The *Sc7/Sc14* gene family of *Schizophyllum commune* codes for extracellular proteins specifically expressed during fruit-body formation

FRANK H. J. SCHUREN,* SIGRIDUR A. ÁSGEIRSDÓTTIR, ERIKA M. KOTHE,† JOSÉ M. J. SCHEER and JOSEPH G. H. WESSELS

Department of Plant Biology, Biological Centre, University of Groningen, Kerklaan 30, 9751 NN Haren, The Netherlands

(Received 1 April 1993; revised 13 May; accepted 20 May 1993)

The *Sc7* and *Sc14* genes are specifically expressed in the dikaryon of the basidiomycete fungus *Schizophyllum commune* during fruiting. These genes are closely linked (within 6 kb) and highly similar in gene structure and nucleotide sequence (70% identical nucleotides in their coding regions). The encoded proteins (204 and 214 amino acids, respectively) have 87% similarity in amino acids (56% of the amino acids are identical). They contain putative signal sequences for secretion, are rich in aromatic amino acids which are generally located at similar positions, and they are generally hydrophilic. Inspection of databanks showed similarities with pathogenesis-related proteins (PR1) from plants, testis-specific proteins from mammals and venom allergen proteins from insects. An antibody raised against a *Sc7* fusion protein showed the presence of the *Sc7* protein in the culture medium and in the fruit bodies where it is apparently loosely associated with hyphal walls.

Introduction

Fruit-body formation is the most conspicuous developmental process in higher fungi. In the basidiomycete *Schizophyllum commune* the genetic system controlling dikaryon and fruit-body formation is known in some detail (Wessels, 1992). The primary control is by interactions between mating-type genes, some of which have recently been cloned and sequenced (Stankis *et al.*, 1992; Specht *et al.*, 1992). Furthermore, two other genes that regulate the fruiting process, *FBF* and *THN*, have been identified (Springer & Wessels, 1989; Wessels *et al.*, 1991*a*).

Regarding genes that are possible targets for regulation in this system, eight cDNA clones for mRNAs specifically and abundantly produced in the dikaryon during fruit-body formation were isolated (Mulder & Wessels, 1986). It was shown that the expression of the

corresponding genes is under control of the mating-type genes (Ruiters *et al.*, 1988), the *FBF* gene (Springer & Wessels, 1989) and the *THN* gene (Wessels *et al.*, 1991*a*). Two of these regulated genes, *Sc1* and *Sc4*, are expressed at particularly high levels and belong to a family of genes coding for hydrophobins, small cysteine-rich hydrophobic proteins (Schuren & Wessels, 1990). These proteins are excreted into the cell walls of fruit-body hyphae where they form high molecular mass insoluble complexes (Wessels *et al.*, 1991*b*). Another member of this gene family, *Sc3* (Schuren & Wessels, 1990), is regulated by the *THN* gene only and the encoded hydrophobin forms insoluble complexes in the walls of aerial hyphae (Wessels *et al.*, 1991*b*).

Here we report the nucleotide and amino acid sequences of two other genes, *Sc7* and *Sc14*, which are abundantly expressed in the dikaryon during fruit-body formation (Mulder & Wessels, 1986; Ruiters *et al.*, 1988). By using an antibody against a bacterial fusion protein, we show that the *Sc7* protein is excreted into the medium and loosely bound to fruit-body hyphae.

Methods

Strains, culture condition and labelling. Surface cultures of *S. commune* monokaryon 4-40 (*A43 B43*, CBS 340.81) and a dikaryon, obtained by crossing monokaryon 4-40 with the co-isogenic mono-

* Author for correspondence. Tel. 50 632322; fax 50 632273.

† Present address: Department of Molecular Genetics, University of Marburg, Karl von Frisch Strasse, W 3550 Marburg, Germany.

The nucleotide sequence data reported in this paper have been submitted to the GenBank/EMBL/DBJ databases and have been assigned the accession numbers M81722 (*Sc7*) and M81723 (*Sc14*).

karyon 4-39 (*A41 B41*, CBS 341.81), were grown at 24 °C in continuous light for 4 d as described by Wessels *et al.* (1991*b*). For labelling experiments, the mycelial mats were transferred after 3 d of growth to minimal medium (Dons *et al.*, 1979) containing either 565 kBq [¹⁴C]lysine ml⁻¹ (specific activity > 11 GBq ml⁻¹) or 565 kBq [¹⁴C]tyrosine ml⁻¹ (specific activity > 11 GBq ml⁻¹) and cultivation was continued for 24 h. Alternatively, labelling was done from the beginning of cultivation with 1.89 MBq [³⁵S]sulphate ml⁻¹ (specific activity 1000 to 1500 GBq ml⁻¹; MgCl₂ replacing MgSO₄ in the medium).

