The Candida albicans Sup35p protein (CaSup35p): function, prion-like behaviour and an associated polyglutamine length polymorphism

Catarina Resende,1,2 Steven N. Parham,1 Caroline Tinsley,1 Paulo Ferreira,1,2 Júlio A. B. Duarte2 and Mick F. Tuite1

Author for correspondence: Mick F Tuite. Tel: +44 1227 823699. Fax: +44 1227 763912. e-mail: M. F.Tuite@ukc.ac.uk

1 Research School of Biosciences, University of Kent, Canterbury, Kent CT2 7NJ, UK
2 Centro de Genética e Biologia Molecular, Universidade de Lisboa, Portugal

The Sup35p protein of Saccharomyces cerevisiae is an essential translation factor whose prion-like properties give rise to the non-Mendelian genetic element [PSI+]. In this study the SUP35 gene from the related yeast species Candida albicans has been characterized. The CaSUP35 gene encodes a protein (CaSup35p) of 729 aa which shows 65% amino acid identity to the S. cerevisiae Sup35p protein (ScSup35p), with the C-terminal region showing greater identity (79%) than the N-terminal region. The full-length CaSup35p can functionally replace ScSup35p in S. cerevisiae although complementation is only complete when CaSup35p is overexpressed. Complementation only requires expression of the CaSup35p C domain. In S. cerevisiae the full-length CaSup35p is unable to establish a prion-like aggregated state even in the presence of endogenous ScSup35p prion ‘seeds’, thus confirming the existence of a species barrier in fungal prion propagation. Subcellular localization studies in C. albicans show that although CaSup35p is normally ribosome-associated, when not ribosome-associated, it does not form pelletable high-molecular-mass aggregates characteristic of the ScSup35p in [PSI+] strains. Unlike the ScSup35p, the CaSup35p N domain contains a number of polyglutamine repeats although it does contain seven copies of the peptide GGYQQ that is repeated in the ScSup35p N domain. Analysis of the CaSUP35 gene from 14 different strains of C. albicans identified four naturally occurring polymorphisms associated with changes in the length of the largest of the polyglutamine repeats. These findings have important implications for the evolution of fungal prion genes.

Keywords: SUP35 gene, polyQ, translation termination, [PSI+] determinant

INTRODUCTION

Prions are a remarkable class of proteins that were originally identified as the protein-only infectious agents associated with a range of debilitating neurodegenerative disorders of mammals, the transmissible spongiform encephalopathies otherwise known as prion diseases (Prusiner, 1982). More recently, studies with the yeast Saccharomyces cerevisiae have identified several prion proteins that are able to act as stable, transmissible determinants of phenotype. Of these yeast prion proteins, two in particular have been intensively studied: Ure2p, which gives rise to the [URE3] determinant, and Sup35p, which gives rise to the [PSI+] determinant (reviewed by Wickner & Chernoff, 1999).

The SUP35 gene was originally identified in an S. cerevisiae screen for recessive mutations which impaired the maintenance of translation fidelity (Hawthorne & Leupold, 1974). The encoded Sup35p protein was subsequently shown to be a subunit of the eukaryotic release factor forming a heterodimer with a second subunit, Sup45p (also called eRF1) that functions in
translation termination at all three stop codons (Stansfield et al., 1995b; Zhouravleva et al., 1995). Wickner (1994) suggested that Sup35p, an essential component of the translation machinery, might be a prion protein responsible for the translation-termination-related [PSI-] phenotype of S. cerevisiae. This hypothesis has now been confirmed by a wide range of biochemical and genetical studies (reviewed by Wickner & Chernoff, 1999).

[PSI-] strains have an elevated level of suppressor tRNA-mediated readthrough of nonsense codons compared with wild-type [psi−] strains (Cox, 1965; Firoozan et al., 1991). Analysis of a mutation located in the SUP35 gene (Doel et al., 1994), coupled with the demonstration that the deletion of the non-essential N-terminus of the protein (Ter-Avanesyan et al., 1994) eliminated both the [PSI-] phenotype and the increased de novo appearance of [PSI+] following SUP35 overexpression (Chernoff et al., 1993), showed that SUP35 is the [PSI+] determinant.

In [PSI+] strains, the majority of the cellular pool of Sup35p, not associated with Sup45p and/or the ribosome, is present as part of a high-molecular-mass aggregate which is incompatible with its translation release factor activity, leading to the translation-termination-associated phenotype characteristic of [PSI+] mutants (Patino et al., 1996; Paushkin et al., 1996, 1997). The conversion of soluble Sup35p to the aggregated prion form is promoted by a yet-to-be fully defined conformational change in Sup35p and is driven by the accumulating altered conformer (Serio et al., 2000; Sparrer et al., 2000). Such a self-catalysed propagation of an altered protein conformer parallels the behaviour of the mammalian prion protein PrP (reviewed by Prusiner et al., 1998).

Analysis of the known eukaryotic Sup35p protein sequences suggests they all contain two distinct regions: an evolutionarily conserved C-terminus, which in S. cerevisiae is essential for translation termination and hence cell viability, and a less conserved N-terminal extension of varying length and amino acid composition (Kushnirov et al., 1988; Wilson & Culbertson, 1988). The prion conversion of Sup35p is promoted by this N-terminal prion-forming domain (PrD) of Sup35p (reviewed by Tuite, 2000), a region of the protein molecule which is not required for eRF3 activity (Ter-Avanesyan et al., 1993). Within the Sup35p PrD lie a series of five nonapeptide repeats with the core motif GGYQQ. Deletion of the repeat-containing region leads to an inability of the otherwise functional Sup35p to be propagated in its prion-like form (Ter-Avanesyan et al., 1993) and we have recently shown that deletion of even one of the nonapeptide repeats leads to loss of the prion-like behaviour of Sup35p (Parham et al., 2001). Intriguingly the human prion protein PrP also contains a series of sequence-related octapeptide repeats near its N-terminus which have also been implicated in mammalian prion propagation (Flechsig et al., 2000).

