Review TRENDS in Biochemical Sciences Vol.29 No.12 December 2004

Rpb4 and Rpb7: subunits of RNA polymerase II and beyond

Mordechai Choder

Department of Molecular Microbiology, Rappaport Faculty of Medicine, Technion – Israel Institute of Technology, Haifa, Israel 31096

RNA Polymerase II (pol II) is a large multi-subunit complex that is responsible for the synthesis of all eukaryotic mRNAs. Its correct and timely recruitment to promoter regions is a crucial step of transcription regulation, involving complicated and well-controlled networks of protein–DNA and protein–protein interactions. The best-studied pol II is the yeast complex consisting of 12 subunits (Rpb1–12). Rpb4 and Rpb7 form a dissociable heterodimer (Rpb4/7). The unique location of Rpb4/7 within the transcription initiation complex, and its capacity to interact with various transcription factors, suggest that it provides important links to the network of interactions that control transcription initiation. Moreover, Rpb4/7 executes some non-transcriptional activities, including mRNA transport. Hence, Rpb4/7 functions at the interface of transcriptional and post-transcriptional machinery.

The regulation of transcription is crucial to gene expression and has been intensively investigated. RNA polymerases are complex structures that can carry out transcription independently of other factors *in vitro* [1]. However, *in vivo*, the polymerases are recruited to specific promoters at the right time by extremely complex machineries (see, for example, Ref. [2]). Among the three different RNA polymerases, RNA polymerase II (pol II), which is responsible for the transcription of all eukaryotic mRNAs and some small nuclear RNA genes, is the best characterized.

Rpb4 and Rpb7: the RNA pol II context

Pol II is composed of 12 subunits, Rpb1–Rpb12 [1]. The crystal structures of the 12 subunits of yeast pol II have recently been resolved [3,4], revealing two distinct structures: (i) a ten-subunit core, which constitutes the bulk of pol II structure, that includes the catalytic active site; and (ii) a two-subunit complex comprising Rpb4 and Rpb7 (Rpb4/7; GenBank accession numbers NC_001142 and NC_010692, respectively; Figure 1). The additional structural constituent of pol II, the C-terminal domain (CTD) of Rpb1, is not seen in the crystal structure because it is unstructured. The CTD plays a pivotal part during