Isolation of genomic clones. Isolation of the *Sc7* and the *Sc14* cDNA clones was described by Mulder & Wessels (1986). The genomic clone for *Sc7* was isolated from a genomic library of *S. commune* strain 4-39 in λ EMBL4 (Schuren & Wessels, 1990). A 1.8 kb *Bam*HI fragment and a 0.5 kb *Bam*HI–*Sal*I fragment, together encompassing the entire *Sc7* gene and flanking sequences, were fused in plasmid p7BBS. A genomic clone for *Sc14* was isolated from a cosmid library of the coisogenic strain 4-40 (Giasson *et al.*, 1989). A 9.5 kb *Nco*I cosmid fragment containing the *Sc14* gene was subcloned. This clone (p147Nc) turned out to contain both the *Sc14* and the *Sc7* genes.

Sequencing strategy. Subclones of both the *Sc7* and the *Sc14* cDNA and genomic clones were constructed in pUC vectors. By sequencing these clones, independently from both strands, the nucleotide sequences of both genes were determined. Sequencing was performed with the T7 DNA polymerase kit from Pharmacia using the dideoxy chain termination method (Sanger *et al.*, 1977). Transcript termini were determined by S1-nuclease experiments (Schuren & Wessels, 1990). Nucleotide and amino acid sequences were compared with the EMBL (release 34.0) and SWPROT (release 24.0) data libraries using the FASTN and FASTP programs (Lipman & Pearson, 1985). Multiple alignments were performed with the CLUSTAL program (Higgins & Sharp, 1988).

Production of *Sc7* fusion-protein in *E. coli*. An *A*luI restriction fragment of the *Sc7* cDNA clone, containing the entire coding sequence of *Sc7*, except for the first two amino acids, plus 17 nucleotides downstream of the stop codon, was subcloned into the *Sma*I site of pUC18. This insert was cloned as a *Eco*RI–*Sal*I fragment into a pRIT2T vector (Pharmacia), resulting in an in-frame fusion with the Protein A gene. An overnight culture of *E. coli* N4830-1 cells containing the *Sc7* fusion gene construct was diluted 100 times to a final volume of 15 ml LB medium (Sambrook *et al.*, 1989) containing 50 μ l ampicillin ml⁻¹ and grown at 30 °C to early stationary phase (OD₆₀₀ = 1). The thermo-inducible λ *Pr* promoter was switched on by adding an equal amount of medium preheated to 54 °C and incubation continued at 42 °C for 90 min. The cells were then lysed by sonication in 3 ml ice-cold TST buffer (50 mM-Tris/HCl, pH 7.6, 150 mM-NaCl, 0.05% v/v, Tween 20) and the clear supernatant was applied to an IgG Sepharose-6 Fast Flow column (Pharmacia). The fusion protein was eluted from the column with 0.1 M-glycine/HCl buffer, pH 3; 1 ml fractions were collected and immediately neutralized with 0.1 M-Tris/HCl, pH 9.6. The presence of Protein A/*Sc7* fusion protein was detected by Western blotting using horseradish peroxidase anti-peroxidase antibodies and peroxidase activity was visualized with 4-chloro-1-naphthol according to Harlow & Lane (1988).

Antisera. Rabbits were injected five times at 10 to 14 d intervals with the *Sc7* fusion protein or with purified Sc3 hydrophobin isolated from the medium of a *S. commune* monokaryon (this laboratory, unpublished). Blood was collected before the first immunization and 14 d after the last injection.

Isolation of proteins and immunodetection. Hyphae were fragmented in an X-press (Biotec) and proteins were extracted by boiling fragmented hyphae and isolated hyphal walls (Wessels *et al.*, 1991*a*) in

SDS sample buffer (2%, w/v, SDS, 0.05 M-Tris/HCl, pH 6.8, 10%, w/v, glycerol, 0.001%, w/v, bromophenol blue) for 10 min. Extracellular proteins present within fruit bodies were collected by first infiltrating isolated fruit bodies with 50 mM-Tris/HCl, pH 7.5, for 1 h under reduced pressure and recovering the infiltrated fluid by centrifugation. The fruit bodies were then ground with a pestle and mortar and the remaining proteins were extracted with hot SDS sample buffer as described above. Proteins excreted into the medium were concentrated by precipitation with 10% (w/v) trichloroacetic acid (TCA) and solubilized with concentrated trifluoroacetic acid (TFA), before dissolution in SDS sample buffer (De Vries *et al.*, 1993). Proteins were separated on 12.5% PAGE gels (Laemmli, 1970) and transferred to polyvinylidene difluoride (PVDF) membranes (Immobilon-P, Millipore). The blots were blocked with 5% (w/v) skimmed milk proteins and 1% (w/v) polyvinylpyrrolidone (PVP) in phosphate-buffered saline (PBS, pH 7.6; per litre 8 g NaCl, 0.2 g KCl, 1.45 g Na₂HPO₄·2H₂O, 0.23 g NaH₂PO₄·H₂O). Incubation with antiserum and secondary antibodies (GAR-AP, Boehringer) was carried out in blocking buffer, for 1 h each, at room temperature and the blots were washed with PBS between incubation steps. Immunoreaction was visualized by the alkaline phosphatase reaction, using BCIP and NBT as substrates (Harlow & Lane, 1988).