Studies on Sup35p have recently been extended to a variety of other fungal species, including Candida albicans and Pichia methanolica (Santoso et al., 2000; Kushnirov et al., 2000; Chernoff et al., 2000). These studies have indicated that the N-terminal regions of these Sup35p homologues have diverged considerably from the S. cerevisiae Sup35p PrD at the amino acid sequence level although they are still Gln/Asn rich. Interestingly, the CaSup35p NM sequence reported by Santoso et al. (2000) also contained a number of short poly(Gln) sequences. Replacing the ScSup35p N or NM region with the corresponding heterologous Sup35p region did not prevent the chimaeric Sup35p from forming a prion-like aggregate in S. cerevisiae, although such aggregates could only be induced by the overexpression of the corresponding heterologous Sup35p sequence. That overexpression of the ScSup35p NM region did not induce the formation of the chimaeric Sup35p aggregates was taken as evidence for a ‘barrier’ to interspecies prion propagation in yeasts. Furthermore, Santoso et al. (2000) showed that replacing a region of the CaSup35p N domain with residues 8–26 from the ScSup35p N domain overcame this barrier to seeding by ScSup35p NM. Further studies by Chien & Weissman (2001) have indicated that this ‘species barrier’ is probably due to the different Sup35p N domains forming distinct, non-interacting conformations with the S. cerevisiae protein. While the aggregation of chimaeric Sup35p protein can be seeded by homologous N domains in S. cerevisiae, there is no evidence that the full-length heterologous Sup35p behaves in the same way in S. cerevisiae or is able to establish a prion-like state in its natural host.

To further address the question of the conservation of function and prion-like properties of Sup35p in fungal species other than S. cerevisiae, we have isolated and characterized the full-length SUP35 gene from C. albicans (CaSUP35). This species of yeast, although closely related to S. cerevisiae, is pathogenic in man and is frequently encountered as a commensal of the human digestive system and vaginal tract (Smith, 1989). It has thus evolved to colonize distinct ecological niches to those of S. cerevisiae.

**METHODS**

**Bacterial and yeast strains used.** Escherichia coli strains used were DH5α [supE44 lacU169 (ΔlacZ M15) bsdR17 recA1 endA1 gyrA96 thi-1 relA] (Hanahan, 1983) and JM109 [recA1 supE44 endA1 bsdS17 gyrA96 relA thi Δ(lac–proAB)] (Yanisch-Perron et al., 1985).

Genotypes of the S. cerevisiae strains used in this work and their origins are shown in Table 1. The diploid strain MT26 was obtained by mating the two haploid strains BSC783/4a and BSC783/4c. MT26.4 is a derivative of MT26 in which one of the SUP35 alleles was deleted and replaced with a kantMX4 allele, as described below. The haploid strains MT26.4/3c and MT26.4/5c were recovered from the meiotic progeny of MT26.4 transformed with pDR4, a centromeric plasmid containing the URA3 and the C. albicans SUP35 gene (see below and Table 3). The haploid strain MT700/94 [PSI+] was recovered from the meiotic progeny of the diploid strain MT700 transformed with pYK810, a centromeric plasmid...
Serat 30 manipulator (model MSM System). Yeast cultures were grown et al procedures and genetic techniques were used (Kaiser et al (Kikuchi). Yeast growth conditions.

Standard yeast media, cultivation containing the URA3 and the S. cerevisiae SUP35 genes (Kikuchi et al., 1988; Table 3). The MT700/9d [psi] derivative was created by growth in the presence of 3 mM GuHCl (Tuite et al., 1981). The diploid strain MT700 was obtained by crossing MT26/3c with BSC772/18c. Table 2 provides details of all laboratory and clinical strains of C. albicans used in this study.

**Table 1. Genotypes of strains of S. cerevisiae used in this study**

<table>
<thead>
<tr>
<th>Strain</th>
<th>Genotype</th>
</tr>
</thead>
<tbody>
<tr>
<td>BSC483/1a</td>
<td>MATα SUQ5 ade2-1 his5-2 lys1-1 can1-100 PNM2 ura3-1 sal3 [psi−]</td>
</tr>
<tr>
<td>JD53-15</td>
<td>MATα SUQ5 ade2-1 his5-2 lys1-1 can1-100 PNM2 ura3-1 sal3 [psi−]</td>
</tr>
<tr>
<td>PFI-28</td>
<td>MATα SUQ5 ade2-1 his5-2 lys1-1 can1-100 PNM2 ura3-1 leu2-3/112 [psi−]</td>
</tr>
<tr>
<td>BSC783/4a</td>
<td>MATα SUQ5 ade2-1 his3-11 his3-15 ura3-1 leu2-3 leu2-112 [psi−]</td>
</tr>
<tr>
<td>BSC783/3c</td>
<td>MATα SUQ5 ade2-1 his3-11 his3-15 ura3-1 leu2-3 leu2-112 [psi−]</td>
</tr>
<tr>
<td>BSC783/4a sal3-4</td>
<td>MATα SUQ5 ade2-1 his3-11 his3-15 ura3-1 leu2-3 leu2-112 sal3-4 [psi−]</td>
</tr>
<tr>
<td>MT26</td>
<td>MATα/MATα SUQ5/SUQ5 ade2-1 his3-11 his3-15 ura3-1 ura3-1 leu2-3 leu2-112</td>
</tr>
<tr>
<td></td>
<td>[psi−]</td>
</tr>
<tr>
<td>MT26.4</td>
<td>As for MT26 but also sup35::kanMX4/SUP35+</td>
</tr>
<tr>
<td>MT26.4/3c</td>
<td>MATα sup35::kanMX4 SUQ5 ade2-1 his3-11 his3-15 ura3-1 leu2-3 leu2-112 [psi−]</td>
</tr>
<tr>
<td>MT26.4/5c</td>
<td>MATα sup35::kanMX4 SUQ5 ade2-1 his3-11 his3-15 ura3-1 leu2-3 leu2-112 [psi−]</td>
</tr>
<tr>
<td>BSC772/18c</td>
<td>MATα SUQ5 ade2-1 his4-166 ura3-1 leu2-2 lys-1 can1-100</td>
</tr>
<tr>
<td>MT700/9d [PSI+]</td>
<td>MATα SUQ5 ade2-1 his3 ura3-1 leu2 sup35::kanMX</td>
</tr>
<tr>
<td>MT700/9d [psi−]</td>
<td>MATα SUQ5 bis4 lys1 ura3-1 leu2 sup35::kanMX</td>
</tr>
</tbody>
</table>

