The Conserved and Non-conserved Regions of Rpb4 Are Involved in Multiple Phenotypes in *Saccharomyces cerevisiae*

Vinaya Sampath*, Nambudiry Rekha§, N. Srinivasan§, and Parag Sadhale‡**

*From the Department of Microbiology & Cell Biology and §Molecular Biophysics Unit, Indian Institute of Science, Bangalore 560012, India

Rpb4, the fourth largest subunit of RNA polymerase II in *Saccharomyces cerevisiae*, is required for many phenotypes, including growth at high and low temperatures, sporulation, pseudohyphal growth, activated transcription of a subset of genes, and efficient carbon and energy metabolism. We have used deletion analysis to delineate the domains of the protein involved in these multiple phenotypes. The scRpb4 protein is conserved at the N and C termini but possesses certain non-conserved regions in the central portion. Our deletion analysis and molecular modeling results show that the N- and C-terminal conserved regions of Rpb4 are involved in interaction with Rpb7, the Rpb4 interacting partner in the RNA polymerase II. We further show that the conserved N terminus is required for efficient activated transcription from the *INO1* promoter but not the *GAL10*- or the *HSE*-containing promoters. The N terminus is not required for any of the stress responses tested: growth at high temperatures, sporulation, and pseudohyphal growth. The conserved C-terminal 23 amino acids are not required for the role of Rpb4 in the pseudohyphal growth phenotype but might play a role in other stress responses and activated transcription. From the deletion analysis of the non-conserved regions, we report that they influence phenotypes involving both the N and C termini (interaction with Rpb7 and transcription from the *INO1* promoter) but not any of the stress-responsive phenotypes tested suggesting that they might be involved in maintaining the two conserved domains in an appropriate conformation for interaction with Rpb7 and other proteins. Taken together, our results allow us to assign phenotype-specific roles for the different conserved and non-conserved regions of Rpb4.

Transcriptional regulation by RNA polymerase II and its associated proteins lies at the core of eukaryotic gene expression. *Saccharomyces cerevisiae* RNA polymerase II is a complex of 12 subunits, Rpb1–12, named in the order of decreasing size. Rpb1, Rpb2, and Rpb3-Rpb11 are homologs of *Escherichia coli* RNA polymerase β’, β, and α subunits, respectively. Rpb5, 6, 8, 10, and 12 are shared by all three RNA polymerases I, II, and III. Rpb9 and Rpb4 are the two non-essential subunits in this complex (1). Rpb4 forms a sub-complex with Rpb7 that is easily dissociated from the polymerase under mild denaturing conditions and is involved in promoter dependent initiation of transcription (2). This sub-complex is associated sub-stoichiometrically with the rest of the polymerase under logarithmic growth phase, but under stationary phase conditions the stoichiometry increases to one (3, 4). These observations have led to the hypothesis that this sub-complex forms the eukaryotic counterpart of the bacterial σ subunit (2).

The Rpb4-Rpb7 sub-complex has been reported from many species (5–8). Recently Rpb4-Rpb7 like sub-complex has also been demonstrated in RNA polymerases I and III (9, 10). Rpb7 is very highly conserved across species (11, 12), and the lethality of *serpBΔ* can be rescued by overexpression of homologs from other species.1 Rpb4 is conserved from archaea to humans with 12 Rpb4-like sequences reported in the GenBank™ database (13). Till date, only the *Homo sapiens* (Hs), *Arabidopsis thaliana* (At), *Drosophila melanogaster* (Dm), *Schizosaccharomyces pombe* (Sp), and *Methanococcus jannaschii* (Mj) proteins have been shown to be bona fide Rpb4 homologs (5–8).

Rpb4Δ yeast strain grows at moderate temperatures, albeit slowly, and exhibits a variety of stress response defects (14). It cannot survive at temperatures above 32 °C or below 15 °C. It also shows poor recovery from stationary phase. Previous work from our group has shown that Rpb4Δ strains are also defective for sporulation and show enhanced pseudohyphal morphology, two hallmark responses to nutritional starvation (15). Transcriptional activity and global mRNA synthesis in *rpb4Δ* strains is significantly reduced on prolonged exposure to high temperatures (*in vitro* and *in vivo*) or stationary phase conditions (3, 16, 17). In addition, in certain genetic backgrounds *rpb4Δ* strains exhibit Na⁺/Li⁺ ion sensitivity (18). Some but not all of these phenotypes can be rescued by overexpression of Rpb7 suggesting that one of the roles of Rpb4 is to stabilize the interaction of Rpb7 with the rest of the polymerase (19–22).

Apart from these stress response defects, *rpb4Δ* also affects transcription of many genes even under moderate growth conditions. Using promoter-reporter studies, we have shown previously that *rpb4Δ* strains exhibit defects in activated transcription of a subset of genes. This defect is partially rescued by overexpression of cognate transcriptional activators (23). Whole genome transcriptional analysis of *rpb4Δ* strains reveals defects in carbon and energy metabolism at moderate temperatures and, additionally, the transcription of ribosomal protein genes in response to mild heat shock (15).

