GmbH, Hilden, Germany) and treated with DNase (Amplification Grade DNase I, Sigma Aldrich, St. Louis, Missouri, USA) to completely remove genomic DNA. Transcript level was analyzed by semiquantitative RT–PCR. RT was carried out with 50 ng of total RNA using the iScript cDNA Synthesis kit (Bio-Rad). One microlitre of cDNA was used for each PCR amplification with the following *geo1*-specific primers: 5′-CTGCCTACGACGCCTGCCCCA-3′ (forward) and 5′-ACAACAGGGAGGACGCAGCAAGT-3′ (reverse). The 18S rRNA was used as the endogenous control, since its expression level was not affected by the fungal growth conditions, and it was amplified in the same reaction as the target gene using the Ambion Competimer technique (QuantumRNA, Universal 18S Internal Standard, Ambion Life Technologies, Carlsbad, California, USA). For each primer–target combination, the linear range of amplification was determined empirically, and a cycle number in the middle of the linear range was chosen for use in subsequent experiments (28 cycles). Different ratios of 18S primers to competimers were also tested to obtain an amplification efficiency for 18S cDNA most similar to that of the amplicon under study.

Fig. 1. RP–HPLC (reversed-phase high-performance liquid chromatography) purification of the cerato–ulmin-like protein from *Geosmithia pallida* isolate IVV7. The main peak is indicated as RP1. Column: Dionex, C4, 5 μm, 250 mm × 4.6 mm. Solvent A: 10 mmol/L trifluoroacetic acid (TFA) in water; Solvent B: 10 mmol/L TFA in acetonitrile. Elution gradient is indicated as (—).
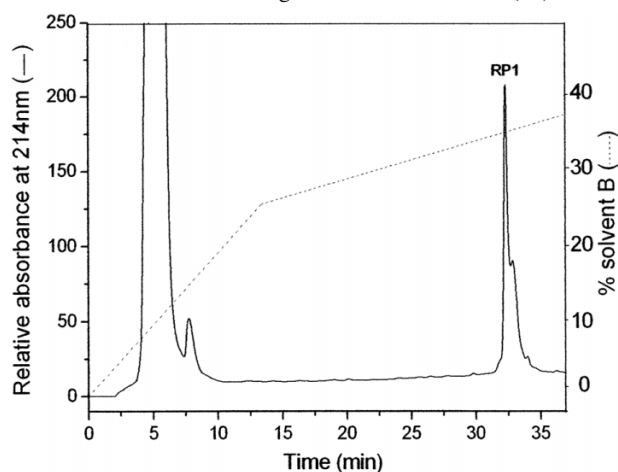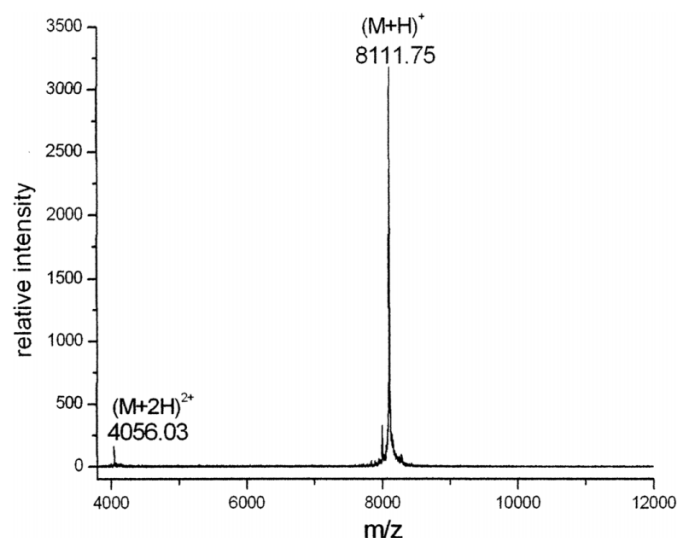

Fig. 2. MALDI-TOF (matrix-assisted laser desorption ionization time-of-flight) mass spectrum of the purified *Geosmithia pallida* protein. Analysis was performed using a LINEAR method. The mass spectrometer was calibrated with a 6.6–18 kDa protein standard. $(M+2H)^{2+}$ represents the double-charged ion of the relative protein.