- Provided by F. C. Odds (Aberdeen, UK).

**Table 2. Strains of C. albicans**

<table>
<thead>
<tr>
<th>Strain</th>
<th>Origin/source</th>
</tr>
</thead>
<tbody>
<tr>
<td>2005E</td>
<td>GlaxoWellcome, UK</td>
</tr>
<tr>
<td>10261</td>
<td>ATCC</td>
</tr>
<tr>
<td>W10</td>
<td>D. Soll, University of Iowa, USA</td>
</tr>
<tr>
<td>IGC 2427</td>
<td>Tree leaves, Gulbenkian Institute, Portugal</td>
</tr>
<tr>
<td>WHH1, 3, 4, 7</td>
<td>Clinical samples, J. Reuther, William Harvey Hospital, UK</td>
</tr>
<tr>
<td>5α</td>
<td>Skin (Spain)</td>
</tr>
<tr>
<td>78/011α</td>
<td>Oral cavity (UK)</td>
</tr>
<tr>
<td>81/175α</td>
<td>Fingernail (USA)</td>
</tr>
<tr>
<td>81/191α</td>
<td>Blood (USA)</td>
</tr>
<tr>
<td>81/192α</td>
<td>Vagina (USA)</td>
</tr>
<tr>
<td>J930510α</td>
<td>Oral cavity, AIDS patient (Germany)</td>
</tr>
</tbody>
</table>

**Yeast growth conditions.** Standard yeast media, cultivation procedures and genetic techniques were used (Kaiser et al., 1994). Tetrad dissection was performed on a Singer Micro-manipulator (model MSM System). Yeast cultures were grown at 30 °C. The presence of the [PSI] determinant was monitored phenotypically by suppression of the ochre mutation ade2-1 in strains carrying the ochre suppressor tRNA α gene SUQ5. This allows for a direct visualization of allosuppression by colony-colour and/or adenine prototrophy (Cox, 1965). Thus [psi−] strains are red on complete (YPD) medium and do not grow on medium lacking adenine (SD—Ade) whereas [PSI+] strains are white or pink on YPD medium and grow on SD—Ade after 4 days incubation at 30 °C. To enhance the red colony-colour, strains were grown on 1/4YPD solid medium (4%, w/v, glucose; 1%, w/v, Bactopeptone; 0.25%, w/v, yeast extract; 2%, w/v, agar). Elimination of [PSI+] with 3 mM GuHCl (Tuite et al., 1981) was performed as previously described. Resistance to canavanine was monitored in defined minimal medium supplemented with canavanine (25 μg ml−1). Geneticin (G418)-resistant strains were selected in YPD + G418 (200 μg ml−1).

**Recombinant DNA techniques.** Standard protocols were used for nucleic acid isolation, electrophoresis, fragment purification, restriction enzyme digestion, Southern blotting and PCR (Sambrook et al., 1989). Restriction enzymes and Taq DNA polymerase (Expand) were purchased from Boehringer Mannheim. For Southern hybridization analysis of isolated plasmid DNA, DNA probes were labelled with DIG-11-dUTP (Boehringer Mannheim) using the random primer method (Feinberg & Vogelstein, 1983) and following the manufacturer’s instructions. Hybridization was carried out using a nylon membrane (Hybond-N; Amersham) according to the manufacturer’s instructions, as was luminescent detection (Boehringer Mannheim) of bound probe. For Southern blot analysis of C. albicans genomic DNA, a CaSUP35 gene probe...
was obtained from the plasmid pDR4 (Table 3) as a 2 kb SpeI–KpnI fragment and labelled with [α-32P]dCTP by the random primer method (Amersham), according to the manufacturer’s instructions.

Isolation and sequencing of the C. albicans SUP35 gene. Plasmid pDR1 (Table 3) containing the C. albicans SUP35 gene (CaSUP35) was isolated from a C. albicans genomic library cloned into the yeast single-copy shuttle Ycp50 vector (Rose et al., 1987). The presence of the Sup35p homologue was expected to restore efficient translation termination and this could be monitored phenotypically by antisuppression of the can1-100 (UAU) mutation in a strain carrying the ochre suppressor tRNase 1° formed as follows. After transformation of BSC483 (Ura-), cells were incubated for a further 18 h at 30 °C, with aeration, in defined liquid minimal medium (SD–Ura). Cells were then plated on to SD–Ura + canavanine (25 µg ml−1) to allow for simultaneous selection of transformants that were Ura+ and canavanine resistant, the latter arising from the reduced efficiency of suppression of the can1-100 allele.

The plasmid pDR1 identified by this strategy further resulted in simultaneous loss of the suppression phenotype for several other ochre mutations in the strain, i.e. ade2-1, his5-2 and lys1-1. pDR1 transformants were then assayed for co-loss of the complementing phenotype simultaneously with loss of the pDR1 plasmid-borne URA3 marker following culturing in non-selective YPD medium. Transformation of a second strain JDS3-15 (Table 1) was used to confirm the complementing phenotype of pDR1.