Because Rpb4 affects many different phenotypes, we hypo-

---

1 S. Singh, N. Rekha, B. Pillai, V. Singh, A. Naorem, V. Sampath, N. Srinivasan, and P. Sadhale, unpublished data.
Functional Analysis of Domains of Rpb4

TABLE I
List of plasmids used in this study

| Plasmid name | Alias | Description | Reference |
|--------------|-------|-------------|-----------|
| pPS2         | Ypplac111 | Yeast cloning vector, LEU2, CEN | (44) |
| pPS12        | pYES2 | Yeast cloning vector, P_GAL1, URA3, 2µ | Invitrogen |
| pPS24        | pPH509 | P_FUS1-LacZ, URA3, 2µ | (45) |
| pPS40        | pGAD424 | GAL4AD fusion vector, LEU2, 2µ | Invitrogen |
| pPS111       | pZHS51 | HSE element-P_CCV-LacZ fusion, URA3, 2µ | (46) |
| pPS121       | pGA8 | P_FUS1-LacZ, URA3, CEN | U. Vijayaraghavan |
| pPS141       | LexABDD-Rpb7 | RBP7 ORF in frame with LexABDD, HIS3, 2µ | (12) |
| pPS143       | pJK103 | LexAop2-LacZ, URA3, 2µ | (12) |
| pPS21       | pGAD424A RI | pPS31-EcoRI site, LEU2, 2µ | This study |

Deletions of Rpb4 in the gene context

| Plasmid name | Description | Reference |
|--------------|-------------|-----------|
| pPS114      | N- and C-terminal truncation of Rpb4 | This study |
| pPS212      | C-terminal 23 aa truncated Rpb4 | (23) |
| pVS378      | N-terminal 32 aa truncated Rpb4 | This study |
| pVS379      | N- and C-terminal truncated Rpb4 | This study |
| pVS366      | Deletion of acidic stretch in Rpb4 | This study |
| pVS386      | Deletion of basic stretch in Rpb4 | This study |
| pVS70       | Deletion of acidic and basic stretch in Rpb4 | This study |
| pVS84       | Deletion of 106–140aa in Rpb4 | This study |
| pVS85       | Deletion of entire non-conserved region in Rpb4 | This study |

Deletions of Rpb4 as fusions to GALA4D in pPS31 vector, LEU2, 2µ

| Plasmid name | Description | Reference |
|--------------|-------------|-----------|
| pVS153      | Full-length Rpb4 gene in pPS2 | (19) |
| pVS288      | C-terminal 23 aa truncation of Rpb4 | This study |
| pRS217      | N-terminal 32 aa truncated Rpb4 | This study |
| pVS366      | N- and C-terminal truncated Rpb4 | This study |
| pVS732      | Deletion of acidic stretch in Rpb4 | This study |
| pVS737      | Deletion of basic stretch in Rpb4 | This study |
| pVS837      | Deletion of acidic and basic stretch in Rpb4 | This study |

Deletions of Rpb4 in upstream sequence

| Plasmid name | Description | Reference |
|--------------|-------------|-----------|
| pVS153      | Full-length Rpb4 ORF in pPS31 | This study |
| pPS212      | C-terminal 23 aa truncation of Rpb4 | This study |

Experimental Procedures

Strains and Growth Conditions

The yeast strains used were (i) SY10 (Mat a, his3Δ200, ura3Δ–52, leu2Δ–3,112, lys2–1,128, rpb4Δ::HIS3); (ii) SYD1011 (Mat a/s, his3Δ200/ his3Δ200, ura3Δ–52/ura3Δ–52, leu2–3,112/leu2–3,112, lys2–1/lys2–1, rpb4::HIS3/rpb4Δ::HIS3); (iii) SY9 (EGY191, Mat a, his3Δ200, trp1–901, ura3Δ–52, leu2::LEU2-LexAop2) (12). These strains were transformed with the appropriate plasmids listed in Table I and assayed for various phenotypes. Yeast strains were grown in YPD (1% yeast extract, 2% peptone, 2% dextrose) or in Synthetic Dropout (SD) medium and manipulated as per standard protocols (24). Yeast transformations were performed using the modified lithium acetate protocol, which does not involve heat treatment of cells (25). All manipulations of plasmids were performed in Escherichia coli strain DH5α (supE44, deltalacI169 (phi 80 lacZ delta M15), hsdR17, recA1, endA1, gyrA96, thi–1, relA1) as per standard protocols (24).

Construction of Plasmids

Mutants of Rpb4 were expressed in their gene context for assays of temperature sensitivity, sporulation, pseudohyphal growth, and activated transcription in rpb4Δ strains. For two-hybrid interaction analysis, the various deletions of the Rpb4 open reading frame (ORF) were expressed as fusions to Gal4 activation domain (GALA4D) in the plasmid pPS31 (pGAD424, Invitrogen).

RPB4 Mutants in Gene Context—The full-length Rpb4 protein is 221 aa long. Construction of the C-terminal 23 aa deletion (Rpb4-(1–198), pBP212) has been detailed elsewhere (23). The Rpb4 upstream sequence encompassing the promoter was amplified using the primer pPS40 (5′-GGGATCCATTTTTCTATATTC-3′ and pRP4 RIdelr (5′-CAGCGTTGACTACTGCCGTAGTTTC-3′) which was cloned in pGEM-Teasy (Promega). The EcoRI fragment from this construct was used to replace the EcoRI fragment in pPS114 (Rpb4, 1–221 aa) and in pBP212 to generate the N-terminal 32-aa deletion (Rpb4-(32–221), pPS378) and the N- and C-terminal deletion (Rpb4-(33–198), pVS85), respectively. The internal deletions of Rpb4 were generated by overlap extension PCR method using deletion-specific primers and the Rpb4 upstream and downstream primers. Rpb4PS2 and Rpb4PS1r (5′-GGCGTCAGCT- CACTCCACGGTTTGACACC-3′) were expressed in their gene context for assays of temperature sensitivity, sporulation, pseudohyphal growth, and activated transcription in rpb4Δ strains. For two-hybrid interaction analysis, the various deletions of the Rpb4 open reading frame (ORF) were expressed as fusions to Gal4 activation domain (GALA4D) in the plasmid pPS31 (pGAD424, Invitrogen).
cloned into pPS12 and further sub-cloned as a BamHI fragment in pVS151 to generate the C-terminal 23-aa deletion of Rpb4 (pVS238). The EcoRI-Hind fragment of pPS40 was cloned in pPS31 to generate the N-terminal deletion (Rpb4–33–221), pRS217. The EcoRI-Sall fragment of pBP212 was cloned in-frame in pPS1 to generate the N- and C-terminal deletion (Rpb4–33–198), pVS356. Overlap Extension PCR using the set of deletion-specific primers listed above, primers that bind to the ORF after looking at the residue conservation at the solvent buried sequence from an archaeal homolog. The alignment accuracy was used to generate the RPB4 internal deletions. These PCR products were cloned into pGEM-Teasy. The RPB4 deletion ORFs were sub-cloned using the sites flanking the ORF into pVS151. Due to the cloning strategies, the C-terminal truncation in the gene context is shorter than the truncation in the ORF context by 1 residue at the 198th position. However, to avoid confusion, both these constructs have been labeled Rpb4–(1–198).

