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

FRANK H. J. SCHUREN,* SIGRIDUR A. ÁSGEIRS-DÓTTIR, ERIKA M. KOTHE,† JOSEF 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 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., 1991a).

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., 1991a). 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., 1991b). 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., 1991b).

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-
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 \( \lambda EMBL4 \) (Schuren et al., 1990). A 1.8 kb BamHI fragment and a 0.5 kb BamHI-SalI fragment, together encompassing the entire Sc7 gene and flanking sequences, were fused in partial pBluescript IIIE. A genomic clone for Sc14 was isolated from a cosmid library of the tominogen strain 4-40 (Gissason et al., 1989). A 9.5 kb NcoI cosmid fragment containing the Sc14 gene was subcloned. This clone (p14G7) was 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 database (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 ApI 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 Smal site of pUC18. This insert was cloned as a EcoRI-SalI fragment into a Pri2T vector (Pharmacia), resulting in an in-frame fusion with the Protein A gene. An overnight culture of E. coli N4830-A cells containing the Sc7 fusion gene construct was diluted to 10^7 cells/ml ice-cold Tris/HCl, pH 9.6. The presence of Protein A/Sc7 fusion protein was detected by Western blotting using horseradish peroxidase anti-peptide 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 rhodopsin 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.

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 (Benoist 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...
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.
Deduced amino acid sequences

The Sc7 gene contains an open reading frame that encodes a protein of 204 amino acids (mol. mass 22,328 Da), whereas the Sc14 gene encodes a protein of 214 amino acids (mol. mass 23,661 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
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 (4–5% 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. 3. Genomic organization of the Sc7 and Sc14 genes and their transcripts. (a) Map of the 9.5 kb Sc7 fragment. Black arrows indicate the positions of the genes and the direction of transcription. Relevant restriction enzyme sites are indicated: B, BglII; E, EcoRI; N, ScaI; P, PstI; V, EcoRV. (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.

### 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
monokaryon and dikaryon were separated by protein. Therefore, we focused on the identification of the Sc7 proteins. The patterns were determined using the parameters of Kyte and Doolittle (1982). The patterns represent hydrophobic regions. Peaks above the line denote hydrophilic regions.

![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.](image)

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. 4b). 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 that the ratio is 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 Sc7 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. 6a), as expected, whereas a signal with the Sc7 antiserum was present only in the dikaryon medium (Fig. 6b). Similar signals with the Sc7 antiserum were present in hot-SDS-models of the mycelium of the dikaryon (bearing fruit bodies) but not in the extracts of the monokaryon (Fig. 6c). 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, apoplastic fluid was collected after infiltrating fruit bodies with 0.01 M-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 apoplastic fluid (Fig. 8, lane 1) and a very distinct immunological
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. 8. Presence of Sc7p in fruit-body extracts after SDS-PAGE. Lane 1, apoplastic 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., 1991 b) 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