Results and discussion

Purification and partial sequencing of the *G. pallida* protein

The *G. pallida* protein was purified by RP–HPLC from culture filtrate obtained from 2 large-scale fungal cultures (500 mL each) as described in Materials and methods (Fig. 1). The filtrate was concentrated to 5 mL, and 500 μL aliquots were applied to HPLC. Fractions corresponding to the main peak (RP1) were collected, and protein concentration, as determined by the bicinchoninic acid method, was 1.2 mg/mL eluate. The final yield of protein thus obtained was 2.4 mg. MALDI-TOF mass spectrometry was used to determine the exact molecular mass of the purified GEO1 pro-

Fig. 3. Nucleotide sequence of the *Geosmithia pallida* IVV7 *geo1* gene and deduced amino acid sequence of the encoded protein. The translation start and stop codons are in bold; introns are in lowercase. Putative regulatory elements are highlighted as follows: TATA box, shaded black; CAAT boxes, shaded dark grey; stress-responsive element, shaded light grey. The 8 conserved cysteine residues are indicated by asterisks and the signal peptide is underlined.

```
-462 CATGGGGCTTCGACACGGAACAAACACCCATCTCGACGTGGCTGACGG
-414 GACGAAGTTGGCAATTCCGTCCAACTATGGCTACCACCCCGATTTCCA
-366 GGCCCGCCTCATTTACTCGAGATCGGTAGAGAACAATCAATATTGACA
-318 AGGCCGACTCACATGCACGCGCCCATCGTGTACTTGGCAAATCATCTT
-270 GATTTTCGGGCCGGGGGCTTGCATGGAACCCGCACGTTGGCGGGGTCG
-222 AGAGAGATGATGCCTCTTCAGTTTTACGGCGAGACATCATGTCTCTTC
-174 AGTCTTACGACGGGATCATCTGGCGAGACGATTTGGGAGTATAAATAT
-126 ATCCATCGGATCCCCTTGTCCCAGTCAATCTACAACCATCAGCTCTCC
- 78 CATCATCCAACAAACATCTCACTTCAAACAAACATACTTTCAAATTTC
- 30 AGCTCTCAGCTCTCAACTCTCATCATCAAAATGAAGTCCTTTGCCATC
                                      M  K  S  F  A  I

  19 ACTGCCGTTCTCTTTGCTGCCGCTGCCATGGCTGGCCCTCTCGAGGTC
      T  A  V  L  F  A  A  A  A  M  A  G  P  L  E  V

  67 CGGACTGGTGGCGGTGGTGGCGGCGGCGGCGGTGGTGGTTCTGCCTAC
      R  T  G  G  G  G  G  G  G  G  G  G  G  G  S  A  Y

 115 GACGCCTGCCCCAGCGAGCTCCTCAGCAACCCTGAGTGCTGCGATGTC
      D  A  C  P  S  E  L  L  S  N  P  E  C  C  D  V
            *                              *  *

 163 GATGTCCTCGGGCTCCTCTCGATCAACTGCGCGCCCCgtgagttttttc
      D  V  L  G  L  L  S  I  N  C  A  P
                                    *

 211 aaagcagctgctcttgcatatgtccggctgacctttttcattacagCCG
                                                     P

 259 ACAACACTCCTACTTCTGCCGCGGACTTCAAGAGTCAGTGCGCTAAGC
      D  N  T  P  T  S  A  A  D  F  K  S  Q  C  A  K
                                                *

 307 GTGGAAAGTCTGCCACTTGCTGCGTCCTCCCTGTTgtaagatatcctg
      R  G  K  S  A  T  C  C  V  L  P  V
                            *  *

 355 ctgctctaactgcacggttcccggtactaatcacgactcttcaatgtg

 403 tagCTTGGCCAGGCCGTCGGCTGCCAGGCCCCTGCTGGCTCTGGCTAA
        L  G  Q  A  V  G  C  Q  A  P  A  G  S  G
                          *

 451 GTGCCGGGTTACTCTCGGGCCATTTGATTCACGAGGGCCAGTCCTTTT
 499 GTTTACCCGTGTGTATTTGCTTGCTGTTGATCTTTTACGCCGTGGACT
 547 TCAATGGGAAGTGTTTGGGAGCCGTCTTACACCGCCTACATGTACATG
 595 GCTGTGGCCATAACCGACAAGGACGCTTTCATCCTCCCCTGCAAACTG
 643 TCGTTTCTGTTCATAGTCCAATCGCAAATGCAACCATGGCGCTATAGC
 691 CGAACTCC
```

tein. Results (Fig. 2) showed that the mass of the mature protein was 8111.75 Da. Prior to sequencing, the GEO1 protein was subjected to carboamidomethylation and trypsin hydrolysis, followed by RP–HPLC purification of the tryptic peptides. Edman sequencing of the entire carboamidomethylated protein determined an amino acid sequence extremely rich in glycine, which was followed by the SAYDACP sequence. One of the most abundant tryptic peptides was also sequenced and showed the sequence ATCCVLPVL.