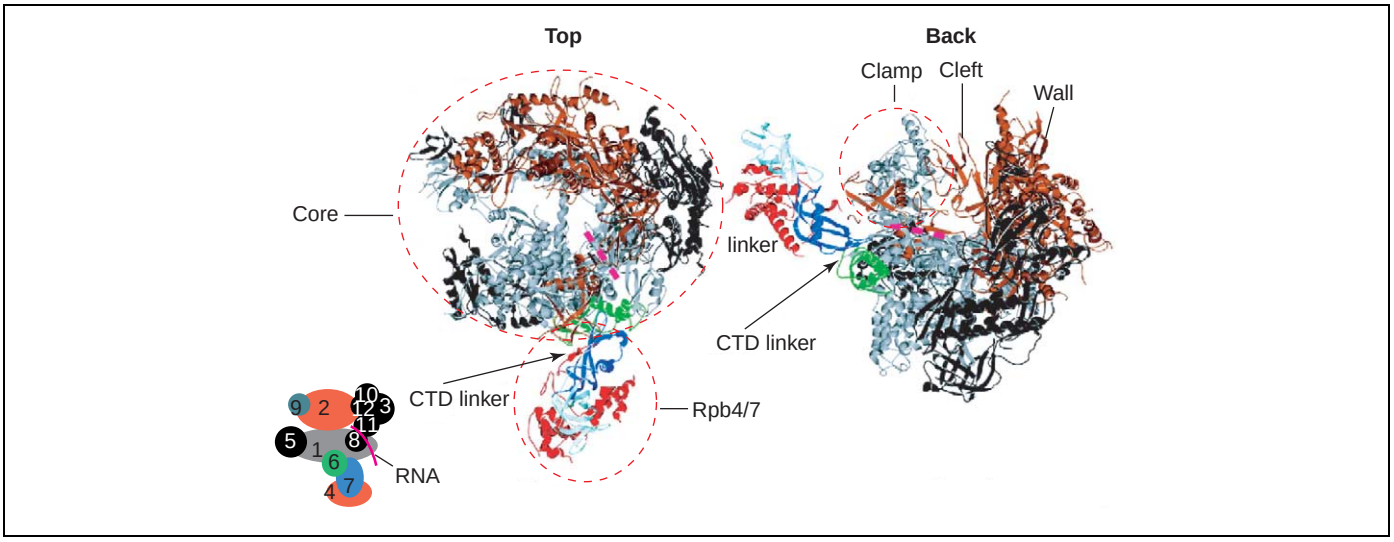

Figure 1. Backbone model of 12-subunit polymerase II (pol II) reveals two distinct structures. A ribbon representation of the pol II structure from the top view and back view is shown (orientation is as in Ref. [57]). The two encircled regions indicate the 'core' containing ten subunits and the Rpb4/7 heterodimer. A third region – the C-terminal repeat domain (CTD) of Rpb1 – is unstructured, and is therefore not detected in the crystal; the proposed position of its 'linker' is indicated by arrows. The CTD itself probably protrudes away from the core. The two, large subunits 1 and 2 form opposite sides of a central 'cleft'. The cleft accommodates the template DNA and contains the active center [43]. The cleft is constricted at one end by a protein 'wall'; the other side of the cleft is formed by a mobile 'clamp' (indicated in the model of the back view), the mobility of the clamp seems to be restricted by the presence of Rpb4/7. The location the proposed RNA exit groove 1 (thick, red broken line) is indicated. The N-terminal half of Rpb7, which contains the RNP domain (a conserved fold involved in RNA-binding, see Box 3 for details), is dark blue; the C-terminal half of Rpb7, which contains the oligo-binding fold, is light blue. The inset shows the color code for the other subunits and also their approximate positions. The assumed direction of the emerging transcript is depicted as a pink line in the inset diagram. Reproduced, with permission, from Ref. [4]. © (2003) National Academy of Sciences, U.S.A.

Corresponding author: Mordechai Choder (choder@technion.ac.il).

Available online 22 October 2004

 0968-0004/$ - see front matter doi:10.1016/j.tibs.2004.10.007

different stages of the transcription cycle by acting as an assembly platform for various proteins that affect both transcription and the processing of the emerging transcript (for a recent review, see Ref. [5]). Two of the structural components of pol II have been studied intensively, whereas Rpb4/7 has received less attention. The Rpb4/7 structure is located at a strategic region, that is, near the transcript-exit groove and adjacent to the CTD linker region (Figure 1). In the larger context of the transcription initiation complex, Rpb4/7 is situated near the general transcription factor TFIIB [4] and physically interacts with the largest subunit of TFIIF [6] and the TFIIF-interacting CTD-phosphatase Fcp1 (GenBank accession number NC_001145) [7,8]. Likewise, the Rpb4 paralog in RNA polymerase III (pol III), Rpc17, interacts with Brf1, which is the pol III counterpart of TFIIB [9], and the Rpb7 paralog in RNA polymerase I (pol I), Rpa43, binds the general initiation factor Rrn3 [10,11]. Although these observations suggest an important role for Rpb4/7 in transcription regulation, the functions of these subunits have only recently begun to be deciphered.

Rpb4 and Rpb7 were first identified in *Saccharomyces cerevisiae* as the fourth (Rpb4) and seventh (Rpb7) largest subunits of pol II [1,12,13]. By far the most studied among Rpb4 and Rpb7 families are the *S. cerevisiae* counterparts that were used in early studies as the prototypes of their orthologs and paralogs (Box 1). Rpb4 and Rpb7 normally associate

Box 1. Rpb4/7 represents an ancient family that carries ubiquitous function

In both structure [3,4] and function [49], Rpb4 and Rpb7 [two subunits of RNA polymerase II (pol II)] are homologous to the two subunits of the archaeal RNA polymerase, RpoE and RpoF (the *Methanococcus jannaschii* orthologs of Rpb4 and Rpb7). It seems, therefore, that this family of heterodimers evolved before *eukarya* and *archae* separated. Indeed, a Rpb4/7-like heterodimer has been found in each of the three forms of eukaryotic RNA polymerase that probably evolved from the archaeal origin [9,11,36,50–52]. In general, the key features common to all Rpb4 and Rpb7 paralogs and orthologs that have been examined to date are their ability to stably interact with each other, often also outside the context of pol II [11,15,29,33,35,36,38,49,52–54]. Another feature that has been found in several Rpb4/7 homologs is their capacity to bind RNA *in vitro* [33,36]. However, the *in vivo* relevance of this RNA-binding capacity remains to be determined.

Throughout archae and eukarya, the sequence and structure of Rpb7 paralogs and orthologs are highly conserved. Their functional conservation could be manifested by the capacity of Schizosaccharomyces pombe Rpb7, Candida albicans Rpb7, Drosophila melanogaster Rpb7 and Homo sapiens HsRpb7 to replace Rpb7 in Saccharomyces cerevisiae [21,32,40]. Interestingly, there is a strong evolutionary selection for maintaining only the basic folds of the RNA-binding domains (Figure I), again suggesting that RNA-binding capacity is pertinent to Rpb7 function. However, the capacity of Rpb4/7 and its homologs to bind to RNA *in vivo* remains to be determined.

The Rpb4 family is the least conserved of the subunits of RNA polymerases. RpoF, which seems to be the family prototype, is a short (107-amino acid) protein that tightly interacts with RpoE. However, some RpoF homologs are much longer than RpoF, with extensions in either their N or C termini. The existence of these long extensions suggests that, during evolution, Rpb4 homologs have been used as vehicles for recruiting other functional domains into the transcription or post-transcription arena (an interesting example was reported in Ref. [52]). This was made possible owing to the unique location of Rpb4 within pol II complex, away from the core ten-subunit complex [3,4].