Temperature Sensitivity, Sporulation, and Pseudohyphal Growth

The assays for temperature sensitivity, sporulation, and pseudohyphal growth using rpba4 strains containing plasmids expressing either RPB4 deletion mutants or appropriate controls were done essentially as described earlier (15, 23). Assays for temperature sensitivity were done on SD medium plus 2% dextrose plates at 25 °C and 37 °C. Assays for sporulation and pseudohyphal growth were performed at 25 °C on 1% potassium acetate plates and synthetic low ammonia dextrose (SLAD) plates, respectively.

Activated Transcription

The ability of haploid rpba4 strains containing RPB4 deletion plasmids to activate transcription from the GAL10, INO1, and HSE promoters were assayed using promoter-LacZ reporters as earlier described (23). β-Galactosidase assays were performed by glass bead disruption method essentially as described previously (26).

Two-hybrid Analysis

Two-hybrid analysis of interaction between LexA DNA binding domain (DBD)-Rpb7 fusion and the GALAD fusions of Rpb4 and deletion mutants was performed in strain Y19 (EGY191) transformed with LexAop2–LacZ reporter plasmid (JK103) (12). This particular combination of plasmids was preferred as the extended multiple cloning sites in the pPS31 plasmid allowed the construction of various GALAD-Rpb4 fusions. However, it precludes the use of the LexA reporter integrated in EGY191. Strains to be assayed for interaction were grown in SD plus 2% galactose plus 1% xylose plus tryptophan until mid log phase and harvested for β-galactosidase assays. β-Galactosidase assays were performed by chloroform permeabilization method as described previously (27).

Comparative Modeling of S. cerevisiae Rpb4p Using an Archaeal Homolog

Although, during the course of this work, structures of the multisubunit complex of RNA polymerase from S. cerevisiae has been published (28, 29), the structures are of low resolution (4.1 and 4.2 Å) and for Rpb7 and Rpb4 the positions of Ca atoms alone are available. Hence detailed analysis of the side-chain interactions between Rpb4 and Rpb7 is not possible. So we generated models of Rpb4 and Rpb7 using archaeal complex structure, and we ensured that the features of the model built, such as similarity of the positioning of secondary structures alone are consistent with the features observed in the low resolution structures of yeast complex.

The structure of Rpo_E1 and Rpo_F1 proteins from M. jannaschii is available in the Protein Data Bank (30). We have used this structure as a template to model Rpb4 and Rpb4 from S. cerevisiae, respectively. The sequence identity between archaeal Rpb4 and S. cerevisiae Rpb4 is about 21%. The Rpb4 proteins are significantly different, with iterative position-specific iterative blast failing to pick up the other as homologs about 21%. The Rpb4 proteins are significantly different, with iterative position-specific iterative blast failing to pick up the other as homologs about 21%.

The model generated for the heterodimeric interaction between Rpb4 and Rpb7 suggests that these residues are important for the interaction. Western blotting of protein extracts from these strains with anti-Gal4 antibodies showed that, although the N-terminal truncation does not affect the stability of the protein, the C-terminal truncation reduces protein levels below detectable limits under conditions of our assay (data not shown). However, as shown below (Fig. 3C) this destabilization does not completely eliminate the functioning of the protein. The two-hybrid result suggests a role for the conserved N-terminal region in interaction with Rpb7. A role for the C-terminal Rpb4 residues in interaction with Rpb7 cannot be predicted based on these results alone.

Molecular Modeling of Rpb4-Rpb7 Complex Predicts a Role for the Rpb4 N and C Termini in the Interaction with Rpb7—To understand the structural basis of this interaction, we generated a molecular model of the scRpb4-Rpb7 complex based on the crystal structure of archaeological homologs (30). The model for scRpb7 and its homologs from other lower eukaryotic homologs shows very high structural conservation with the archaeal homolog and has been detailed elsewhere. The C-terminal region of scRpb4 shows high sequence similarity with the homologous mJF protein. The N-terminal region of Rpb4 is poorly conserved between these homologs (as against the homology among eukaryotic Rpb4 proteins). The alignment arrived at by consulting the best match of the residues in the region among all the known homologs of Rpb4 was used for model building. The model generated for the heterodimeric interaction between Rpb4 and Rpb7 for S. cerevisiae proteins is shown in Fig. 2B. The model for scRpb4 is shown in blue, and the regions interacting with Rpb7 are colored orange. The non-conserved re-
**FIG. 1.** Sequence features of Rpb4. A, alignment of Rpb4 sequences from eukaryotes. The protein sequences of Rpb4 homologs from *S. cerevisiae* (S.c), *Schizosaccharomyces pombe* (S.p), *Homo sapiens* (H.s), *Arabidopsis thaliana* (A.t), and *Drosophila melanogaster* (D.m) were aligned using ClustalW and modified for print representation using ESPript 2.0 (43). The residues outlined in black boxes are identical across all homologs, whereas the residues outlined in white boxes are similar across all homologs. B, deletions generated in this study. In the first set of deletions, the conserved regions (denoted by black-filled regions) were truncated either at the N or the C terminus or both simultaneously keeping the rest of the protein intact. In another set of deletions, the non-conserved regions were deleted singly (the Acidic stretch is denoted by diagonal lines, the Basic stretch is denoted by waves, and the aa 106–140 is denoted by crossed hatches) or in combination.
Functional Analysis of Domains of Rpb4