Cloning and sequence analysis of the *geo1* gene

To isolate the gene encoding the GEO1 protein, degenerate primers were designed on the 2 known regions of the protein according to the codon usage of *Aspergillus fumigatus*, a filamentous ascomycete whose genome has been completely sequenced (http://www.kazusa.or.jp/codon/cgi-bin/showcodon.cgi?species=5085). The sequences of the primers were as follows: 5′-TC(C/G)GC(C/A)TACGACGC(C/T)TGCCCC-3′

Fig. 4. MUSCLE alignment of the *Geosmithia pallida* GEO1 protein with the best BLAST hits. Sequences were aligned starting from the first cysteine residue to remove the signal peptide. Conserved amino acids are highlighted in grey and the 8 cysteine residues are indicated by asterisks. The sequences used are as follows: *Verticillium dahliae* VDAG_07851 (EGY16687) and VDAG_01586 (EGY17904); *Ophiostoma ulmi* cerato–ulmin (CU) (Q06153); *Ophiostoma novo-ulmi* CU (CAD58391); *Trichoderma atroviride* ABS59366, ABS59365, ABS59367, ABS59371, and EHK49783; *Metarhizium anisopliae* MAA_01182 (EFZ04108); *Trichoderma reesei* hfb1 (P52754); *Verticillium albo-atrum* VDBG_05155 (XP_003004042); *Glomerella graminicola* GLRG_11190 (EFQ36046) and GLRG_10074 (EFQ34930); *Fusarium oxysporum* FOXB_09102 (EGU80400); *Claviceps purpurea* cpph1 (CAD10781); *Hypocrea virens* ABS59377; *Cryphonectria parasitica* CRP (P52753); *Claviceps fusiformis* TH1 (Q9UVI4); *Magnaporthe oryzae* MGG_10105 (EHA51087). UniProt or GenBank accession numbers are in parentheses. The sequences from *T. atroviride* and *H. virens* are identified by their accession numbers, as gene names were not available.

```
GEO1                    CPSE--LLSNPECCDVDVLGLLSINCAPPDNTPTSAADFKSQC-AKRGKSATCCVLPVLGQAVGCQAPAGSG-
V.dahliae_VDAG_07851    CPSG--LSMTPVCCATNVLGLLALNCDAPSKTPTSAKDFQKIC-ADAGTAAKCCTLNLLNQGVLCQVPVGVAA
O.ulmi_CU               CTGL--LQKSPQCCNTDILGVANLDCHGPPSVPTSPSQFQASCVADGGRSARCCTLSLLGLALVCTDPVGI--
O.novo-ulmi_CU          CTGL--LQKSPQCCNTDILGVANLDCHGPPSVPTSPSQFQASCVADGGRSARCCTLSLLGLALVCTDPVGI--
T.atroviride_ABS59366   CPPGL-LYTNPQCCDVDVLGVADLNCVVPPRGPSNCKTFSGIC-ASIGKEPKCCAVPILGQALLCTDPVGSRN
M.anisopliae_MAA_01182  CPSG--LYSNPQCCSTVIIGVIALDCKSPSEVPRDGTDLRNIC-AKTGQKAACCVIPVAGQALLCETAPGA--
T.atroviride_ABS59365   CPNG--LYSNPQCCDTLVLGIVGLGCEVPTQTPRDGVDFKNIC-AKTGDQAVCCVVPVAGQDLLCQTAVGIQG
T.reesei_hfb1           CPPG--LFSNPQCCATQVLGLIGLDCKVPSQNVYDGTDFRNVC-AKTGAQPLCCVAPVAGQALLCQTAVGA--
T.atroviride_ABS59367   CPAGL-LYSNPQCCSVDVLGIADLDCSSPSKTPKDASDFQDIC-DTAGKKAQCCAVPIAGQAVLCQDVIG---
T.atroviride_ABS59371   CSGGGGLYSNLQCCSVDVLGIADLDCPPSKTPVSGSDFKSIC-QASGAKPKCCVLPILGQGLLCEDAIGTA--
V.albo-atrum_VDBG_05155 CPSGL-LNGSPQCCATDVLGVASLDCSVPSRTPSDGADLKAIC-AEAGKSALCCSVPVAGQSLLCTAVIG---
T.atroviride_EHK49783   CAAG--LYSTPQCCAVDVLGVADLNCAGLVGTVTTAAEFKANC-AAIGQEARCCVLPVLGQDVLCQTPPGL--
G.graminicola_GLRG_11190 CSG---LYGSAQCCATDILGIANLDCANPYEVPTTADNFSSIC-SEIGQRARCCILPILDQGILCNTPAGVQD
F.oxysporum_FOXB_09102  CPDS--LYGNPQCCSVDVLGVADVECDSPTESPTDAENFQAIC-AASGQRARCCVLPVLGQALVCLTPVGVSN
V.dahliae_VDAG_01586    CSG---LYGNPQCCATDVLGVASLDCQNP-KSANNANDFKQSC-ASTGKSAFCCTLPVAGQAVLCNKPVGV--
C.purpurea_cpph1        CPSG--LYSNPQCCSTDVLGVASLDCVNPSSKPRSGDNFKSTC-AAKGKHASCCVIPVAGQGVLCQPAM----
G.graminicola_GLRG_10074 CSG---LYSSVQCCATSVLDLADLTCRPPPKVPTSAANFGKIC-ADIGQRARCCVLPALGLGVLCQTPAGVTY
H.virens_ABS59377       CPGG--LESNPQCCSTDVLGIADLDCANPSSPVTDVQSFRAVC-AAGGQRARCCAIPVAGQALLCESPVGI--
C.parasitica_CRP        CSST--LYSEAQCCATDVLGVADLDCETVPETPTSASSFESIC-ATSGRDAKCCTIPLLGQALLCQDPVGL--
C.fusiformis_TH1        CPSG--LYSVPQCCATDVLGVADLDCGNPSRQPTDSSDFASVC-AAKGQRARCCVLPLLGQAVLC---TGA--
M.oryzae_MGG_10105      CEGNGLLYSSAQCCATDVLGVADLDCAVPPSLPTSASGFTDIC-AALGQRARCCVLPLAGQAVLCQAPVGA--
                        *         **         *                    *         **         *
```