The cloned CaSUP35 gene was sequenced by the chain-termination method (Sanger et al., 1977) by either Advanced Biotechnology Centre (Charing Cross & Westminster Medical School, London, UK) or MWG Biotech (Milton Keynes, UK). A variety of subclones derived from pDR1 were generated to facilitate gene localization and DNA sequencing. Both strands of the CaSUP35 gene were sequenced and the sequence data deposited at GenBank (accession no. AF020554). The N domains of the CaSUP35 gene were PCR-amplified from 14 different C. albicans strains (Table 2), using ‘high fidelity’ Pwo polymerase (Roche) with a 3′ to 5′ proofreading activity, and oligonucleotide primers NTCaSUP35I (ACTGACGGAATCCGGTCTGGAAGTGGTCTTC) and NTCaSUP35III (CGGAGCTGAGCTGGAGAT). Both strands of each of the CaSUP35 N domains were sequenced and the sequence data deposited with GenBank (accession nos AY028660–AY028673).

Plasmid constructions. Plasmid pDR2 was made by subcloning a 5.2 kb BamHI–SalI fragment from plasmid pDR1 into plasmid pRS416 (Sikorski & Hieter, 1989). This DNA fragment carried a region showing homology to the S. cerevisiae SUP35 (ScSUP35) gene probe as detected by the Southern blot analysis (data not shown). A 1.3 kb EcoRV fragment that did not hybridize to the ScSUP35 gene was deleted from the plasmid pDR2 to produce the URA3-based centromeric plasmid pDR4. The multicopy plasmid pDR6, containing the complete CaSUP35 gene, was constructed by subcloning the 3.9 kb BamHI–SalI fragment from pDR4 into the LEU2d-2µ plasmid pMA3a (Spalding & Tuite, 1989). Plasmid pUKC1530 contained the complete CaSUP35 gene, including the promoter, cloned into pRS313 (Sikorski & Hieter, 1989) as a BamHI–SalI fragment generated from plasmid pDR4.

Plasmid pUKC1530-C, expressing the C domain of CaSup35p, was constructed as follows. The fragment containing the CaSUP35 C domain was PCR-amplified from plasmid pDR4 using oligonucleotide primers CaSUP35SpeI (GAGAGCAGCTGACGAAGATTGTTCAAG) and CaSUP35SacI (GAGAAAAGCTCCCCTGGTACTCATGCATTGC) bearing extensions with SpeI and SacI restriction sites respectively (underlined). The suitably digested amplified fragment was then inserted into the plasmid pUKC1508 (with the SUP35 promoter inserted into pRS313) to generate pUKC1530-C.

Construction of an S. cerevisiae SUP35 gene deletion strain for plasmid shuffling. The sup35::kanMX4 disruption was constructed by a directed PCR-mediated gene deletion strategy (Wach et al., 1994) using the plasmid pFA6a-KanMX4 as template. The PCR primers used were: S1, TGCTCGGAATTTAATCGAT-CACTAGTGACGGAAGATTGTTCAAG and S2, ATATTTTAGGTGCACTTTACTTAATTGTTCGGAGAATTGACGAGCTGAGCTGGAGAT. Oligonucleotides S1 and S2 were designed such that their 5′ regions were complementary to the sequence flanking the chromosomal SUP35 coding sequence, while the 3′ portions were complementary to the sequence flanking the kanMX4 cassette. The resulting PCR-amplified 1.59 kb kanMX4 cassette, with flanking extensions homologous to the sequences flanking the ScSUP35 gene, was then transformed into strain MT26 (Table 1) and sup35::
Sup35 gene polymorphism in C. albicans

kanMX4/SUP35 heterozygous transplacements selected on YPD + G418 (200 μg ml⁻¹) plates and verified by PCR (data not shown). The PCR primers used were: A1, CGCATTAGCCTCGCTATTTTG; A2, GCTGATGTGCTGGGCTGG; A3, GAGGATCAAGGTACCCAC; A4, GATGATGATGATGCGAG; K2, GATTGCATGCATCTCGG; and K3, TATGGGACTCGCTCGGTG. Primers A1–A4 were complementary to the ScSup35 sequence while K2 and K3 were complementary to the kanMX4 cassette and were designed for PCR verification of the expected sup35::kanMX4 replacement. The resulting kanamycin-resistant diploid strain, MT26.4, was then transformed with pDR4 (Table 3), and two haploid strains with the sup35::kanMX4 disruption carrying the pDR4 plasmid (designated MT26.4/3c and MT26.4/5c; Table 1) were identified amongst the meiotic progeny.

Plasmid shuffling to test CaSup35 function. SUP35 genes were introduced either into the S. cerevisiae haploid strain MT700/9d (both [PSI⁺] and [psi⁻] derivatives: Table 1; Parham et al., 2001), which carried both the sup35::kanMX allele and the viability-maintaining, single-copy plasmid pYK810 (Kikuchi et al., 1988) bearing the ScSUP35 and URA3 genes, or into the [psi⁻] strains MT26.4/3c or 5c carrying the plasmid pDR4 (Table 3). These strains were transformed with the various plasmid constructs carrying the HIS3 marker and one or other of the different SUP35 genes. His⁺ transformants were selected in SD−His medium and, single transformants were then spread onto both solid SD−His medium and YPD medium containing 5-fluoroorotic acid (1 g l⁻¹), selecting for strains that had lost the URA3-containing plasmid (Kaiser et al., 1994).

Western blot analysis. Yeast protein extracts were prepared and fractionated into soluble and pelletable (i.e. insoluble) fractions by centrifugation essentially as described by Parham et al. (2001). Protein samples were analysed using 10% SDS-polyacrylamide gels and electrophoretically transferred to nitrocellulose (Sartorius) for Western blot analysis, employing polyclonal antiserum raised against ScSup35p expressed in E. coli, essentially as described by Stansfield et al. (1995a, b). Anti-rabbit secondary antibody was purchased from Amersham. The ECL reagent (Amersham) was used for detection of bound antibody according to the manufacturer’s protocols. Molecular masses of proteins were estimated using pre-stained molecular mass markers (Sigma).