The conserved N- and C-terminal aa of Rpb4 are involved in interaction with Rpb7. A, two-hybrid analysis of Lex-A DBD-Rpb7 interaction with GAL4AD Vector, Rpb4 (1-221), and the indicated conserved region truncations. Plotted here are the average β-galactosidase activities and standard deviations thereof from two independent experiments with three different transformants each. B, Molecular Model of scRpb4-Rpb7 generated based on the crystal structure of mjF-mjE homologs. The Rpb4 model is represented in blue with the Rpb7 interaction regions colored orange. The six helices in the Rpb4 structure have been numbered H1-H6. The residues not modeled due to lack of similarity are represented as dotted lines. The Rpb7 model is represented in gray with the regions involved in interaction with Rpb4 in green. The Rpb7 interaction regions on Rpb4 map to the N and C termini.

S. cerevisiae undergoes sporulation to form tetrads (four haploid spores enclosed in an ascus) in response to severe starvation usually involving lack of nitrogen and fermentable carbon source. We analyzed the ability of the Rpb4 truncations to rescue the sporulation defect of rpb4Δ strains as compared with the wild-type protein. This was measured as percentage of tetrads in each population after 3 days in sporulation medium. As can be seen from Fig. 3B, the N-terminal truncation (1-198) shows similar results as compared with the wild-type protein. The 33-198 mutant proteins are defective for sporation as is the rpb4Δ strain.

In the presence of a fermentable carbon source and limiting amounts of a nitrogen source, rpb4Δ strains show pronounced cell elongation, unipolar budding patterns, and irregular colony morphology associated with pseudohyphal growth phenotype.
The conserved C-terminal aa but not the N-terminal aa of Rpb4 are required for stress response. A, haploid rpb4Δ strains with vector (rpb4Δ), full-length Rpb4 (1–221), or the deletions of conserved regions were assayed for temperature sensitivity. Photographs were taken after 3-day incubation at permissive (25°C) and non-permissive (37°C) temperatures. B, diploid rpb4Δ strains with vector (rpb4Δ), full-length Rpb4 (1–221), or the deletions of conserved regions were assayed for pseudohyphal growth on SLAD (low nitrogen-containing SD media) plates. Shown here are representative colony morphologies for the strains 18 h after spotting.

Some Activated Promoters Require Both the Conserved Regions of Rpb4 for Effective Transcription—Apart from its stress-specific roles, Rpb4 is required for transcription from many activated promoters like the PGAL10-, PINO1-, and HSE-containing promoters (Fig. 4, A and B) but reduces transcription significantly from the PINO1 promoter (Fig. 4C). The activity of the N-terminal deletion is nearly 70–80% of the full-length protein activity for the GAL10- and HSE-containing promoters but less than 20% for the INO1 promoter. These results suggest that the INO1 promoter requires both the N- and C-terminal regions of Rpb4 for effective transcription, whereas transcription from PGAL10- and HSE-containing promoters depends on only the C-terminal conserved regions of Rpb4.

Deletion of the Non-conserved Residues in Rpb4 Affects Interaction with Rpb7—scRpb4 has certain unique stretches in its primary sequence that are not conserved among the different homologs. After analyzing the requirement of the conserved regions of Rpb4 in various phenotypes, we wanted to test if the non-conserved (aa 66–80 and 107–140), and weakly conserved 95–105 residues play any role in these phenotypes. Toward this aim we generated a series of deletions in Rpb4 that lack these stretches singly or in combination (Fig. 1B) and expressed them as described for the conserved region deletions above.

We used two-hybrid analysis to analyze the effect of deletion of non-conserved regions on interaction of Rpb4 with Rpb7. Analysis of the interaction between Rpb4 mutant-GAL4AD and the LexA DBD-Rpb7 shows that loss of any one of the non-conserved regions of Rpb4 affects interaction with Rpb7 (Fig. 1C).
To rule out that the differences observed did not arise from differential stability of these mutant fusion proteins, we ascertained their steady-state levels using anti-Gal4 antibodies. We find that all the mutant fusion proteins are expressed in this strain background as expected (data not shown). This suggests that these non-conserved regions are also involved in interaction with Rpb7 either directly or indirectly (also see “Discussion”).