(forward) and 5′-GAGGACGGGGAGGACGCAGCA(G/A)GT-3′ (reverse). Amplification of *G. pallida* isolate IVV7 genomic DNA generated a single band of 236 bp, which was sequenced. The complete sequence of the *geo1* gene, including 462 bp upstream and 248 bp downstream of the coding region, was obtained by means of a PCR-based genome-walking strategy. Primers designed on the putative open reading frame were then used for the RT–PCR amplification of the corresponding cDNA, whose sequence was determined and compared with the genomic sequence to detect the presence of introns. Primers used to this aim were 5′-AAAT-GAAGTCCTTTGCCATCA-3′ (forward) and 5′-GAGAG-TAACCCGGCACTTAGC-3′ (reverse). The predicted open reading frame consisted of 450 bp and was interrupted by 2 introns of 56 and 64 bp, respectively (Fig. 3). A putative TATA box at position –136 and 2 putative CCAAT boxes at positions –72 and –394 were identified in the 462 bp region upstream of the ATG codon. A putative stress response element (Marchler et al. 1993) at position –115 could also be involved in the regulation of the *geo1* gene. No polyadenylation signal(s) were detected in the 248 bp region downstream to the stop codon. The predicted primary translation product (Fig. 3) was 109 amino acids long and had a putative signal peptide of 23 amino acids. Database searches (BLASTp) with the deduced GEO1 amino acid sequence showed significant homologies to proteins of the class II hydrophobins family (Pfam PF06766). The alignment of GEO1 with the best hits from the BLAST search showed the presence of 8 cysteine residues in conserved positions (Fig. 4), whose spacing (X_{40}-C-X_9-CC-X_{11}-C-X_{16}-C-X_8-CC-X_{10}-C-X_7, where C indicates cysteine and X any other amino acid) was in accordance to the class II hydrophobin consensus $X_{17–67}$-C-$X_{9–10}$-CC-X_{11}-C-X_{16}-C-$X_{6–9}$-CC-X_{10}-C-$X_{3–7}$ (Wösten and de Vocht 2000).