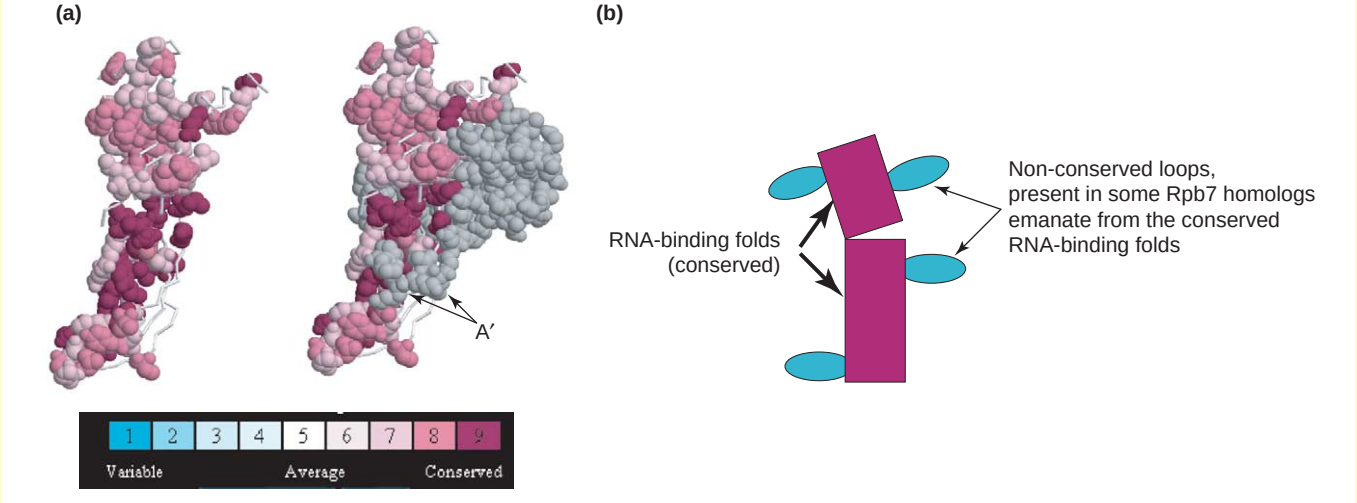

Figure I. Two putative RNA-binding folds of RpoE, the archeal ortholog of Rpb7, form a continuous and conserved path. **(a)** Resolution of the RpoE structure [35] is shown as a space-fill model, with evolutionary conservation scores represented by a color code onto the Van der Waals surface of RpoE. The model was generated using the X-ray crystal structure of the RpoE/F heterodimer (PDB code: 1go3). The evolutionary conservation of RpoF was analyzed using the maximum likelihood-based algorithm 'Rate4site' and the 'ConSurf' server (http://conseq.bioinfo.tau.ac.il/) [60]. The analysis yielded a phylogenetic tree, which was composed of 27 unique sequences of the Rpb7 family. The color of each sphere represents the degree of conservation of that particular residue; blue is the non-conserved and dark purple is the most conserved. Only the most conserved residues (scored 7–9 in the color-code bar) are shown. The gray spheres shown in the right-hand structure represent RpoF, the archeal ortholog of Rpb4. The models are oriented such that the RpoF distal face is facing the reader. Note that the conserved residues form a continuous path along the two putative RNA-binding domains of RpoE [35]. The variable residues are scattered throughout the molecule (not shown). It was found that the anti-parallel β strands of the two putative RNA-binding folds are conserved among 46 homologs, whereas some loops between the β strands vary and can contain extensions of various lengths (see supplementary material to Ref. [36]). This selective conservation suggests that there was a strong evolutionary selection for maintaining only the basic folds of the putative RNA-binding domains, which reinforces its possible biological significance. **(b)** Illustration of the selective conservation. The putative RNA-binding domains that are conserved are schematized as purple rectangles. In this model, as in the RpoE/F crystal structure, there is ~45° between their longitudinal axes when the structure is oriented as in Figure 1a (left). Ovals represent the non-conserved regions, present in some Rpb7 homologs, that protrude away from the conserved 'core' and, thereby, do not interfere with the RNA-binding folds.

with their cognate polymerase as a heterodimer [3,4,14,15]. Nevertheless, Rpb7 is an essential protein (like most pol II subunits) [13,16], whereas Rpb4 is dispensable under optimal environmental conditions but essential under some adverse conditions [12,14,17–21]. Understanding how the environment determines whether or not Rpb4 would be required for life is still an important challenge (see later).