Results and Discussion

Nucleotide sequences of the *Sc7* and *Sc14* genes

The complete nucleotide sequences of *Sc7* and *Sc14* and the predicted amino acid sequences are shown in Figs 1 and 2. The two genes are structurally similar, closely linked (within 6 kb) and divergently transcribed (Fig. 3). At the DNA level the homology in the coding parts of *Sc7* and *Sc14* is 70%. The GC-content in the coding regions of both genes is quite high (64% for both genes), but is much lower in the 5'- and 3'-noncoding regions (55% for *Sc7*, 54% for *Sc14*). A TATA-box is present in the promoter region of both genes around 30 nucleotides upstream of the transcription start point. In the *Sc7* promoter region the sequence CGAAT is present starting at position –85, which is quite similar to the CCAAT box found in many genes of eukaryotes between nucleotides –70 and –90 (Benoit *et al.*, 1980). The sequences around the ATG initiation codons of both genes fit the consensus of Kozak (1984) with an A at position –3.

The intron positions were determined from a comparison of the cDNA and genomic sequences. Both the *Sc7* and the *Sc14* genes contain five introns, four of which are at exactly the same positions with respect to the amino acid sequences (Fig. 3). The first intron of both genes is rather large (134 bp for *Sc7*, 103 bp for *Sc14*), whereas the other introns are smaller (49 to 76 bp) and more similar in size to those found in the *Sc1*, *Sc3* and *Sc4* genes of *S. commune* (Schuren & Wessels, 1990) and of fungal genes in general (Gurr *et al.*, 1987). Intron 5 of *Sc14* is located in the 3'-noncoding region of the gene. All introns start with GT and end with AG, but

```

-125  ccctgcgagg tcgtgctggg catcgcaagc gagcaaggtt cgaatcagcc ttactgcac ctgaacgcac gtaaacttct ctcacccgt
                                     1
-35  aaggtataaa ggccgtgcgc tcggcgagtc aagagcaagc acttctgctt ctcttcagtc ccttgccttt caccgacccc gacaga ATG
                                     →
                                     10                                20
46  Lys Leu Thr Val Ile Leu Leu Thr Ala Val Leu Ala Ala Ser Ala Ser Pro Ala Pro Val Asp Val Asp Ala Arg
AAG CTC ACC GTC ATC CTC CTC ACC GCC GTC CTC GCT GCA AGC GCT AGC CCT GCG CCT GTC GAT GTC GAT GCT CGC

                                     30                                40                                50
121 Ala Pro Val Ala Leu Asp Ser Arg Ser Pro Ile Asp Ile Asp Ser Arg Ser Ala Asp Ala Leu Ala Asn Arg Ala
GCG CCG GTC GCC CTC GAC TCT CGC TCG CCG ATT GAC ATC GAC TCC CGC TCC GCC GAC GCC CTC GCC AAT CGC GCG

                                     60                                70
196 Ala Pro Pro Gln Ser Glu Ile Asp Gln Trp Leu Lys Ala His Asn Asn Glu Arg Ala Gln His Gly Ala Val Ala
GCC CCG CCC CAG TCC GAG ATC GAC CAG TGG CTC AAG GCC CAC AAC AAC GAG CGT GCG CAG CAC GGC GCG GTT GCG

                                     80                                90                                100
271 Leu Val Trp Asn Gln Thr Leu Ser Asp Lys Ala Ala Asp Trp Ala Ser Gln Cys Ile Trp Glu His Ser Asn Ser
CTC GTG TGG AAC CAG ACC CTG TCG GAC AAG GCA GCG GAC TGG GCG AGC CAG TGC ATC TGG GAG CAC TCG AAC TCG

Gly Gln Asn Leu Ala Ala Trp Phe S
346 GGC CAG AAT CTG GCT GCG TGG TTT T gtgagtaatc tcgagctcga gagctgatgg gtcgggggtg gaaataaatg cgaagaggg

                                     110
431 gagaagaaag ggagatcacg acacgacgag gaatctgtga tggagggaag atgcactgac acgcatcta tgcagCG CCC CAG GCA AAC
                                     er Pro Gln Ala Asn

                                     120                                130
520 Lys Pro Met Asn Ile Ser Gln Gly Val Gly Gly Trp Asn Ala Glu Glu P
AAG CCC ATG AAC ATC TCT CAA GGC GTT GGA GGC TGG AAC GCC GAG GAG C gtgagtcgtg catgctgtta tctcgatgaa

                                     140
596 catgtttttg acatgcctca ccgtcgag CC GAC TAC AAC ACG ACG ACG TAC TCG GGC GCT GGG CAT TGG ACT CAG gtgag
ro Asp Tyr Asn Thr Thr Thr Tyr Ser Gly Ala Gly His Trp Thr Gln

                                     150
676 cgtaa ccctcggtga cgagttgacg cagcacttac cctcaatct cgtcgccatc tag GTC GTG TGG AAG AGC ACC ACG AGC GTC
Val Val Trp Lys Ser Thr Thr Ser Val

                                     160                                170
763 Gly Cys Ala Ala Tyr Ser Cys Pro Pro Gly Thr Leu Gly Arg Lys Pro Thr Asp Pro Trp Lys T
GGG TGC GCC GCG TAC TCC TGC CCA CCT GGC ACG CTC GGG CGT AAG CCG ACG GAT CCG TGG AAG A gtgagctgtc gcc

                                     180
840 ttgcgc tctcttgcaa tcaagaaaat tatcagctga ctcttcattc cag CG CTC TGG TAT TAC GTC TGC AAC TAC TAC AGG
hr Leu Trp Tyr Tyr Val Cys Asn Tyr Tyr Arg

                                     190
917 Pro Gly Asn Val Ser Pro Arg Asp Lys Tyr
CCG gtgagtgatt cgtgcacgca aaggc tagag aatgtgccca ccgatatact catagGGC AAC GTC TCA CCG CGC GAC AAG TAC

200
1007 Tyr Pro Ile Asn Val Gln Pro Ter
TAC CCT ATC AAT GTG CAG CCC TAA aggaggcct tgggaagctc gaggaaggt atagtataca cgcgccacct aggaaggcag at

1093 cgaattga cggaattgg aaagggaag aagaaagtat gatacgcgga attgtacgct cccccagtct gcttgtttgc gtcaagatgg ctc

*
1184 ggcaatg tgacagtgat tgctgaaat cagttggaca ccgggcacaa aggaaggatg agtacgacaa gatttcgggc gccagattt cgga