The theoretical molecular mass of the mature protein was calculated with the Compute pI/MW tool (ExPaSy Proteomics Server). Taking into account the 8 hydrogen atoms in the cysteine residues involved in the formation of 4 disulfide bridges, the estimated molecular mass was 8109.04 Da and, therefore, consistent with the MALDI-TOF result of 8111.75. The theoretical isoelectric point was 4.33. Hydrophobins are characterized by a series of biochemical properties such as lowering of surface tension of water, efficiency in foam production and stabilization, and formation of surface membranes (Linder 2009). In particular, the membranes formed by class II hydrophobins are readily soluble in ethanol and in sodium dodecyl sulfate, while those of class I can be solubilized only in TFA or formic acid (Wösten 2001; Linder et al. 2005). To demonstrate that GEO1 shared the properties of class II hydrophobins, the appearance of a protein solution after agitation, its foaming capability, and the dissolution of protein aggregates thus formed by ethanol were tested. Results showed that an aqueous solution of GEO1 (200 μg/mL) became milky after shaking (Figs. 5*a* and 5*b*) and that a dense foam developed when the solution was agitated 5 min by vortexing (Fig. 5*e*), which was comparable with that obtained with a solution of the class II hydrophobin cerato–ulmin at the same concentration (Fig. 5*f*). Moreover, the foam thus obtained was stable for at least 96 h at room temperature (data not shown). When the protein solution was poured in a Petri dish, a film formed at the air–water interface that was immediately dissolved upon addition of ethanol (Figs. 5*g*–5*i*). Finally, as hydrophobin solutions on a hydrophobic surface take a trapezoid-like profile with a membrane developing on top (Szilvay et al. 2007), the shape change of a 40 μL drop of a GEO1 aqueous solution on a polytetra-fluoroethylene surface was followed over 1 h in ambient conditions and compared with the shape of a drop of distilled

Fig. 5. Biochemical properties of GEO1. Appearance of a GEO1 aqueous solution (200 μg/mL) before (*a*) and after (*b*) vortexing; profile of a 40 μL drop of distilled water (*c*) and of a GEO1 solution (*d*) on a hydrophobic polytetrafluoroethylene surface after 30 min in ambient conditions; foaming capability of GEO1 (*e*) and cerato–ulmin (*f*) solutions (200 μg/mL) after 5 min vortexing; formation of GEO1 aggregates at the water–air interface (*g*); and dissolution of the film formed upon ethanol addition (*h* and *i*).
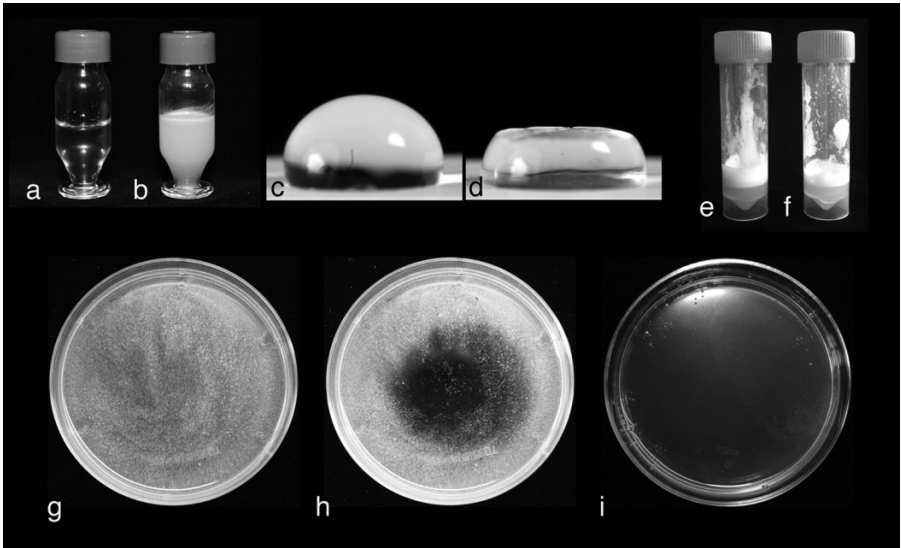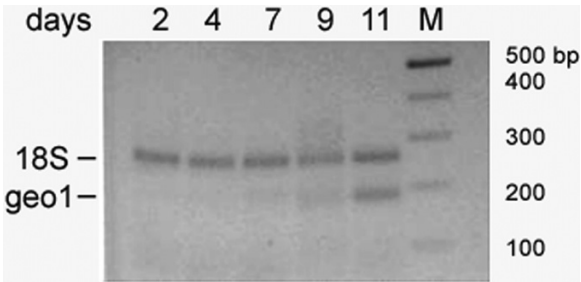

Fig. 6. Southern blot analysis of the *geo1* gene. Lanes: 1, Genomic DNA from *Geosmithia pallida* isolate IVV7 double-digested with *Eco*RI and *Hin*dIII and hybridized with an 803 bp *geo1* probe; 2, DNA molecular weight marker II DIG-labeled (Roche Applied Science).