One of the first observations that singled Rpb4/7 out as unusual among pol II subunits was its sub-stoichiometric association with the pol II complex [22] and its propensity to dissociate from the complex [15,23]. Although the significance of these features is still unclear, it seems likely that they are functionally related. Indeed, the interaction of the Rpb4/7 heterodimer with the core of pol II is complex and varies according to the environmental conditions. Thus, in exponentially proliferating cells, only ~20% of the pol II complexes contain Rpb4/7 [17], and the pol II complexes from such proliferating cells could not form a crystal. By contrast, in starving cells that enter stationary phase, Rpb4/7 binds pol II in a stoichiometric manner [17], and 2D crystal structures of the 12-subunit complex purified from such starved cells has been obtained [24] (Figure 1). It seems that neither Rpb4 nor Rpb7 undergo phosphorylation [22], and that some as yet unknown modification(s) of the ten-subunit complex seem(s) to be responsible for its starvation-specific enhanced interaction with Rpb4/7 [25]. Although understanding the biological significance of the complex interaction behavior of pol II with Rpb4/7 is likely to be important for understanding Rpb4/7 function, it has received relatively little attention in the literature. Several rationales to the complex interaction of Rpb4/7 with the core pol II can be, or have been, proposed:

- First, Rpb4/7 protects the pol II complex. Thus, Rpb4/7 is recruited to the pol II complex to prevent adverse conformational changes that might occur during periods of long-term starvation [17] (see later).
- Second, the interaction of Rpb4/7 with pol II is promoter specific [21,26,27], so that this heterodimer would be recruited in only 20% of the initiation events in optimally proliferating cells.
- Third, the heterodimer and pol II might interact only during some specific stages of the transcription cycle. Thus, at any given time during exponential proliferation, only 20% of the pol II molecules are in the interactive stage, whereas, during stationary phase, most pol II molecules are in that stage.

It is important to note, however, that the *Schizosaccharomyces pombe* and human orthologs of Rpb4/7 seem to associate stiochiometrically with their cognate pol II complex during the exponential growth phase [28,29]. The mechanistic understanding of these differences between the organisms is still missing.

Beyond the pol II context

In *S. cerevisiae*, Rpb4 is present in excess over the other pol II subunits [25]. Careful quantitative immuno-blotting experiments of *S. pombe* pol II subunits have revealed that, whereas the level of *S. pombe* Rpb7 (SpRpb7) is only slightly greater than that of the pol II complex, the level of *S. pombe* Rpb4 (SpRpb4) is sevenfold higher than that of SpRpb7 [30]. This feature, and the inability of over-expressed *RPB7* to suppress the sensitivity of *RPB4*Δ cells to the high temperature (e.g. 37°C) easily tolerated by wild type [14,20,29,32], suggests that Rpb4 maintains some of its functions independently of Rpb7. Sedimentation of whole-cell extracts through sucrose or glycerol gradients reveal that Rpb4 does not sediment as a free protein, but appears to be associated with pol II and, to a much greater extent, with smaller complexes [25,30]. These observations suggest that Rpb4 or Rpb4/7 has a role outside the context of pol II, which is consistent with the post-transcriptional functions assigned to Rpb4 (Box 2).

As pol II is a nuclear complex, it is not surprising that Rpb4 and Rpb7 are localized to the nucleus in

Box 2. New roles for Rpb4 and Rpb7

Whereas the Rpb4/7 heterodimer is a subunit of RNA polymerase, recent works have uncovered some non-transcriptional roles.

New role in DNA repair

Rpb4 was reported to have a dual role in regulating transcription-coupled repair: (i) repressing the Rpb9-mediated sub-pathway and (ii) facilitating the Rad26-mediated sub-pathway that also operates in non-transcribed regions [55]. Moreover, Rpb7 was implicated in the DNA-repair mechanism induced in response to bleomycin – an anti-tumor drug that kills cells (mammalian as well as yeast) by introducing lesions into replicating DNA [56]. Both observations suggest a role for Rpb4/7 in a transcription-coupled DNA-repair mechanism. An interesting question that remains to be answered is: how can Rpb4 affect repair in non-transcribed DNA regions?

New possible role in transcription termination

Another potential new role for Rpb4/7 in transcription termination was uncovered by identifying Nrd1 [an RNA-binding protein implicated in 3′-end formation of small nucleolar and small nuclear RNAs transcribed by RNA polymerase II (pol II)] as a Rpb7-interacting protein. Similar interactions, both physical and two-hybrid ones, were observed between *Schizosaccharomyces pombe* Rpb7 and Seb1, an *S. pombe* ortholog of Nrd1, suggesting that this interaction is conserved throughout evolution [16]. A direct role for Rpb4 in transcription termination remains to be determined.

New role in mRNA export

Recently, a post-transcriptional role was assigned to Rpb4: it was found that Rpb4 is required for efficient mRNA transport [18]. This Rpb4 activity is detectable only during stress and is essential for survival during stress. The involvement of Rpb4 in transcription and in mRNA export can be uncoupled genetically by specific mutations in *RPB4*; each mutation compromises only one of these two functions. Interestingly, over-expression of *NSP1* – a component of the nuclear pore complex – suppressed the sensitivity of *RPB4Δ* cells to mild temperature stress [20]. It is possible that the high level of Nsp1 changed the nuclear pore complex, thus enabling it to export mRNA even in the absence of Rpb4.