Preparation of yeast ribosomes and post-ribosomal supernatants. Both S. cerevisiae and C. albicans strains were grown at 30 °C in YPD to a density of 1–2×10⁸ cells ml⁻¹ and harvested by centrifugation. Total cell lysates were then prepared in A buffer (25 mM Tris pH 7.2, 50 mM KCl, 5 mM MgCl₂, 5 mM β-mercaptoethanol, 10 mM PMSF) essentially as described by Stansfield et al. (1992) and loaded onto a centrifuge rotor. The resulting post-ribosomal supernatant was divided into aliquots and the ribosomal pellet resuspended in A buffer.

Quantification of nonsense suppression levels in vivo. Yeast strains were transformed with one or other of the two ‘read-through’ assay vectors, pUKC815-L or pUKC817-L (Stansfield et al., 1995a). The single-copy vector pUKC815-L carries the lacZ gene under the control of the constitutive PGK promoter. The read-through vector pUKC817-L has an in-frame premature ochre (TAA) termination codon at the PGK–lacZ junction. The degree of nonsense codon suppression was determined essentially as described by Stansfield et al. (1995a).

Units of β-galactosidase activity were calculated according to Miller (1972). β-Galactosidase activity was expressed as a percentage of the control activity (i.e. that of strains bearing the control vector pUKC815-L).

RESULTS

Isolation and analysis of the CaSUP35 gene

The C. albicans SUP35 gene homologue (CaSUP35) was identified in a genomic library by its ability to functionally complement the termination phenotype of the S. cerevisiae sal3-4 allosuppressor allele of the SUP35 gene (Crouzet et al., 1987). A 2.6 kb C. albicans genomic DNA fragment from the single clone identified (plasmid pDR1) was completely sequenced on both strands and the resulting DNA sequence (GenBank accession no. AF020554) identified an open reading frame of 729 residues (Fig. 1a), with a predicted molecular mass of 80.6 kDa. The C-terminal 430 aa sequence (amino acids 300–729) was highly homologous to the corresponding region of S. cerevisiae Sup35p (ScSup35p) (79% amino acid identity) as well as to the translation elongation factor eEF1A of C. albicans (35% amino acid identity).

Fig. 1. The C. albicans Sup35p protein. (a) Amino acid sequence of the N domain (amino acids 1–143) of CaSUP35 deduced from the CaSUP35 gene sequence (GenBank accession no. AF020554). The three regions identified are equivalent to the N, M and C domains defined in the ScSup35p sequence (Ter-Avanesyan et al., 1993) based on amino acid identity and the position of the indicated Met residues at residues 1 and 144. The three poly(Gln) repeats are underlined. (b) The seven copies of the GGYQQ repeat in the putative N domain of CaSup35p.

Alignment between the ScSup35p and CaSup35p sequences identified the region between amino acids 1 and 143 of the CaSup35p as corresponding to the ScSup35p PrD (Fig. 1a). Although the putative CaSup35p PrD shows little primary amino acid sequence identity to the
ScSup35p PrD, it is similarly rich in Gln and Asn residues (ScSup35p 53/123, 43%; CaSup35p 74/143, 53%). The core motif of the ScSup35p PrD oligopeptide repeat, GGYQQ, is present as five perfect and two imperfect copies in the CaSup35p N domain (Fig. 1b). An additional striking feature of the CaSup35p putative PrD is the presence of three distinct poly(Gln) tracts, the largest containing a run of 13 Gln residues (amino acids 27–39; Fig. 1b).

Hybridization experiments with digested total genomic DNA from C. albicans strain 2005E, probed with a 2 kb SpeI–Kpn1 fragment containing the CaSUP35 gene, demonstrated that CaSUP35 was present as a single copy sequence (data not shown). Northern blot analysis confirmed that CaSUP35 encodes a single 2400 nt mRNA transcript (data not shown).

**Table 4. Efficiency of nonsense suppression in strains carrying different combinations of S. cerevisiae and C. albicans SUP35 gene alleles**

<table>
<thead>
<tr>
<th>Strain</th>
<th>SUP35 gene*</th>
<th>Readthrough (%)†</th>
</tr>
</thead>
<tbody>
<tr>
<td>BSC783/4a</td>
<td>SUP35</td>
<td>0.82±0.28</td>
</tr>
<tr>
<td>BSC783/4a sal3-4</td>
<td>sal3-4 [ScSUP35]</td>
<td>23.92±8.15</td>
</tr>
<tr>
<td>BSC783/4a sal3-4</td>
<td>sal3-4 [CaSUP35]</td>
<td>4.19±0.70</td>
</tr>
<tr>
<td>BSC783/4a</td>
<td>SUP35 [CaSUP35]</td>
<td>9.66±6.80</td>
</tr>
<tr>
<td>BSC783/4a sal3-4</td>
<td>sal3-4 [CaSUP35]</td>
<td>8.02±2.47</td>
</tr>
<tr>
<td>MT26.4/3c</td>
<td>sup35::kanMX [CaSUP35]</td>
<td>12.85±4.11</td>
</tr>
</tbody>
</table>

* Plasmid-borne genes are shown in square brackets. The plasmid carrying the ScSUP35 gene was pYK810 and the plasmid carrying the CaSUP35 gene was pDR4, both single-copy, CEN-based plasmids.

† β-Galactosidase activity is expressed as a percentage of the control, i.e. the relevant strain bearing the control vector pUKC815-L. The data are the means ± SD of four independent assays using two different transformants.
with both the *sal3-4* allele of *SUP35* (see above) and the *sup35::kanMX* disruption. The co-expression of CaSup35p in a strain expressing wild-type ScSup35p also resulted in a detectable decrease in termination efficiency. These data suggest that the CaSup35p interacts less efficiently with components of the endogenous translation termination machinery, most likely eRF1 (Stansfeld *et al.*, 1995b).