Fig. 7. Time-course analysis by semiquantitative RT–PCR of the *geo1* gene transcription after 2, 4, 7, 9, and 11 days of growth of *Geosmithia pallida* isolate IVV7 in liquid medium. The *18S* ribosomal gene was used as the endogenous control. M, molecular weight marker (GeneRuler DNA Ladder Mix, Fermentas Life Sciences).
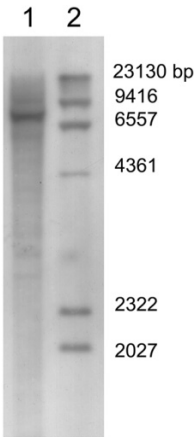
water. After 30 min the drop showed a trapezoid-like shape (Fig. 5*d*) and a film could be seen on the flat surface, while the drop of distilled water maintained a round shape (Fig. 5*c*).

The *geo1* copy number in the fungal genome was determined by Southern hybridization using as a probe an 803 bp fragment of the gene labeled with digoxigenin-dUTP. A single band of 7.3 kb was obtained (Fig. 6), indicating that *geo1* is a single-copy gene. However, due to the low homology between hydrophobins, we cannot exclude that *G. pallida* could possess still other, unidentified, hydrophobin genes as shown for other species (Fuchs et al. 2004; Askolin et al. 2005; Kubicek et al. 2008). Finally, semiquantitative RT–PCR was used for the determination of *geo1* transcript levels after 2, 4, 7, 9, and 11 days of growth in liquid Takai medium. The gene for the 18S ribosomal RNA was chosen as the constitu-

tive control. Time-course of *geo1* expression is shown in Fig. 7. The *geo1* transcript appeared after 7 days of culture and increased up to 11 days, while the level of the 18S transcript remained constant throughout the entire experiment. The fungi of the genus *Geosmithia* are entomochoric, i.e., they depend strictly on insect vectors for dispersal; however, they do not possess sticky conidia as other entomochoric species, such as the Ophiostomas. Conidia produced by *Geosmithia* spp. are instead dry and hydrophobic such as those produced by airborne fungi (Kolařík et al. 2008). As hydrophobins can mediate the attachment of fungi to hydrophobic surfaces, a possible role for GEO1 could be to favour the dissemination of the fungus by virtue of the hydrophobicity conferred to the conidia, which would allow establishment of hydrophobic interactions between the chitinous exoskeleton of the insect vectors and the conidia themselves (Wösten 2001; Temple and Horgen 2000). This is the case for entomopathogenic fungi, such as *Beauveria bassiana* (Zhang et al. 2011), where the adhesion of conidia to the host's surface is mediated by nonspecific hydrophobic and electrostatic interactions involving hydrophobin rodlet layers on the conidial

cell wall. Further work is underway in our laboratory to analyze all the species of the genus *Geosmithia* known so far for the presence of GEO1, with the aims of (*i*) verifying if species with different geographic origin and (or) lifestyle could have different forms of the protein and (*ii*) comparing the phylogenetic relationships thus obtained between the species with those already known for the genus.

Acknowledgements

References

Aimanianda, V., Bayry, J., Bozza, S., Kniemeyer, O., Perruccio, K., Elluru, S.R., et al. 2009. Surface hydrophobin prevents immune recognition of airborne fungal spores. Nature, **460**(7259): 1117–1121. doi:10.1038/nature08264. PMID:19713928.

Altschul, S.F., Gish, W., Miller, W., Myers, E.W., and Lipman, D.J. 1990. Basic local alignment search tool. J. Mol. Biol. **215**(3): 403–410. doi:10.1016/S0022-2836(05)80360-2. PMID:2231712.

Askolin, S., Penttilä, M., Wösten, H.A.B., and Nakari-Setälä, T. 2005. The *Trichoderma reesei* hydrophobin genes *hfb1* and *hfb2* have diverse functions in fungal development. FEMS Microbiol. Lett. **253**(2): 281–288. doi:10.1016/j.femsle.2005.09.047. PMID:16243453.