Concluding remark

The possible function of Rpb4/7 in anchoring a processing factor to the pol II apparatus, together with its role in mRNA transport and its cytoplasmic localization during stress, suggests that Rpb4/7 is involved in coupling between the various steps of mRNA synthesis and transport, and that it functions both within the context of pol II and beyond.

S. cerevisiae [18]. However, human Rpb4 and Rpb7 [HsRpb4 (GenBank accession number NM_004805) and HsRpb7 (GenBank accession number U20659), respectively] are found in both the nucleus and the cytoplasm [21,26]. Similarly, a small portion of Rpb4 is found in the cytoplasm of *S. pombe* [30]. Recently, it has been observed that the localization of Rpb4 is strongly influenced by the environmental conditions: in response to stresses of high temperatures, ethanol or starvation – when Rpb4 becomes essential – a large portion of Rpb4 molecules is exported to the cytoplasm [18]. Heat shock-induced relocation of Rpb4 is also a characteristic of *S. pombe* cells [30]. The stress-induced export has been correlated with the new role of Rpb4 in mRNA export (Box 2).

Roles of Rpb4 and Rpb7 in transcription

The ability of Rpb4/7 to dissociate from pol II in a reversible manner has been exploited to demonstrate that Rpb4/7 is required for promoter-directed initiation of transcription *in vitro* [15,33]. However, Rpb4/7 is not required for stable recruitment of polymerase to active pre-initiation complexes, suggesting that the heterodimer mediates an essential step subsequent to promoter binding [33]. In contrast with its role in promoter-dependent initiation, under conditions of optimal temperature, Rpb4/7 does not seem to be required for elongation or for promoter-independent initiation [when single-stranded poly(rC) [(rCytidine)$_n$] is used as the template] [15,23,25]. *In vivo*, Rpb4 and Rpb7 do not contribute equally to the function of the heterodimer. At moderate temperatures and otherwise optimal growth conditions under which some *RPB4Δ* strains proliferate comparably to wild-type cells (this feature is strain dependent) [17], the overall transcriptional capacity of the two strains is similar [17,19,31]. By contrast, when *RPB4Δ* cells are exposed to the stresses of temperatures (over 30°C or under 12°C), ethanol, or some starvation conditions, they lose their ability to transcribe most, if not all, genes [14,17–19,31,34]. However, Rpb4 is not required under all stress conditions. For example, cells lacking *RPB4* exhibit normal transcriptional capacity during oxidative stress, osmotic stress [31] or nitrogen starvation [34]. It seems, therefore, that Rpb4 has an essential role in some adverse conditions, but not all stressful conditions.

The temperature-sensitive phenotype of *RPB4Δ* cells can be partially suppressed by the over-expression of *RPB7* (at 34°C but not at 37°C), indicating that Rpb7 can function in transcription independently of Rpb4 even under moderate temperature stress [14,20,31]. Moreover, yeast cells carrying HsRPB7 instead of *RPB7* can survive at moderate temperatures; however, at temperature extremes and during maintenance at stationary phase, these cells lose their viability more rapidly than wild-type cells do. Together, these results suggest that Rpb7 plays a key part in transcription and that Rpb4 enables, enhances or modifies the function of Rpb7.

High-resolution crystal structure of the archaeal ortholog of Rpb4/7

Archaeal *Methanococcus jannaschii* has a single RNA polymerase consisting of 12 subunits that display considerable homology to pol II subunits. Of the 12 subunits, MjRpoF (referred to here as RpoF) and MjRpoE (referred to here as RpoE) are the orthologs of Rpb4 and Rpb7, respectively. Silvia Onesti and her colleagues determined the structure of the RpoE/F heterodimer to a resolution of 1.75 Å [35]. Their study provided the first detailed structural information about a member of the Rpb4/7 family, forming a basis for understanding the structure and possible function of the entire family of these heterodimers (Box 3). Remarkably, a partially conserved RNA-binding 'path' was identified along the surface of the two RNA-binding domains [35] (Box 1). In many RNA-binding proteins (e.g. in most proteins containing an S1 motif), RNA binding is spread over more than one domain, with the RNA-binding sites forming a continuous surface. This similarity to the S1 subfamily reinforces the possible significance of the putative RNA-binding surface in RpoE. The ability of RpoE/F to bind RNA has been demonstrated *in vitro* [36]. Similarly, Rpb4/7 and Rpa14/43 (the pol I Rpb4/7 paralog) have also been shown to bind RNA *in vitro* [33,36]. Thus, it seems that at least one RNA-binding domain of Rpb7 is functional *in vitro* and could be important for the heterodimer functions *in vivo*. However, there is no indication that RpoF has an RNA-binding site of its own. Recently, it was reported that Rpa14 – a member of the Rpb4 family (Box 1) – has no detectable RNA-binding activity of its own, whereas Rpa43 – a member of the Rpb7 family that heterodimerized with Rpa14 – can bind RNA independently of Rpa14 [36].