**CaSup35p is ribosome-associated but does not form high-molecular-mass aggregates in *C. albicans***

CaSup35p contains an epitope that is recognized by our anti-ScSup35p polyclonal antibody (Fig. 2; lanes 2–5) and we therefore were able to use this antibody to study the subcellular location of the CaSup35p. Studies with ScSup35p (Didichenko *et al.*, 1991) have shown that when translation elongation is ‘frozen’ by the addition of cycloheximide, ScSup35p is found associated with the ribosome. We confirmed that this was also the case for CaSup35p in *C. albicans*, with the CaSup35p being clearly present in the ribosomal fraction (Fig. 3a).

Given the known propensity for ScSup35p to form high-molecular-mass aggregates in [PSI+] strains of *S. cerevisiae* (Patino *et al.*, 1996; Paushkin *et al.*, 1996) we next determined the state of the CaSup35p in *C. albicans* in non-cycloheximide-treated cells. In this analysis several different naturally occurring clinical strains of *C. albicans* were studied (see below). In each strain, the majority of the detectable CaSup35p was in the soluble fraction, with little if any being detected in the high-molecular-mass fraction (Fig. 3b). This would suggest that these various strains do not contain CaSup35p in a highly aggregated form indicative of the prion state.

Analysis of the CaSup35p levels in both the yeast and hyphal phases of *C. albicans* indicated that this trans-
The CaSUP35 gene does not sustain the [PSI+] determinant in S. cerevisiae

It has been previously shown by others that overexpression of the CaSUP35 N domain in S. cerevisiae does not induce the de novo conversion of endogenous ScSUP35 to form the [PSI+] determinant (Santoso et al., 2000; Chien & Weissman, 2001). To determine whether overexpression of the full-length CaSUP35 in S. cerevisiae might be more effective as a cross-species seeding agent we utilized the strain MT700/9d. This haploid strain carries a sup35::kanMX disruption, with viability being maintained by the single-copy plasmid pYK810 carrying the wild-type ScSUP35 gene (see Methods). Starting with either a [PSI+] or a [psi−] derivative of this strain, in conjunction with a plasmid-shuffling strategy (Parham et al., 2001) we introduced either a single-copy plasmid (pUKC1530) or a multicopy plasmid (pDR6) carrying the full-length CaSUP35 gene into both strains. Following elimination of the ScSUP35-bearing plasmid pYK810, the CaSUP35 was unable to induce or sustain the [PSI+] phenotype in either strain, even when introduced on a multicopy plasmid (Fig. 4a). In a control experiment, the introduction of the ScSUP35 gene on plasmid pYK810 maintained [PSI+] after plasmid shuffling (Fig. 4a). The weak termination defect (i.e. allosuppressor) associated with the pDR4 transformant (see above) was also observed with the pUKC1530 transformant although, as with the pDR4 transformant, this phenotype was not eliminated following growth on 3 mM GuHCl (Tuite et al., 1981) confirming this was not a [PSI+]-associated phenotype. In strains carrying the multicopy plasmid pDR6, this allosuppression phenotype was no longer detected. Thus, full-length CaSUP35 cannot be seeded by ScSUP35 [PSI+] aggregates in vivo to undergo the necessary conformational change required to sustain Sup35 aggregates (i.e. the [PSI+] determinant) in S. cerevisiae.

We also repeated the experiment using a single-copy plasmid that only expressed the C domain (amino acids 300–729; Fig. 1a) of CaSUP35, i.e. plasmid pUKC1530-C. Post-shuffling, viable Ure− transformants were obtained, confirming that it is the C domain of CaSUP35 that also carries the essential function as for ScSUP35 (see above). The resulting transformants clearly had a [psi−] phenotype (Fig. 4b; right panel) but no termination defect – as indicated by a weak suppression phenotype – was observed, in contrast to the results with pUKC1530 (Fig. 4a). This would suggest that the M domain of CaSUP35 may interfere with the interaction with the endogenous Sup45p (eRF1). When the C domain of CaSUP35 was co-expressed with full-length Sc SUP35 in a [PSI+] strain (Fig. 4b; left panel), no antisuppressor phenotype was observed, indicating that the expressed heterologous C domain could not support the [PSI+] related termination defect. Western blot analysis of these strains failed to detect the expressed C domain of CaSUP35, indicating that the epitope recognized by the anti-ScSUP35 antibody mostly likely lies in the N/M domain of CaSUP35.

Natural variation in the N domain of CaSUP35

To determine whether the amino acid sequence of the N domain of CaSUP35 was conserved between different strains of this pathogenic fungal species we next PCR-amplified the region of the CaSUP35 gene encoding amino acids 1–143 from 14 different clinical strains obtained from a variety of sources (Table 2) using a high-fidelity Taq polymerase (Pwo) with a 3’ to 5’ exo-proofreading activity. Among the 14 strains analysed...
eight different CaSUP35 alleles were observed (Fig. 5). Most significantly, there were four alleles with differing numbers of Gln residues in the largest of the poly(Gln) repeat regions in the N domain (Fig. 1a; amino acids 27–30) with lengths of 7, 9, 11 and 13 residues. The other four variants noted were single amino acid substitutions at positions 2, 6 and 12 in the region N-terminal to the variable-length poly(Gln) repeat region, and at position 93 within one of the copies of the repeated GGYQQ motif.