```

Fig. 1. Nucleotide sequence and deduced amino acid sequence of the *Sc7* gene. Coding nucleotides are in upper-case letters, all other sequences in lower-case letters; they are numbered with reference to the transcription start point (+1). Above the coding regions the predicted amino acid sequence is given and numbered with reference to the first methionine (+1). The major transcription start point is marked with an arrow, the polyadenylation site is marked with an asterisk. A possible CCAAT-box is overlined twice. The putative TATA-box is overlined. Internal intron sequences possibly involved in splicing are underlined.

sequences similar to the internal consensus sequence YGCTAAC thought to be involved in intron splicing

(Gurr *et al.*, 1987) are only present in introns 1 and 4 of both *Sc7* and *Sc14*.

```

-99  gctcaccagg tacgatcgcg aaccgagaat tcggtcgaag ctaccacag cgctcgcagt cctcgcgtat aaaggcaggc ccatgcactc
      1
-9   → cgtagagtc agcaacttac acttgcttct acactcccga tctccagcgc gtccgaac ATG AAG CTC AAC ATC GCG ATC CTC CTC
      10      20      30
176  Ala Ala Leu Ala Ala Thr Ala Ser Ala Ser Pro Ala Gly Asp Ile Ser Glu Val Ala Val Asp Ala Asp Ala Ser
      GCC GCG CTC GCC GCC ACC GCT TCC GCG AGT CCA GCT GGC GAT ATC AGC GAA GTT GCC GTC GAC GCG GAT GCC AGC

      40      50
151  Glu Leu Leu Val Asp Pro Arg Ala Glu Leu Pro Val Ala Pro Arg Ala Ser Phe Phe Ile Asp Ala Arg Gln Asp
      GAG CTC CTC GTC GAC CCG GCG GCC GAA CTC CCC GTC GCA CCC CGT GCC TCC TTC TTC ATC GAC GCC GCG CAA GAC

      60      70      80
226  Thr Asn Gly Thr Ser Gln Asp Glu Ile Asp Gln Trp Leu Thr Ala His Asn Asp Glu Arg Ala Gln His Gly Pro
      ACC AAC GGC ACC TCC CAA GAC GAG ATC GAC CAG TGG CTG ACC GCG CAC AAC GAT GAG GCG GCG CAG CAC GGG CCG

      90      100
301  Val Pro Leu Val Trp Asn Gln Asp Leu Gln Asn Ala Ala Met Ser Trp Ala Ser Arg Cys Val Tyr Lys His Asn
      GTC CCG CTC GTG TGG AAC CAG GAT CTC CAG AAC GCG GCG ATG AGC TGG GCG AGC CGT TGT GTG TAT AAG CAC AAC

      110
376  Arg Gly Gly Gln Asn Ile Ala Ala Arg Tyr A
      CGG GCG GGG CAG AAT ATC GCT GCG AGG TAC A gtgagtggca gggttttcctc gaactgggat ggggacgatt ttctagaatg gg

      120      130
457  atgcggga cgggataaac ggagcaagcg cgctgaccgc accgccacc cag AC ACG AGA GCA AAC TTC CCG CGG GAA ATC GAC
      sn Thr Arg Ala Asn Phe Pro Arg Glu Ile Asp

      140
538  Arg Ala Val Ala Gln Trp Asn Asn Glu Arg G
      CGC GCG GTC GCG CAG TGG AAC AAC GAG CGT G gtcagttttc cgctgtgcct ttgcacagag cgttcgcact tacacctaca

      150      160
618  catcgcttc cacag GC GAG TAT AAC GCC ACG ACC TTC AAG GGC GCC GGC CAC TGG ACG CAG GTC GTG TGG AAG CAC
      ly Glu Tyr Asn Ala Thr Thr Phe Lys Gly Ala Gly His Trp Thr Gln Val Val Trp Lys His

      170      180
695  Ser Arg Asn Leu Gly Cys Ala Ala Tyr Ser Cys Pro Gln Gly Thr Leu Gly Lys Lys Pro Gly Asp Lys Trp Lys
      AGC AGG AAC CTC GGG TGC GCG GCC TAC TCG TGC CCG CAG GGC ACG CTG GGG AAG AAG CCG GGC GAT AAG TGG AAG

      190
770  S T gtgcgtactt gttgcacgcc gtatagcgtt ctatctccaa ttgaatcgca tggagggatg acgtaacgcc tcgcag CC CTT TGG TAC
      er Leu Trp Tyr

      200      210
856  Tyr Val Cys Asn Tyr Asp Pro Lys
      TAC GTC TGC AAC TAC GAT CCC AAG gtgagtagcc cagtgcagga cttggcgacg atggacagcc cactcaccag agcgacg GGC
      Gly

      220      230
940  Asn Val Val Pro Ala Ser Lys Tyr Tyr Pro Ser Asn Val Gln Pro Ter
      AAC GTC GTC CCG GCG TCC AAG TAC TAC CCA AGC AAC GTG CAG CCT TGA aggcagaccg aggcggggcg agcgtgtcaa ggc

1021  gaacaat atctgggaag gggctctgagt gttgaaggca tgcgaagggt tgaagaaag tcgagtattt ggacattgta gccattgac tgca

1113  ttagga aggtcttggt ttacaaaagg aaaccttgt acttctcacg gaagatgccg ctgtcagatt caatcgggat catggccagg cgctg

1204  gacga accgaccaa ggggacgaag aagcaggacg cggccagtaa aaggacccc gaatatcgc atataggccg tccagcggtg cggata