Dyer, P.S. 2002. Hydrophobins in the lichen symbiosis. New Phytol. **154**(1): 1–4. doi:10.1046/j.1469-8137.2002.00387.x.

Edgar, R.C. 2004. MUSCLE: multiple sequence alignment with high accuracy and high throughput. Nucleic Acids Res. **32**(5): 1792–1797. doi:10.1093/nar/gkh340. PMID:15034147.

Fuchs, U., Czymmek, K.J., and Sweigard, J.A. 2004. Five hydrophobin genes in *Fusarium verticillioides* include two required for microconidial chain formation. Fungal Genet. Biol. **41**(9): 852–864. doi:10.1016/j.fgb.2004.04.004. PMID:15288021.

Gasteiger, E., Gattiker, A., Hoogland, C., Ivanyi, I., Appel, R.D., and Bairoch, A. 2003. ExPASy: the proteomics server for in-depth protein knowledge and analysis. Nucleic Acids Res. **31**(13): 3784–3788. doi:10.1093/nar/gkg563. PMID:12824418.

Izumitsu, K., Kimura, S., Kobayashi, H., Morita, A., Saitoh, Y., and Tanaka, C. 2010. Class I hydrophobin BcHpb1 is important for adhesion but not for later infection of *Botrytis cinerea*. J. Gen. Plant Pathol. **76**(4): 254–260. doi:10.1007/s10327-010-0247-0.

Karlsson, M., Stenlid, J., and Olson, A. 2007. Two hydrophobin genes from the conifer pathogen *Heterobasidion annosum* are expressed in aerial hyphae. Mycologia, **99**(2): 227–231. doi:10.3852/mycologia.99.2.227. PMID:17682775.

Kim, S., Ahn, I.-P., Rho, H.-S., and Lee, Y.-H. 2005. *MHP1*, a *Magnaporthe grisea* hydrophobin gene, is required for fungal development and plant colonization. Mol. Microbiol. **57**(5): 1224–1237. doi:10.1111/j.1365-2958.2005.04750.x. PMID:16101997.

Klimes, A., Amyotte, S.G., Grant, S., Kang, S., and Dobinson, K.F. 2008. Microsclerotia development in *Verticillium dahliae*: regulation and differential expression of the hydrophobin gene *VDH1*. Fungal Genet. Biol. **45**(12): 1525–1532. doi:10.1016/j.fgb.2008.09.014. PMID:18951989.

Kolařík, M., Kostovčík, M., and Pažoutová, S. 2007. Host range and diversity of the genus *Geosmithia* (*Ascomycota*: *Hypocreales*) living in association with bark beetles in the Mediterranean area. Mycol. Res. **111**(11): 1298–1310. doi:10.1016/j.mycres.2007.06.010. PMID:18024099.

Kolařík, M., Kubátová, A., Hulcr, J., and Pažoutová, S. 2008. *Geosmithia* fungi are highly diverse and consistent bark beetle associates: evidence from their community structure in temperate Europe. Microb. Ecol. **55**(1): 65–80. doi:10.1007/s00248-007-9251-0. PMID:17940820.

Kolařík, M., Freeland, M., Utley, C., and Tisserat, N. 2011. *Geosmithia morbida* sp. nov., a new phytopathogenic species living in symbiosis with the walnut twig beetle (*Pityophthorus juglandis*) on *Juglans* in USA. Mycologia, **103**(2): 325–332. doi:10.3852/10-124. PMID:20943528.

Kubicek, C.P., Baker, S.E., Gamauf, C., Kenerley, C.M., and Druzhinina, I.S. 2008. Purifying selection and birth-and-death evolution in the class II hydrophobin gene families of the ascomycete *Trichoderma/Hypocrea*. BMC Evol. Biol. **8**(1): 4. doi:10.1186/1471-2148-8-4. PMID:18186925.

Linder, M.B. 2009. Hydrophobins: protein that self assemble at interfaces. Curr. Opin. Colloid Interface Sci. **14**(5): 356–363. doi:10.1016/j.cocis.2009.04.001.

Linder, M.B., Szilvay, G.R., Nakari-Setälä, T., and Penttilä, M.E. 2005. Hydrophobins: the protein-amphiphiles of filamentous fungi. FEMS Microbiol. Rev. **29**(5): 877–896. doi:10.1016/j.femsre.2005.01.004. PMID:16219510.