**DISCUSSION**

**CaSup35p is a functional homologue of ScSup35p**

The *C. albicans* SUP35 gene encodes a protein showing a significant level of overall amino acid identity to *S. cerevisiae* SUP35p (65% identity). As would be expected, given this high degree of amino acid sequence conservation, CaSup35p is able to restore viability to an otherwise lethal disruption of the homologous gene in *S. cerevisiae*. The association of CaSup35p with the ribosome confirms the functional equivalence of CaSup35p with ScSup35p. To date, the only known function of SUP35p in yeasts is as a subunit of the translation termination release factor and, as such, it interacts with Sup45p (eRF1) to mediate translation termination at all three stop codons (Stansfield et al., 1995a; Zhuravleva et al., 1995). We have also recently identified a close homologue of SUP45p in *C. albicans* (A. Brandao-Burch & M. F. Tuite, unpublished data). Although mRNA decoding in *C. albicans* is non-standard (Santos & Tuite, 1995), there is no evidence that *C. albicans* decodes any of the three stop codons in any way other than as stop even though such genetic code deviations are relatively common compared to sense codon reassignments.

The ScSup35p molecule appears to have three distinct functional domains, N, M and C. These were originally annotated on the basis of the location of the first three Met residues in the ScSup35p sequence, Met1, Met124 and Met 254 (Kushnirov et al., 1988; Wilson & Culbertson, 1988). Although the equivalent N region of the CaSup35p contains an additional Met residue (Met 15; Fig. 1a), the Met residues at positions 144 and 300 in the CaSup35p sequence correspond to the Met residues used to define the N, M and C regions of ScSup35p. The C domain of ScSup35p (amino acids 254–685) is known to carry the function with regard to translation termination (Ter-Avanesyan et al., 1993) and we show here that the equivalent C domain of CaSup35p also carries the same essential function (Fig. 4b). Both the CaSup35p and the ScSup35p C domains show significant amino acid identity to translation elongation factor eEF1A (Kushnirov et al., 1988; Wilson & Culbertson, 1988; Ter-Avanesyan et al., 1993). While the role of eEF1A is well defined in the translation elongation cycle, the exact role of ScSup35p, and more specifically its essential C domain, in translation termination remains to be established. The presence of GTP-binding domains, which are conserved between eEF1A and SUP35p from both *S. cerevisiae* and *C. albicans*, implies that GTP hydrolysis is required for SUP35p activity although, as with other translation factors with such G domains, this may simply regulate binding of the substrate, which, in this case, would be SUP45p (eRF1) and/or the ribosome.

The failure of the CaSup35p to fully restore efficient translation termination in either an *S. cerevisiae* sup35 null mutant or a *sal3-4* mutant, as defined by the weak nonsense suppressor phenotype, is not a consequence of lower cellular levels of the heterologous SUP35p (Fig. 2). Rather, it may simply reflect lower efficiency of binding to SUP45p (eRF1) or another, yet-to-be-defined, component of the translation termination machinery. Over-expression of CaSup35p does however restore efficient termination (Fig. 4a), suggesting that a combination of lower efficiency of initial binding between CaSup35p and ScSup45p with the elevated level of CaSup35p in this strain ensures that a level of functional CaSup35p–ScSup45p complex is formed sufficient to allow for efficient termination of translation.

**The N domain of CaSup35p**

CaSup35p has an N domain of 143 aa which shows many of the sequence features of the well-defined N-terminal prion-forming domain (PrD) of ScSup35p (Tuite, 2000). In particular, there is a high content of Gln and Asn residues (53%) and a low content of charged residues (5%). Between the PrD and the essential C domain lies the functionally undefined M domain, whose role in ScSup35p function in termination is dispensable although it remains to be established whether this part of the protein contributes to the prion-like behaviour of ScSup35p. The ScSup35p M domain is 130 aa in length and is highly charged, with Lys and Glu representing 36% of the residues. The corresponding region in CaSup35p is some 26 residues longer, with 26% of the residues being Lys and Glu. Alignment between the putative M regions of published fungal Sup35p sequences shows little evidence of evolutionary conservation of primary sequence, although a consistent feature is a high proportion of Lys and Glu residues (data not shown).

Santoso et al. (2000) have recently shown that a chimeraic Sup35p containing the NM domain of CaSup35p (amino acids 1–299) and the C domain of ScSup35p could not be incorporated into endogenous ScSup35p prion aggregates in a [*PSI*] strain. Furthermore, these authors have now shown that this apparent species barrier is probably the consequence of differences in conformation and hence seeding potential (Chien & Weissman, 2001). In our experiments reported here, we demonstrate that full-length CaSup35p is similarly unable to be seeded by the endogenous ScSup35p prion even when overexpressed (Fig. 4a). This would suggest that the lack of the CaSup35p C domain in the recombinant molecules used in the cross-seeding experiments of Santoso et al. (2000) did not alter the inter- or intramolecular interactions required for aggregate form-
ation, ruling out the need for an interaction between the CaSup35p N and C domains in prion aggregation. This would in turn rule out a role for the C domain in establishing the ‘species barrier’ to fungal prion propagation. We have recently shown that replacing the CaSup35p M domain in a hybrid Sup35p molecule with the ScSup35p M domain does not overcome this potential ‘species barrier’ (C. Resende, S. N. Parham & M. F. Tuite, unpublished data). However, Santoso et al. (2000) showed that a CaSup35pN–GFP fusion protein, when expressed at high levels, was able to form prion-like foci in both [PSI+] and [psi−] cells with similar kinetics. This finding demonstrates that the CaSup35pNM region can direct self-aggregation in vivo, but its polymerization requires no interaction with the endogenous ScSup35p.