```

Fig. 2. Nucleotide sequence and deduced amino acid sequence of the *Sc14* gene. Intron 5, which is in the 3'-nontranslated region, is underlined. No clear polyadenylation site was detected. For other details, see legend to Fig. 1.

Deduced amino acid sequences

The *Sc7* gene contains an open reading frame that encodes a protein of 204 amino acids (mol. mass 22328 Da), whereas the *Sc14* gene encodes a protein of 214 amino acids (mol. mass 23661 Da). Alignment of

the *Sc7* and *Sc14* sequences (Fig. 4) shows an overall homology of 56%, rising to 64% in the C-terminal halves of both proteins. When conserved amino acid substitutions are also taken into account the homology increases to 87%. Both *Sc7p* and *Sc14p* contain hydrophobic N-terminal sequences of about 20 amino acids with an

Fig. 3. Genomic organization of the *Sc7* and *Sc14* genes and their transcripts. (a) Map of the 9.5 kb *Nco*I fragment. Black arrows indicate the positions of the genes and the direction of transcription. Relevant restriction enzyme sites are indicated: B, *Bgl*II; E, *Eco*RI; N, *Nco*I; P, *Pvu*II; V, *Eco*RV. (b) Structure of both genes. Introns are indicated in black, whereas open blocks indicate exons. The translational start and termination sites are indicated with ATG and TAA/TGA. Black vertical arrows mark the putative TATA-boxes; the open vertical arrow marks a putative CCAAT-box. The open space in the *Sc14* gene is introduced to align the amino acid sequences.

apolar core and a lysine at position 2, indicating signal sequences for secreted proteins (Watson, 1984).

Both Sc7p and Sc14p are rich in alanine (13–14%), aspartate (6%) and asparagine (6–7%) residues. Also the aromatic amino acids tryptophan and tyrosine are quite abundant (each 4–5%) as are charged amino acids (16% for Sc7p, 20% for Sc14p). Overall, Sc7p is slightly acidic and Sc14p slightly basic. Sc7p has four potential *N*-glycosylation sites (Spiro, 1973) at positions 80, 118, 134 and 190, whereas Sc14p has only two at positions 61 and 143 (Figs 1, 2 and 4a). Except for the N-terminal signal sequences both proteins are rather hydrophilic and hydropathy plots are quite similar (Fig. 5). If the amino acid alignments (Fig. 4a) are considered, it is noticeable that all four cysteine residues, all eight tryptophan residues and seven out of nine tyrosine residues are conserved with respect to their positions, whereas two phenylalanine/tyrosine pairings and one tryptophan/tyrosine pairing are present also. According to the Dayhoff amino acid similarity index (Dayhoff, 1978), the conservation of these amino acids indicates that both proteins have similar structures.

The codon usage of *Sc7* and *Sc14* is clearly biased as is often found in highly expressed genes from filamentous fungi (Gurr *et al.*, 1987). C is preferred as third base (48% in *Sc7*, 52% in *Sc14*), whereas an A is rarely used (4–5%) and a G is clearly preferred over T as a third base (35% G, only 9–13% T). Because of the strong codon bias some codons (TTA, CTA, ATA, GTA, ACA, AAA and CGA) are not used at all.