**The Non-conserved Regions Are Not Involved in Stress-specific Roles of Rpb4**—We next determined the requirement of the non-conserved regions of Rpb4 in rescue of temperature sensitivity, sporulation defect, and predisposition to pseudohyphal growth of rpb4Δ strains. These deletions were made in the context of the Rpb4 gene and transformed into rpb4Δ haploid and diploid strains along with the appropriate controls. Deletion of most of the non-conserved regions does not affect the ability of Rpb4 to support growth of haploid (Fig. 6A) or diploid rpb4Δ strains (data not shown) at higher temperatures of 37°C. Only the deletion of the entire Sc-specific region (from aa 66–140) moderately compromises growth at higher temperatures. This effect is seen only in haploid and not in diploid rpb4Δ strains (data not shown). We analyzed the ability of these deletions to support sporulation in diploid rpb4Δ strains in comparison to full-length Rpb4. As can be seen from Fig. 6B, deletion of none of the non-conserved regions affects the ability of Rpb4 to rescue the sporulation defect of rpb4Δ strains. To assay for their ability to reverse the pseudohyphal growth predisposition of rpb4Δ strains, these mutants were spotted on pseudohyphae-inducing SLAD plates. The difference in morphology assayed after 18 h shows that (Fig. 6C) none of the internal deletions of RPB4 can be distinguished from wild-type RPB4 suggesting that these regions are not involved in this phenotype. Taken together, the results in Fig. 6 suggest that the non-conserved regions of Rpb4 do not play a significant role in stress response.

**Transcription from Some Activated Promoters Requires the Non-conserved Regions of Rpb4**—We also analyzed the requirement for these non-conserved regions of Rpb4 in activated transcription from the P_GAL10-, P_INO1-, and HSE-containing promoters as explained above. As can be seen from Fig. 7 (A and B), deletion of the non-conserved regions either singly or in combination does not affect activated transcription from the P_GAL10- and HSE-containing promoters (activity for all deletions nearly 80–100% of the full-length activity). However, activated transcription from P_INO1 is affected by deletion of the non-conserved regions (Fig. 7C). The deletion of both the basic and acidic regions (aa 66–105) and the deletion of the entire Sc-specific region (from 66 to 140) have more pronounced effects (activity is reduced to ~10% of full-length protein) on activated transcription than deletion of individual stretches (activity is between 20 and 40% of full-length protein). This suggests that the non-conserved regions of Rpb4 are involved in activated transcription from the P_INO1 promoter but not the P_GAL10- and HSE-containing promoters.

**DISCUSSION**

Regulation of gene expression has been thought to involve contact between the DNA-bound activators and the Holoenzyme components (co-activators and general transcription factors) leading to increased recruitment of the polymerase (1). Modulation of RNA polymerase II composition could be an additional level of regulation in gene expression (37). The Rpb4-Rpb7 sub-complex is an ideal candidate for such a regulation, because the association of the complex with the rest of the polymerase is sub-stoichiometric under logarithmic growth conditions and becomes stoichiometric under stationary phase conditions (2). This sub-complex is also easily dissociated from the rest of the polymerase under mild denaturing conditions and is absent from polymerase purified from a mutant of Rpb1, Rpb6, and from rpb4Δ strains (3, 38, 39). Rpb4 plays a role in many stress-responsive phenotypes like growth at high and low temperature, sporulation, pseudohyphal growth, etc. (14, 15). It is also involved in activated transcription from a subset of promoters, in carbon and energy metabolism at moderate temperatures (15, 21, 23). Delineation of the roles played by this subunit in multiple pathways requires an understanding of its domainal organization.
A comparison of the Rpb4 primary sequence of the eukaryotic homologs reveals that the N-terminal 45 aa and the C-terminal 80 aa of the proteins can be denoted as “conserved regions” (Fig. 1A). However, the scRpb4 has unique stretches in the central region (aa 66–140) that have interesting features like a high density of basic (aa 66–80) or acidic (aa 95–105) residues. Our analysis of various deletion mutants in these regions (Fig. 1B) was aimed at understanding their roles in controlling the multiple phenotypes observed in rpb4Δ strain.

We have shown that the Rpb7 interaction domain in Rpb4 involves the conserved N-terminal 32 aa, because deletions of this domain abolishes two-hybrid interaction with Rpb7 (Fig.
2A). This observation is validated by molecular modeling of the
scRpb4•Rpb7 complex (Fig. 2B) based on the crystal structure
of mF•mE subunit complex, which reveals that the scRpb4
N-terminal and C-terminal regions are involved in the interac-
tion. The N-terminal region forms a β-α-β-β′-α–helix loop with
the conserved three-stranded β-sheet of Rpb7. The C-terminal re-
gions fold into α-helices that are involved in contact with Rpb7.
The C-terminal 23 aa maps onto a helix after a region that has
conserved surface-exposed residues and residues involved in
binding Rpb7 in the heterodimer. The truncation of the C-
terminal 23 residues could affect the folding of the regions
preceding it (156–185), thereby affecting the binding to Rpb7.
This region of the Rpb4 structure has structural similarity to
the HRDC domain of the RecQ helicase (30). Hence, this region
could be a structural domain such that any truncations in this
region might result in misfolding and thereby compromise
binding to Rpb7.

The role for the conserved N- and C-terminal regions of Rpb4
in interaction with Rpb7 is also validated by other observations
made previously: (i) Rpb4 and Rpb7 from many different sys-
tems interact with each other in vivo suggesting that a con-
served region is used for interaction. (ii) The spRpb4•Rpb7 and
AtRpb4/N25.9/19.5 homologs form hybrid dimers with scRpb4•Rpb7
in vitro (6, 7) suggesting that Rpb4•Rpb7 interaction is conserved
among these species. (iii) Using two-hybrid analysis, the
hsRpb7 interaction region in hsRpb4 has been mapped to the
N-terminal 1–92 aa (5). Although this region encompasses
some of the conserved regions, the C-terminal 23 aa conserved
region is not involved in this interaction. This suggests that the
hsRpb4 and scRpb4 have evolved differently to interact with
their respective Rpb7 proteins. In keeping with this, the
hsRpb7 protein interacts very weakly with scRpb4 (5).