Given the similarity in amino acid composition between the PrD of ScSup35p and CaSup35p, and in particular the presence of related repeat sequences (with the ‘core’ motif of GGYQQ) in both, it is perhaps surprising that the CaSup35p is not seeded by ScSup35p. Santoso et al. (2000) demonstrated that replacement of a short region of the CaSup35pN domain with the 19 aa between residues 8 and 26 from the ScSup35pN domain generated a CaSup35pN–GFP fusion protein whose aggregation could now be seeded by the endogenous ScSup35p. Their data suggest that a region of the N domain of Sup35p N-terminal to, but excluding, the oligopeptide repeats, plays a key role in driving prion conversion of Sup35p. This region shows no amino acid sequence identity between the ScSup35p PrD and the CaSup35p PrD, and the corresponding region in the ScSup35p N domain has a significantly higher proportion of Gln and Asn residues than the corresponding CaSup35p region. These differences may be sufficient to inhibit the conformational conversion event even though the heterologous N domain may be able to physically interact via the conserved oligopeptide repeats (see below).

As suggested by the experiments of Chien & Weisemann (2001) the failure to efficiently recruit the newly synthesized CaSup35p molecules into the pre-existing ScSup35p aggregates in a [PSI+] strain is probably due to the different conformations these two proteins take up in the cell. More specifically, the underlying problem may be the different folding intermediates that predispose the nascent polypeptide chains to conformational conversion in vivo. Structure prediction suggests that both the ScSup35p and the CaSup35p N domains are highly flexible and unstructured regions. However, an important requirement for prion establishment and maintenance in S. cerevisiae is a dependence upon a certain ‘wild-type’ level of the molecular chaperone Hsp104 (Chernoff et al., 1995). Thus, in S. cerevisiae, the newly synthesized CaSup35p may be unable to access one or more of the components of the protein-folding machinery to which the ScSup35pN domain has been evolutionarily tuned.

One of the most intriguing features of the ScSup35p PrD is the presence of five copies of a 9 aa repeat with a conserved core motif of GGYQQ. This highly repeated core motif is also present in multiple copies in the N domain of the CaSup35p (Fig. 1b). Deletion analysis of the equivalent ScSup35p repeats has led us to propose that these repeats mediate the Sup35p–Sup35p interactions that precede the conformational conversion event that leads in turn to aggregate formation (Parham et al., 2001). Even deletion of one repeat, although not preventing interaction with the wild-type ScSup35p PrD, does prevent interaction with similarly deleted ScSup35p–PrD molecules. That the CaSup35p is unable to interact with the wild-type ScSup35p to form aggregates would suggest that multiple copies of the GGYQQ motif are not sufficient for the important intermolecular interactions required for prion propagation. However, it might be that the conformational conversion of the oligomeric intermediates (Serio et al., 2000) rather than the intermolecular interactions per se is the step that is blocked or at least less efficient. The transient intramolecular interactions between the CaSup35p PrD and the molecular chaperone Hsp104 that are essential for [PSI+] maintenance in S. cerevisiae (Chernoff et al., 1995) may therefore not be optimal for efficient propagation of a CaSup35p-based prion in the absence of the homologous folding factor.

**Polyglutamine repeat expansion in the N domain of CaSup35p?**

One feature of the CaSup35p that differentiates it from the ScSup35p sequence is the presence of several poly(Gln) repeat regions. Expansion of such poly(Gln) repeats has been identified with a number of inherited human diseases, in particular Huntington's disease, fragile X syndrome and myotonic dystrophy (reviewed by Zoghbi & Orr, 2000), and this is due to the underlying protein having an increased propensity to aggregate in the brain. Our finding that there is a significant level of polymorphism in the length of one (but not all) such tracts in the CaSup35p N domain (Fig. 5) raises the interesting possibility that CaSup35p aggregation may be modulated by a different mechanism to ScSup35p (which has no such poly(Gln) repeat regions in the PrD). It remains to be established whether or not changes in the Gln repeat number do alter the aggregation behaviour of CaSup35p in C. albicans. It should be noted that the subcellular fractionation analysis of strains carrying the different CaSUP35 alleles (Fig. 3) failed to detect any differences in the levels of solubility of the CaSup35p between the strains although the centrifugation steps we employ only pellet material in excess of 190S (F. Ness & M. F. Tuite, unpublished data). We also note that the variable poly(Gln) repeat region lies within the same region as implicated in the species barrier of yeast prion propagation (Santoso et al., 2000).

The mechanism by which poly(Gln) repeats expand is still far from clear although meiotic gene conversion and/or unequal crossing-over between two homologous chromosomes has been widely suggested (Berg et al., 2000; Jakupciak & Wells, 2000). It is important to note...
that C. albicans is apparently unable to undergo meiosis, lacking many of the genes essential for meiosis in S. cerevisiae (Tzung et al., 2001). This would suggest that the mechanism of poly(Gln) repeat expansion in C. albicans may occur via a non-meiotic event, either through unequal crossing-over between the two sister chromatids of a chromosome during mitosis or due to replication slippage (reviewed by Moore et al., 1999).

None of the strains we studied showed clear evidence of heteroallelism at the CaSUP35 locus since all DNA sequences of the PCR-generated fragments were homogeneous, although were polymorphisms to exist within the primer binding site(s), this might bias such an analysis in favour of one allele. The significance of this high-frequency variation in the length of the poly(Gln) repeat in CaSup35p may be in the rapid evolution of a directed mechanism of protein aggregation akin to the prion-like behaviour of the S. cerevisiae homologue. It now remains to be established whether any of the identified alleles of the CaSUP35 gene encode a form of CaSup35p that is able to behave as a prion in C. albicans.

ACKNOWLEDGEMENTS

This work was supported by Subprograma Ciência e Tecnologia do 2º Quadro Comunitário de Apoio (Portugal), by project grants from the Wellcome Trust, the EC and the BBSRC (M.F.T.) and by a BBSRC postgraduate studentship (S.N.P.). Dr Frederique Ness and Clare Lawrence are acknowledged for helpful discussions.

REFERENCES


Rose, M. D., Novick, P., Thomas, J. H., Bolstein, D. & Fink, G. R.
C. Resende and others


Received 12 October 2001; accepted 29 November 2001.