The structure of Rpb4/7 within the 12-subunit pol II complex: functional implications

Recently, two research groups have independently reported the structure of the 12-subunit pol II complex at resolution of ~4 Å [3,4]. Both groups have found that the structure of Rpb4/7 in the context of the 12-subunit complex is similar to that of its archaeal ortholog RpoE/F. Except for a region of ~70 residues towards the N terminus of Rpb4, the similarity between RpoE/F structure and Rpb4/7 structure was high enough that it was possible to use the archaeal structure (which was known in more detail) as a guide for data analyses.

Rpb4/7 interacts with the core pol II (comprising ten subunits) via two, small, core–dimer interfaces. The main contact point resides in a 'tip' of Rpb7 (see Figure Ib, Box 3). The tip, which is composed of two of Rpb7 outmost loops (based on RpoF structure), interacts with five protein regions that form a pocket in the core pol II. This pocket is lined with three regions of Rpb1, one of Rpb2 and one of Rpb6 [3,4]. The importance of the region of Rpb6 – residues 90–105, which form a conserved α helix – has been demonstrated biochemically: pol II containing a mutant form of Rpb6, the conserved glutamine at position 100 of which was replaced by arginine, binds very poorly to the Rpb4/7 heterodimer [37]. This major contact point between Rpb7 and the pocket in pol II core complex does not involve Rpb4, which explains why Rpb7 can interact with pol II independently of Rpb4 [14]. One group found that the N-terminal region of Rpb4 makes a second contact with the N-terminal region of Rpb1, further stabilizing the

Box 3. Structure of the archeal ortholog of Rpb4/7

Figure I shows the RpoE/F – the *Methanococcus jannaschii* ortholog of Rpb4/7 – crystal structure [35] as a space-filling model, with evolutionary conservation scores represented by a color code onto the Van der Waals surface of RpoE. Thus, the color of each sphere represents the degree of conservation of that particular residue. The model shows that RpoE (the Rpb7 ortholog) has an elongated two-domain structure, each of which contains a putative RNA-binding fold [35]. The N-terminal domain of RpoE, which constitutes the bottom half of the structure, resembles a truncated RNP fold. This RNP fold contains conserved charged and aromatic side chains that, in other RNP domains, are known to be involved in interactions with RNA (see Ref. [35] and references therein). This domain is structurally similar to the B8 domain of phenylalanine tRNA synthetase. The C-terminal domain contains a five-stranded anti-parallel β barrel, which resembles an oligo-binding (OB) fold that is often involved in binding nucleic acids [58]. This OB fold is found in the bacterial ribosomal protein S1 and in several other proteins that bind single-straded RNA, often in a sequence-nonspecific manner [59].

RpoF (the Rpb4 ortholog) shown in gray (as Rpb4 family is relatively unconserved) is composed of one β strand at its N terminus (A′) and six α helices (H1–H6) that pack against one side of RpoE at the hinge between the two RpoE domains. RpoF forms a semi-circular belt that probably stabilizes the RpoE structure. In addition, A′ contributes an extra β- strand to the β sheet of the RpoE truncated RNP motif, thereby, further stabilizing its structure and possibly contributing to its proposed ability to interact with RNA [35].

Note that the upper domain of RpoE rotates by ~45° (Figure I). Visualization of the ~45° rotation of RpoE is emphasized by the dashed lines that are parallel to the horizontal axis of the upper domain. This orientation of the RpoE domains apears to be held together by RpoF. I propose that the relative positions of the two domains of RpoE might be important mechanistically and that, in the absence of RpoF, the two domains are free to move one with respect to the other. Such degree of freedom might affect RpoE function. Under stress, this freedom might compromise Rpb7 activity.