The observation that deletion of any of the non-conserved
regions in Rpb4 also abrogates interaction with Rpb7 (Fig. 5)
throws up interesting possibilities. The simplest explanation
could be that these regions are also involved in the interaction.
However, this possibility will have to be discounted given that
Rpb4•Rpb7 interaction is conserved from archaea to humans
and none of the homologues have these regions. The other pos-
bility is that these regions are involved in maintaining the
conserved N- and C-terminal region helices, Helix 1 and Helix
2 (Fig. 2B), in an appropriate conformation for interaction with
Rpb7. The exact structural fold adopted by these non-conserved
regions is not yet known. Based on the 12-subunit crystal
structures (28, 29) and the model building of scRpb4•Rpb7
interaction, it can be predicted that the basic stretch immedi-
ately following Helix 1 extends the Helix 1 further away from
Helix 2. The other non-conserved regions probably adopt a fold
that will bridge Helix 1 to Helix 2. Therefore, any deletion in
the non-conserved region of the protein will affect the juxtapo-
sition of the helices 1 and 2 of the conserved regions and hence
the interaction between Rpb4 and Rpb7. To further understand
the roles played by the non-conserved regions of Rpb4 in inter-
action with Rpb7, point mutants of Rpb4 defective for interac-
tion with Rpb7 need to be isolated and characterized. A genetic
circuit to isolate these mutants is underway.

We have earlier reported that the conserved C-terminal 23
aa of Rpb4 are required for growth at high temperatures and
efficient sporulation (15). Our analysis of deletions of the con-
served and non-conserved regions of Rpb4 allows us to predict
that the deletion of only the C-terminal conserved region af-
ccts these phenotypes. Lack of the N-terminal region, or any of
the non-conserved regions, does not affect the ability of Rpb4
to function as well as the full-length protein in these phenotypes
(Figs. 3A, 3B, 6A, and 6B). It is to be noted here that the
deletion of the entire Sc-specific region (A66–140) has a mar-
ginal effect on the growth of rpb4Δ strains at 37 °C but not on
sporulation. These observations are in keeping with the fact
that spRpb4 and hsRpb4 complement the temperature sensi-
tivity of the rpb4Δ strain (5, 6). The C-terminal 23 aa of scRpb4
is more conserved in spRpb4 than the hsRpb4, possibly explain-
ing the robust growth of rpb4Δ with spRpb4 as compared with
hsRpb4 (11, 12). Our analysis for the region of Rpb4 involved in
inhibiting pseudohyphal growth of rpb4Δ strains did not pro-
vide any conclusive answers. Deletions of any of the conserved
and non-conserved regions do not affect the ability of Rpb4 to
rescue pseudohyphal growth of rpb4Δ strains suggesting that
these regions are not involved in this phenotype (Figs. 3C and
6C). As these deletions encompass most of the protein, we
surmise that the regions not deleted in this process (aa 141–198)
are the regions involved in this phenotype. This could
possibly be involved in interactions with novel proteins in-
volved in inhibiting pseudohyphal growth.

Many groups, including ours, have shown that overexpr-
ession of Rpb7 can rescue the temperature sensitivity associated
with rpb4Δ suggesting that one of the roles of Rpb4 is to
stabilize interaction of Rpb7 with RNA polymerase II (19, 20,
22). It has also been hypothesized that the temperature sensi-
tivity of rpb4Δ is due to the weakened interaction of Rpb7 with
the rest of the polymerase. The fact that the N-terminal dele-
tion of Rpb4 is compromised for interaction with Rpb7 but is
still able to rescue temperature sensitivity of rpb4Δ (Figs. 2A
and 3A) suggests that Rpb4 and Rpb7 have independent abil-
ities to rescue temperature sensitivity of rpb4Δ.

We find a similar pattern of results for domains of Rpb4
involved in activated transcription from the Pgal10- and HSE-
containing promoters as for the temperature sensitivity and
sporulation. The C-terminal 23 aa might play a role in all these
phenotypes. The N-terminal 32 aa deletion retains ~70–80% of
full-length activity from the Pgal10- and HSE-containing
promoters, whereas the internal deletions are indistinguish-
able from the full-length protein (Figs. 4A and B). On the other hand, transcriptional activation from the
ino1 promoter requires both the N- and the C-terminal re-
gions. Deletion of the N-terminal region of Rpb4 or the internal
deletions of the non-conserved regions lowers activity from the
Pino1 promoter. Deletion of the entire Sc-specific region (aa
66–140) or deletion of the aa 66–105 has more pronounced
effects on the activity than deletions of individual stretches
(Figs. 4C and 7C). As the case with the domains involved in
interaction with Rpb7, it is not possible to differentiate at this
point of time, whether these non-conserved regions are actually
required for the activity from the ino1 promoter or their
deletion affects the conformation of the N- and the C-terminal
regions that are essential for this phenotype. The difference in
the requirements for activation from the gal10- and HSE-
containing promoters and the ino1 promoter could reflect differen-
tial interaction of Rpb4 with other proteins involved in
transcription from these promoters. It is also worthwhile to
note that the ino1 promoter is known to be extremely sensi-
tive to perturbations in RNA polymerase II activity. Reduction
of the levels of Rpb1 or certain mutants of Rpb1 affects activity
of the ino1 promoter (40).