Mankel, A., Krause, K., and Kothe, E. 2002. Identification of a hydrophobin gene that is developmentally regulated in the ectomycorrhizal fungus *Tricholoma terreum*. Appl. Environ. Microbiol. **68**(3): 1408–1413. doi:10.1128/AEM.68.3.1408-1413.2002. PMID:11872494.

Marchler, G., Schuller, C., Adam, G., and Ruis, H. 1993. A *Saccharomyces cerevisiae* UAS element controlled by protein kinase A activates transcription in response to a variety of stress conditions. EMBO J. **12**(5): 1997–2003. PMID:8387917.

Ng, W.L., Ng, T.P., and Kwan, H.S. 2000. Cloning and characterization of two hydrophobin genes differentially expressed during fruit body development in *Lentinula edodes*. FEMS Microbiol. Lett. **185**(2): 139–145. doi:10.1111/j.1574-6968.2000.tb09052.x. PMID:10754238.

Scala, A., Comparini, C., Tegli, S., and Scala, F. 2007. A non-*Ophiostoma* fungus expresses the gene encoding the hydrophobin cerato–ulmin. J. Plant Pathol. **89**(2): 233–240. doi:10.4454/jpp.v89i2.748.

Seidl-Seiboth, V., Gruber, S., Sezerman, U., Schwecke, T., Albayrak, A., Neuhof, T., et al. 2011. Novel hydrophobins from *Trichoderma* define a new hydrophobin subclass: protein properties, evolution, regulation and processing. J. Mol. Evol. **72**(4): 339–351. doi:10.1007/s00239-011-9438-3. PMID:21424760.

Sunde, M., Kwan, A.H.Y., Templeton, M.D., Beever, R.E., and Mackay, J.P. 2008. Structural analysis of hydrophobins. Micron, **39**(7): 773–784. doi:10.1016/j.micron.2007.08.003. PMID:17875392.

Szilvay, G.R., Paananen, A., Laurikainen, K., Vuorimaa, E., Lemmetyinen, H., Peltonen, J., and Linder, M.B. 2007. Self-assembled hydrophobin protein films at the air–water interface: structural analysis and molecular engineering. Biochemistry, **46**(9): 2345–2354. doi:10.1021/bi602358h. PMID:17297923.

Talbot, N.J. 2003. On the trail of a cereal killer: exploring the biology of *Magnaporthe grisea*. Annu. Rev. Microbiol. **57**(1): 177–202. doi:10.1146/annurev.micro.57.030502.090957. PMID:14527276.

Temple, B., and Horgen, P.A. 2000. Biological roles for cerato–ulmin, a hydrophobin secreted by the elm pathogens, *Ophiostoma*

ulmi and *O. novo-ulmi.* Mycologia, **92**(1): 1–9. doi:10.2307/3761443.

van Wetter, M.A., Wösten, H.A.B., Sietsma, J.H., and Wessels, J.G. H. 2000. Hydrophobin gene expression affects hyphal wall composition in *Schizophyllum commune.* Fungal Genet. Biol. **31**(2): 99–104. doi:10.1006/fgbi.2000.1231. PMID:11170739.

Wessels, J.G.H. 2000. Hydrophobins, unique fungal proteins. Mycologist, **14**(4): 153–159. doi:10.1016/S0269-915X(00)80030-0.

Whiteford, J.R., and Spanu, P.D. 2002. Hydrophobins and the interactions between fungi and plants. Mol. Plant Pathol. **3**(5): 391–400. doi:10.1046/j.1364-3703.2002.00129.x. PMID: 20569345.

Wösten, H.A.B. 2001. Hydrophobins: multipurpose proteins. Annu. Rev. Microbiol. **55**(1): 625–646. doi:10.1146/annurev.micro.55.1.625. PMID:11544369.

Wösten, H.A.B., and de Vocht, M.L. 2000. Hydrophobins, the fungal coat unravelled. Biochim. Biophys. Acta, **1469**(2): 79–86. doi:10.1016/S0304-4157(00)00002-2. PMID:10998570.

Zhang, S., Xia, Y.X., Kim, B., and Keyhani, N.O. 2011. Two hydrophobins are involved in fungal spore rodlet layer assembly and each play distinct roles in surface interactions, development and pathogenesis in the entomopathogenic fungus, *Beauveria bassiana.* Mol. Microbiol. **80**(3): 811–826. doi:10.1111/j.1365-2958.2011.07613.x. PMID:21375591.