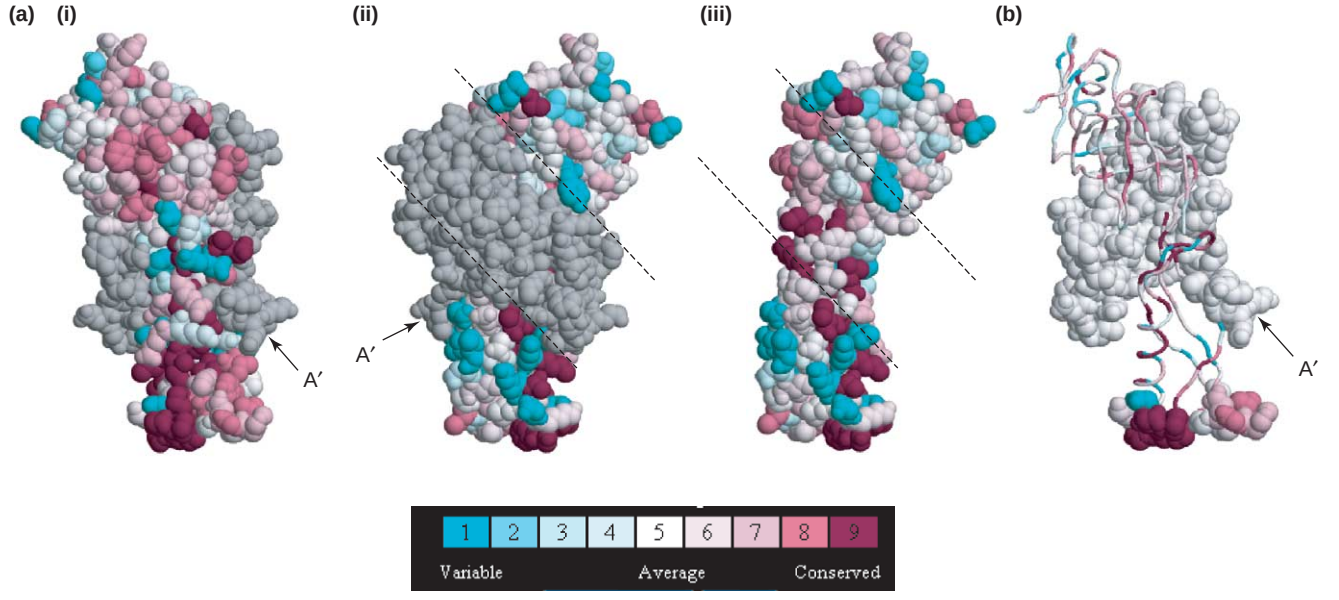

Figure I. Crystal structure of the RpoE/F heterodimer [35] **(a)** Asymmetric distribution of conserved residues in RpoE (the Rpb7 ortholog). (i) The RpoF (the Rpb4 ortholog) distal surface of RpoE (the Rpb7 ortholog). (ii) The RpoF-interacting face of RpoE [shifted 180° around the vertical axis relative to the model in (i)]. (iii) The same structure as in model (ii), but lacking the RpoF. The color of each sphere represents the degree of conservation of that particular residue. The coloring of the model in (i) shows that the RpoF distal surface, which can interact with RNA, is evolutionarily conserved. The RpoF-interacting region in RpoE [between the two parallel broken lines of (iii)] is also conserved. By contrast, outside the RpoF-interacting region and outside the tip (at the bottom of the structure), the RpoF proximal face of RpoE is not conserved. This conservation asymmetry between the two faces suggests that the RpoF distal face is the functional one. **(b)** The polymerase II (pol II)-interacting region, the 'tip', is conserved in the Rpb7 family. Residues that interact with RNA polymerase according to the 12-subunit complex of the yeast pol II [3,4] are shown as colored spheres. The model was generated using the X-ray crystal structure of the RpoE/F heterodimer (PDB code: 1go3). The evolutionary conservation of RpoF was analyzed using the maximum likelihood-based algorithm 'Rate4site' and the 'ConSurf' server (http://conseq.bioinfo.tau.ac.il/) [60]. The analysis yielded a phylogenetic tree, which was composed of 27 unique sequences of the Rpb7 family. RpoF and the position its A′ arm are shown to help with the orientation. The picture was made using RASMOL (http://www.umass.edu/microbio/rasmol).

interaction [4]. In addition, Rpb4 probably stabilizes the core-interacting conformation of Rpb7 [3,35]. The small core-Rpb4/7 contact interfaces can explain why the heterodimer is readily dissociated from pol II in *S. cerevisiae* (although in other organisms the heterodimer–core interaction might be more stable [28,29,38]).

Except for these one or two local points of contact, the heterodimer is protruded away from the core polymerase and is, thus, accessible to other initiating factors. It has been proposed that Rpb4/7 extends the upstream interacting face of pol II, extending the capacity of the polymerase to interact with a variety of transcription factors [3]. Thus, it can serve as a bridge between the polymerase and an initiation factor and can contribute to promoter specificity. Indeed, the heterodimer was reported to interact with some general transcription factors (see earlier). However, the interaction of Rpb4/7 extends beyond interactions with general transcription factors. HsRpb7 was reported to interact with NOV, a putative transcription factor [39]. HsRpb7 also interacts with the N-terminal region of Edwin sarcoma oncogene (called EAD), an activating domain fused to a Fli1 DNA-binding domain due to chromosomal translocation [27,40]. Fli1 itself is a (non-oncogenic) transcription factor.

In nuclear extracts, HsRpb7 was found to co-purify with EAD–Fli1 but not with Fli1, and over-expression of recombinant HsRpb7 specifically increases gene activation by EAD-chimeric transcription factors [27]. These results suggest that HsRpb7 has a selective role. Interestingly, EAD fused to the DNA-binding domain of Gal4 (GenBank accession number NC_001148) can activate transcription of a Gal4-dependent reporter gene in yeast, reflecting the high degree of conservation of the Rpb7 family. *RPB4Δ* yeast cells are defective in EAD-dependent activation, even at 23°C. This defect cannot be corrected by introduction of HsRPB4; however, introduction of a combination of HsRPB4 and HsRPB7 in yeast cells lacking *RPB4* and carrying *RPB7* restores the EAD-dependent activation [40]. Thus, although EAD probably interacts with Rpb7, Rpb4 is still required to materialize the interaction probably by stabilizing the correct Rpb7 conformation. These results also indicate that HsRpb4/7 can be functional in yeast and, if recruited to the transcription arena by EAD, can replace Rpb4/7.