We have not assessed the protein levels of the mutants when
expressed from their own promoters in CEN plasmids, because
they are below the detection limits of the currently available
anti-Rpb4 antibodies. The C-terminal truncation expressed
from the Rpb4 promoter or as the fusion to GALA4AD rescues
the pseudohyphal predisposition of rpb4Δ strains (Fig. 3C and
data not shown). This suggests that, although the truncation
affects protein stability (as assessed by the anti-Gal4AD West-
ern blots) the reduced protein levels are sufficient to rescue this
particular phenotype. The finding that the deletion of the C-terminal region affects all the other phenotypes can be explained either by an absolute role for this region or the requirement for a more stable protein in the other phenotypes. Because the deletions of the conserved N terminus or all of the non-conserved regions do not affect any of the stress-responsive functions of Rpb4, it is possible to surmise that the C-terminal conserved regions of the protein (aa 140–221) are involved in these phenotypes.

Contrary to our expectations, we do not find any specific requirement for the non-conserved regions in S. cerevisiae-specific phenotypes. One explanation for their presence specifically in scRpb4 could be that they help stabilize a weaker conservation of scRpb4 in the N and the C termini. Analysis of the conserved regions between scRpb4 and the other homologs did not suggest any significant difference in the conserved region in the scRpb4 that could be stabilized by these non-conserved regions (data not shown). The other explanation could be that this is evolutionary “baggage” maintained by scRpb4, because it does not affect its function. However, the more complex systems have not maintained these extra sequences. Analysis of the pattern of conservation among closely related species would allow us to predict the evolutionary history of these sequences. Availability of the homologous sequence from Kluveromyces lactis (41) and 5 sequences from closely related Saccharomyces species (42) allowed us to compare the evolution of these regions. Rpb4 sequences from the evolutionarily distant K. lactis (diverged ~150 million years ago) and the Saccharomyces sensu lacto group (S. castellii) and petite negative group (S. kluvyeri) show very high sequence similarity with S. cerevisiae in aa 112–140 region not conserved in higher eukaryotes (Fig. 8, compare with Fig. 1A). They also show a significant sequence similarity in the acidic stretch (aa 95–105). These sequences however lack nearly two-thirds of the “Basic Stretch.” Interestingly, S. castellii sequence has a 24-aa unique insertion in between the acidic stretch. Rpb4 sequences from the Saccharomyces sensu stricto group (S. bayanus, S. mikatae, and S. kudriavzevi) are almost identical to S. cerevisiae sequence throughout as expected from sequences that have diverged only a few million years ago. This analysis shows that there has been a progressive loss of the non-conserved regions, because the organisms have diverged with higher eukaryotes retaining only the functionally important N- and C-terminal regions.

It is interesting to note that, unlike in S. cerevisiae, spRpb4 is an essential protein and in S. pombe and higher eukaryotes, the sub-complex does not dissociate easily from the rest of the polymerase (6, 7). It is not yet clear whether the non-conserved regions of scRpb4 promote easier dissociation from the rest of the polymerase and, therefore, regulate activity under certain conditions. It is possible that, in the absence of the non-conserved regions, the higher eukaryotic Rpb4s have evolved a
stronger interaction with the rest of the polymerase and associated proteins and hence are important for growth and survival under all conditions.

In conclusion, our deletion analysis has allowed us to demarcate the roles for the conserved N- and C-terminal regions. Both these regions are important for interaction with Rpb7 and for transcription from the INO1 promoter. The C-terminal conserved region may be involved in the stress-responsive role of Rpb4 and in activated transcription from a subset of genes. The non-conserved regions do not seem to be involved directly in any of these phenotypes. In fact, these regions seem to be required to maintain the conserved N- and C-terminal regions in an appropriate conformation, and their deletion affects only those phenotypes (interaction with Rpb7 and activated transcription from the INO1 promoter) that require both the N- and the C-terminal conserved regions.

Acknowledgments—We thank Drs. E. Golemis, L. Nover, and U. Vijayaraghavan for the gift of various strains and plasmids and J. Bhat for the gift of anti-Gal4 antibodies used in this study. We are grateful to Dr. P. Cliften and Dr. M. Johnston and the Washington University Genome Sequencing Center for sharing the sequences of related Saccharomyces species prior to publication. P. S. and N. S. thank all members of their respective laboratories for discussions and helpful suggestions.