Fig. 4. Amino acid sequence homologies between (a) Sc7p and Sc14p, (b) Sc7p, Sc14p, PR1a, PR1b, PR1c and p14 proteins. Sc7p and Sc14p sequences are from this paper. PR1a, PR1b and PR1c are PR proteins from tobacco (Cornelissen *et al.*, 1986), p14 a PR protein from tomato (Lucas *et al.*, 1985). Shadings mark the amino acids identical in all aligned sequences, asterisks mark conservative substitutions and black triangles (in a) indicate putative *N*-glycosylation sites in Sc7p and Sc14p. The sequences of the venom allergen 5 form 3 from the white-face hornet (Fang *et al.*, 1988) and of the rat sperm-coating glycoprotein (Brooks *et al.*, 1986) are not shown here, but they also show statistically significant similarity to the Sc7p and Sc14p sequences. The multiple alignments were performed with the CLUSTAL program (Higgins & Sharp, 1988).

Comparison of the amino acid sequences of Sc7p and Sc14p with the SWPROT databank shows statistically significant similarities to the pathogenesis-related proteins PR1a, PR1b and PR1c from tobacco (Cornelissen *et al.*, 1986) and p14 from tomato (Lucas *et al.*, 1985) and some more PR1 proteins from different plants. The

Fig. 5. Comparison of the hydropathy patterns of the Sc7 and Sc14 proteins. The patterns were determined using the parameters of Kyte & Doolittle (1982). A six-amino-acid window was used and plotted against position in the amino acid sequence. Peaks above the line represent hydrophobic regions.

homology between Sc7p (204 aa), Sc14p (214 aa) and the PR1 proteins (130–168 aa) includes 35 identical amino acids in all six sequences and 48 conservative substitutions (Fig. 4*b*). The Sc7 and Sc14 proteins also show significant similarities to the venom allergen antigen 5 forms 2 and 3 (214 and 215 aa) of the white-face hornet (Fang *et al.*, 1988), with 36 identical amino acids and 66 conservative substitutions, and to a sperm-coating glycoprotein from rat (Brooks *et al.*, 1986; 246 aa) and testis-specific proteins from man and mouse (Kasahara *et al.*, 1989; both 243 aa), with 29 identical residues and 54 conservative substitutions (data not shown). Unfortunately, all these proteins are of unknown functions although they all are located extracellularly.

Detection of the Sc7 protein *in vivo*

Previous studies have shown that Sc7 mRNA is more abundantly expressed than Sc14 mRNA, with ratios varying from 2.5:1 (Mulder & Wessels, 1986) to 10:1 (Ruiters *et al.*, 1988). We have repeated these RNA studies and found a ratio of 9:1 (data not shown). Therefore, we focused on the identification of the Sc7 protein *in vivo*.

Proteins from the medium of 4-d-old cultures of monokaryon and dikaryon were separated by PAGE,

blotted onto PVDF membranes and reacted with antisera raised against purified Sc3 hydrophobin and the Sc7/Protein A fusion protein. Prior to electrophoresis the proteins were treated with TFA to effect dissociation of hydrophobins (De Vries *et al.*, 1993). The reaction of the Sc3 antiserum showed the presence of Sc3p in both monokaryon and dikaryon media (Fig. 6*a*), as expected, whereas a signal with the Sc7 antiserum was present only in the dikaryon medium (Fig. 6*b*). Similar signals with the Sc7 antiserum were present in hot-SDS-extracts of the mycelium of the dikaryon (bearing fruit bodies) but not in the extracts of the monokaryon (Fig. 6*c*). Cross-reaction of Sc7 antibodies with Sc14p cannot be excluded. However, because of the higher abundance of Sc7 mRNA the major protein detected most probably is Sc7p and not Sc14p; the weaker signals in Fig. 6(*b*, *c*) may either represent different modified forms of Sc7p or result from the cross-reacting Sc14p. Clearly, these bands are specific for the dikaryon since all of them are absent in the monokaryon, both in the SDS-extract of mycelium and in the culture medium. Pre-immune sera were negative in all cases (results not shown).

The Sc7 protein running at 23 kDa can also be distinguished from the Sc3 hydrophobin, which runs at a slightly higher position (24 kDa), by differential labelling with radioactive precursors. The Sc3 hydrophobin is rich in cysteine and thus heavily labelled with [³⁵S]sulphate (Fig. 7, lane 1) but contains only one tyrosine residue and no lysine residues. Consequently, [¹⁴C]tyrosine only weakly labels Sc3 hydrophobin but heavily labels a protein in the culture fluid of the monokaryon, running at 15 kDa (lane 2). This protein is absent from the medium of the dikaryon, but [¹⁴C]tyrosine now heavily labels a protein running at 23 kDa (lane 3). [¹⁴C]Lysine heavily labels a protein running at the same position (lane 4). Because Sc7 is rich in both tyrosine and lysine we presume that this protein is Sc7p.