REFERENCES
1. Hampsey, M. (1998) Microbiol. Mol. Biol. Rev. 62, 465–503
2. Edwards, A. M., Kane, C. M., Young, R. A., and Kornberg, R. D. (1991) J. Biol. Chem. 266, 71–78
3. Chodor, M., and Young, R. A. (1993) Mol. Cell. Biol. 13, 6984–6991
4. Kolodziej, P. A., Woychik, N., Liao, S. M., and Young, R. A. (1999) Mol. Cell. Biol. 19, 1915–1920
5. Khazak, V., Estojak, J., Cho, H., Majors, J., Sonoda, G., Testa, J. R., and Golemis, E. A. (1999) Mol. Cell. Biol. 19, 7511–7518
6. Larkin, R. M., and Guilfoyle, T. J. (1998) J. Biol. Chem. 273, 5631–5637
7. Werner, F., Eloranta, J. J., and Weinzierl, R. O. (2000) Nucleic Acids Res. 28, 4299–4305
8. Siaut, M., Ziros, C., Levivier, A., Ferri, M. L., Court, M., Werner, M., Callebaut, I., Thuriaux, P., Sentenac, A., and Conesa, C. (2003) Mol. Cell. Biol. 23, 195–205
9. Peyroche, G., Levillain, E., Siaut, M., Callebaut, I., Schultz, P., Sentenac, A., Riva, M., and Carles, C. (2002) Proc. Natl. Acad. Sci. U.S.A. 99, 14670–14675
10. Shipkovsky, G. V., Gadal, O., Labarre-Mariotte, S., Lebedenko, E. N., Miklos, I., Sakurai, H., Proshkin, S. A., Van Mullim, V., Ishihama, A., and Thuriaux, P. (2000) J. Mol. Biol. 295, 1119–1127
11. Khazak, V., Sadhale, P. P., Woychik, N. A., Brent, R., and Golemis, E. A. (1995) Mol. Biol. Cell 6, 769–775
12. Benson, D. A., Karchi-Mizrachi, L., Lipman, D. J., Ostell, J., and Wheeler, D. L. (2003) Nucleic Acids Res. 31, 23–27
13. Woychik, N. A., and Young, R. A. (1989) Mol. Cell. Biol. 9, 2854–2859
14. Pillai, B., Verma, J., Abraham, A., Francis, P., Kumar, Y., Tata, U., Brahmacchari, S. K., and Sadhale, P. P. (2003) J. Biol. Chem. 278, 3339–3346
15. Rosenheck, S., and Choder, M. (1998) J. Bacteriol. 180, 6187–6192
16. Miyao, T., Barnett, J. D., and Woychik, N. A. (2001) J. Biol. Chem. 276, 5049–5054
17. Bourbonnais, Y., Faucher, N., Pallotta, D., and Larouche, C. (2001) Mol. Gen. Genet. 264, 763–772
18. Shpakovski, G. V., Ryabov, V., Sadhale, P., and Johnston, M. (1999) Genet. 78, 149–156
19. Sheffer, A., Varon, M., and Choder, M. (1999) Mol. Cell. Biol. 19, 2672–2680
20. Tan, Q., Li, X., Sadhale, P. P., Miyao, T., and Woychik, N. A. (2000) Mol. Cell. Biol. 20, 8124–8133
21. Maillet, I., Buhler, J. M., Sentenac, A., and Labarre, J. (1999) J. Biol. Chem. 274, 22586–22590
22. Pillai, B., Sampath, V., Sharma, N., and Sadhale, P. (2001) J. Biol. Chem. 276, 19067–19074
23. Austub, F. M., Brent, R., Kingston, R., Moore, D., Seidman, J. M., and Struhl, K. (1987) in Methods in Enzymology, pp. 13.9–13.13, ed., J. Bacteriol. 195, 2411–2417
24. Sherman, F., Egan, G. R., and Lawrence, C. W. (1983) Methods in Cell Biology, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY
25. Rose, M., and Botstein, D. (1993) Methods Enzymol. 101, 167–180
26. Guarente, L. (1983) Methods Enzymol. 101, 181–191
27. Armache, K. J., Tenetkerhagen, H., and Cramer, P. (2003) Proc. Natl. Acad. Sci. U.S.A. 100, 6964–6968
28. Bushnell, D. A., and Kornberg, R. D. (2003) Proc. Natl. Acad. Sci. U.S.A. 100, 6969–6973
29. Sadhale, P., Sharma, N., and Sadhale, P. (2001) Mol. Biol. Cell 12, 1137–1143
30. Thompson, J. D., Higgins, D. G., and Gibson, T. J. (1994) Nucleic Acids Res. 22, 4673–4680
31. Bateman, A., Birney, E., Cerruti, L., Durbin, R., Etwiller, L., Eddy, S. R., Griffiths-Jones, S., Howe, K. L., Marshall, M., and Sonnhammer, E. L. (2002) Nucleic Acids Res. 30, 276–280
32. Blundell, T., Carney, D., Gartner, S., Hayes, F., Howlin, B., Hubbard, T., Overington, J., Singh, D. A., Shibanda, B. L., and Sutcliffe, M. J. (1998) Enzyme. 172, 513–520
33. Srinivasan, N., and Blundell, T. L. (1993) Protein Eng. 6, 501–512
34. Khazak, V., and Srinivasan, N. (2003) BMC. Struct. Biol. 3, 4
35. Berman, H. M., Westbrook, J., Feng, Z., Gilliland, G., Bhat, N. T., Weissig, H., Shindyalov, I. N., and Bourne, P. E. (2000) Nucleic Acids Res. 28, 235–242
36. Sadhale, P., Sharma, N., Pillai, B., Katoch, A., Acharya, N., and Singh, S. K. (1998) J. Biochem. 737–740
37. Ruet, A., Sentenac, A., Fromageot, P., Winsor, B., and Lacroute, F. (1980) J. Biol. Chem. 255, 6450–6455
38. Tan, Q., Prysak, M. H., and Woychik, N. A. (2003) Mol. Cell. Biol. 23, 3329–3338
39. Archambault, J., Jansma, D. B., and Friesen, J. D. (1996) Genetics 142, 737–747
40. Bolotin-Fukuhara, M., Teffano-Nioche, C., Artiguenave, F., Douchateau-Nguyen, G., Lemaire, M., Marreissne, R., Mointrener, R., Robert, C., Termer, M., Wincker, P., and Wosiewisch-Leuel, M. (2000) FEBS Lett. 487, 66–70
41. Ciften, P., Sudarsananam, P., Desikan, A., Fulton, L., Fulton, B., Majors, J., Waterston, R., Cohen, B. A., and Johnston, M. (2003) Science 301, 71–76
42. Gouet, P., Courcelle, E., Stuatt, D. I., and Metoz, F. (1999) Bioinformatics 15, 305–308
43. Gietz, R. D., and Sugino, A. (1988) Gene (Amst.) 74, 527–534
44. Scafe, C., Chao, D., Lepes, J., Hirsch, J. P., Henry, S., and Young, R. A. (1990) Nature 347, 491–494
45. Slater, M. R., and Craig, E. A. (1987) Mol. Cell. Biol. 7, 1906–1916