Sc7p occurs extracellularly in fruit bodies

Assuming Sc7p to be an excreted protein, Sc7p present in mycelial extracts should be primarily derived from hyphal walls or from the interhyphal spaces in the fruit bodies of 4-d-old dikaryotic mycelia. SDS-extracts from purified hyphal walls showed the presence of some Sc7p but most of the Sc7p was removed during the washing procedure (results not shown). Thus, Sc7p was either present intracellularly or, if extracellularly, very loosely bound to the hyphal walls. Therefore, apoplasmic fluid was collected after infiltrating fruit bodies with 50 mM-Tris/HCl, pH 7.5, followed by fragmentation of the fruit bodies and extraction of proteins with hot-SDS buffer. Silver staining detected several proteins in the apoplasmic fluid (Fig. 8, lane 1) and a very distinct immunological

Fig. 6

Fig. 6. Immunodetection of proteins from monokaryon (M) and dikaryon (D) of *S. commune* after SDS-PAGE. (a) Immunoreaction of Sc3 antibodies with proteins in the medium; (b) immunoreaction of Sc7 antibodies with proteins in the medium; (c) immunoreaction of Sc7 antibodies with SDS-extracted proteins of the mycelial mat.

Fig. 7. SDS-PAGE autoradiographs of [35 S]sulphate-, [14 C]lysine- and [14 C]tyrosine-labelled proteins in culture media from monokaryon and dikaryon of *S. commune*. Lane 1, monokaryon labelled with [35 S]sulphate; lane 2, monokaryon labelled with [14 C]tyrosine; lane 3, dikaryon labelled with [14 C]tyrosine; lane 4, dikaryon labelled with [14 C]lysine.

Fig. 7

Fig. 8. Presence of Sc7p in fruit-body extracts after SDS-PAGE. Lane 1, Apoplasmic fluid from fruit bodies, silver staining of proteins; lane 2, as lane 1 but reaction of proteins with Sc7 antibodies; lane 3, intracellular extract, silver staining of proteins; lane 4, as lane 3 but reaction of proteins with Sc7 antibodies.

signal for the presence of Sc7p (lane 2). The hot-SDS extract contained the bulk of the fruit-body proteins (lane 3) but only a faint signal for Sc7p was obtained with this extract (lane 4). We conclude that most of Sc7p

in the fruit bodies occurs loosely bound to the hyphal walls and/or in the space between the hyphae.

Previous studies (Wessels *et al.*, 1991b) have shown the presence of another abundant extracellular protein, the Sc4 hydrophobin, in the fruit bodies of *S. commune*. In contrast to Sc7p, Sc4p appears tightly bound to the hyphal walls. Both proteins may play important roles in interactions between the dikaryotic hyphae which lead to formation of pseudo-parenchymous tissue.

We thank R. C. Ullrich (University of Vermont, USA) for supplying the *S. commune* strain 4-40 cosmid library, E. Wolbert for technical assistance and J. S. Bouwer for assistance with injecting rabbits with the fusion protein. F.S. was supported by the Foundation for Biological Research (BION), which is subsidized by the Netherlands Organization for Scientific Research (NWO). E.K. was supported by grant Ko 1089/1-1 from the Deutsche Forschungsgemeinschaft.

References

- BENOIST, C., O'HARE, K., BREATHNACH, R. & CHAMBON, P. (1980). The ovalbumin gene: sequence of putative control regions. *Nucleic Acids Research* **8**, 127-142.
- BROOKS, D. E., MEANS, A. R., WRIGHT, E. J., SINGH, S. P. & TIVER, K. K. (1986). Molecular cloning of the cDNA for androgen-dependent sperm-coating glycoproteins secreted by the rat epididymis. *European Journal of Biochemistry* **161**, 13-18.
- CORNELISSEN, B. J. C., HOOFT VAN HUIJSDUJNEN, R. A. M., VAN LOON, L. C. & BOL, J. F. (1986). Molecular characterization of messenger RNAs for 'pathogenesis-related' proteins 1a, 1b and 1c, induced by TMV infection of tobacco. *EMBO Journal* **5**, 37-40.
- DAYHOFF, M. O. (1978). A model of evolutionary change in proteins.

- Matrices for detecting distant relationships. In *Atlas of Protein Sequence and Structure*, vol. 5, suppl. 3, pp. 345–358. Edited by M. O. Dayhoff, Washington, DC: National Biomedical Research Foundation.
- DE VRIES, O. M. H., FEKKES, M. P., WÖSTEN, H. A. B. & WESSELS, J. G. H. (1993). Insoluble hydrophobin complexes in the walls of *Schizophyllum commune* and other filamentous fungi. *Archives of Microbiology* **159**, 330–335.
- DONS, J. J. M., DE VRIES, O. M. H. & WESSELS, J. G. H. (1979). Characterization of the genome of the basidiomycete *Schizophyllum commune*. *Biochimica et Biophysica Acta* **563**, 100–112.
- FANG, K. S. Y., VITALE, M., FEHLNER, P. & KING, T. P. (1988). cDNA cloning and primary structure of a white-face hornet venom allergen, antigen 5. *Proceedings of the National Academy of Sciences of the United States of America* **85**, 895–899.
- GIASSON, L., SPECHT, C. A., MILGRIM, C., NOVOTNY, C. P. & ULLRICH, R. C. (1989). Cloning and comparison of *Ax* mating-type alleles of the basidiomycete *Schizophyllum commune*. *Molecular and General Genetics* **218**, 72–77.
- GURR, S. J., UNKLES, S. E. & KINGHORN, J. R. (1987). The structure and organization of nuclear genes in filamentous fungi. In *Gene Structure in Eukaryotic Microbes, SGM Special Publications vol. 26*, pp. 93–139. Edited by J. R. Kinghorn. Oxford: IRL Press.
- HARLOW, E. & LANE, D. (1988). *Antibodies: a Laboratory Manual*. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.
- HIGGINS, D. G. & SHARP, P. M. (1988). CLUSTAL: a package for performing multiple sequence alignment on a microcomputer. *Gene* **73**, 237–244.
- KASAHARA, M., GUTKNECHT, J., BREW, K., SPURR, N. & GOODFELLOW, P. N. (1989). Cloning and mapping of a testis-specific gene with sequence similarity to a sperm-coating glycoprotein gene. *Genomics* **5**, 527–534.
- KOZAK, M. (1984). Compilation and analysis of sequences upstream from the translational start site in eukaryotic mRNAs. *Nucleic Acids Research* **12**, 857–872.
- KYTE, J. & DOOLITTLE, R. F. (1982). A simple method for displaying the hydrophobic character of a protein. *Journal of Molecular Biology* **157**, 105–132.
- LAEMMLI, U. K. (1970). Cleavage of structural proteins during the assembly of the head of bacteriophage T4. *Nature, London* **227**, 680–685.
- LIPMAN, D. J. & PEARSON, W. R. (1985). Rapid and sensitive protein similarity searches. *Science* **227**, 1435–1441.
- LUCAS, J., CAMACHO HENRIQUEZ, A., LOTTSPEICH, F., HENSCHEN, A. & SÄNGER, H. L. (1985). Amino acid sequence of the 'pathogenesis-related' leaf protein p14 from viroid-infected tomato reveals a new type of structurally unfamiliar proteins. *EMBO Journal* **4**, 2745–2749.
- MULDER, G. H. & WESSELS, J. G. H. (1986). Molecular cloning of RNAs differentially expressed in monokaryons and dikaryons of *Schizophyllum commune* in relation to fruiting. *Experimental Mycology* **10**, 214–227.
- RUITERS, M. H. J., SIETSMAN, J. H. & WESSELS, J. G. H. (1988). Expression of dikaryon-specific mRNAs of *Schizophyllum commune* in relation to incompatibility genes, light, and fruiting. *Experimental Mycology* **12**, 60–69.
- SAMBROOK, J., FRITSCH, E. F. & MANIATIS, T. (1989). *Molecular Cloning: a Laboratory Manual*, 2nd edn. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory.
- SANGER, F., NICKLEN, S. & COULSON, A. R. (1977). DNA sequencing with chain terminating inhibitors. *Proceedings of the National Academy of Sciences of the United States of America* **74**, 5463–5467.
- SCHUREN, F. H. J. & WESSELS, J. G. H. (1990). Two genes specifically expressed in fruiting dikaryons of *Schizophyllum commune*: homologies with a gene not regulated by mating-type genes. *Gene* **90**, 199–205.
- SPECHT, C. A., STANKIS, M. M., GIASSON, L., NOVOTNY, C. P. & ULLRICH, R. C. (1992). Functional analysis of the homeodomain-related proteins of the *Ax* locus of *Schizophyllum commune*. *Proceedings of the National Academy of Sciences of the United States of America* **89**, 7174–7178.
- SPIRO, R. G. (1973). Glycoproteins. *Advances in Protein Chemistry* **27**, 349–367.
- SPRINGER, J. & WESSELS, J. G. H. (1989). A frequently occurring mutation that blocks the expression of fruiting genes in *Schizophyllum commune*. *Molecular and General Genetics* **219**, 486–488.
- STANKIS, M. M., SPECHT, C. A., YANG, H., GIASSON, L., ULLRICH, R. C. & NOVOTNY, C. P. (1992). The *Ax* mating locus of *Schizophyllum commune* encodes two dissimilar multiallelic homeodomain proteins. *Proceedings of the National Academy of Sciences of the United States of America* **89**, 7169–7173.
- WATSON, M. E. E. (1984). Compilation of published signal sequences. *Nucleic Acids Research* **12**, 5145–5164.
- WESSELS, J. G. H. (1992). Gene expression during fruiting in *Schizophyllum commune*. *Mycological Research* **96**, 609–620.
- WESSELS, J. G. H., DE VRIES, O. M. H., ÁSGEIRSDÓTTIR, S. A. & SPRINGER, J. (1991a). The *thn* mutation of *Schizophyllum commune*, which suppresses formation of aerial hyphae, affects expression of the *Sc3* hydrophobin gene. *Journal of General Microbiology* **137**, 2439–2445.
- WESSELS, J. G. H., DE VRIES, O. M. H., ÁSGEIRSDÓTTIR, S. A. & SCHUREN, F. H. J. (1991b). Hydrophobin genes involved in formation of aerial hyphae and fruit bodies in *Schizophyllum*. *Plant Cell* **3**, 793–799.