Rpb4/7 is positioned near the CTD linker [3,4,6] (Figure 1). This location might be significant because *RPB7* was demonstrated to be associated genetically with *ESS1*, a prolyl isomerase that binds CTD and affects transcription elongation and termination – probably by modifying CTD conformation [41]. Moreover, Rpb4 was found to interact physically with Fcp1, a CTD dephosphatase [7,8]. The *in vitro* interaction of Fcp1 with the pol II complex is dependent on Rpb4/7 [7], suggesting that Rpb4/7 actually recruits Fcp1. These interactions and the close proximity of Rpb4/7 to the CTD imply that Rpb4/7 influences binding of some CTD-interacting proteins.

Various pol II X-ray structures have revealed a deep cleft that accommodates both the DNA and the active site. One side of the cleft is formed by a domain called the 'clamp' (Figure 1), the conformation of which varies according to the crystal composition. This clamp adopts an open configuration in the ten-subunit structure that lacks Rpb4/7, which leaves room for double-stranded DNA to enter the active-site groove [42]. The structure of the elongating pol II lacking Rpb4/7 with bound template DNA and short transcript (pol IIΔ4/7–DNA–RNA co-crystal) reveals that the clamp changes its conformation and closes the cleft, thus trapping the template DNA inside [43]. Trapping the DNA in the pol II complex probably contributes to the elongation processivity [43,44]. Evidently, Rpb4/7 is not needed for this conformational change. Importantly, then, the current data do not indicate that Rpb4/7 is required for clamp closure during transcription elongation. Nevertheless, the interaction of Rpb4/7 restricts the clamp to the closed position even in the absence of the DNA template [3,4]. It was therefore suggested by both research groups that, if Rpb4/7 is assembled with pol II before initiation, the closed configuration of pol II would prevent the double-stranded promoter DNA from reaching the active site. It was suggested then, that the double-stranded DNA would first bind pol II far above the active site; only after DNA melting occurs would the template strand be able to reach the active site owing to its smaller size and higher flexibility relative to the rigid double-stranded DNA [3,4,44,45]. A similar model has been proposed for the initiation competent *Escherichia coli* RNA polymerase holoenzyme (which includes the σ subunit), the structure of which shows a similar closed clamp [46,47]. It is worth noting, however, that it is not yet clear at what step during transcription initiation the interaction of Rpb4/7 and pol II occurs. The 12-subunit pol II has a greater affinity than pol IIΔ4/7 for the TATA-binding protein–TFIIB promoter complex. It has therefore been proposed that Rpb4/7 assembles with pol II before the interaction of pol II with the promoter and that this dimer–core interaction enhances pol II recruitment [24]. By contrast, it has been reported that pol II can join a stable active pre-initiation complex with the template *in vitro* in the absence of Rpb4/7. Addition of Rpb4/7 would then be necessary to promote efficient initiation following pre-initiation complex assembly [33].

Why is Rpb4 essential under stressful conditions but dispensable under optimal conditions?

There are at least three possible answers to this question that are not mutually exclusive:

(i) Rpb4 is required to enhance the essential role of Rpb7. Thus, under optimal conditions, when gene-expression mechanisms function at their full capacity, this enhancement function is dispensable. For example, the Rpb7 'tip' (discussed earlier; see Box 3, Figure Ib) is placed in a pocket within the core subunit, forming several direct contacts with Rpb1, Rpb2 and Rpb6 (see earlier) that holds them together in the preferred conformation [3,4]. This interaction, enhanced by Rpb4 interaction with both Rpb7 and Rpb1, might lock the active pol II conformation. Hence, in the absence of Rpb4, the 'unlocked' core pol II is especially vulnerable to some stresses. The stabilization role of Rpb4, first proposed more than a decade ago [17], has been challenged using promoter-independent *in vitro* transcription experiments that tested pol II enzymatic activity independently of auxiliary transcription factors. This approach demonstrated that Rpb4, probably through its interaction with Rpb7, is required for *in vitro* pol II elongation activity only under extreme temperatures [25]. Another potential role of Rpb4 in enhancement of Rpb7 function is its possible stabilization of the specific conformation of Rpb7 (discussed in some detail in Box 3).

(ii) Under optimal conditions, Rpb4 is important for the transcription of a few genes whose products are not required for proliferation [34,48]. Under stressful conditions, the products of these genes are required for cell survival.

(iii) Rpb4 has stress-specific roles: it differentially interacts with specific proteins under stress and/or mediates some stress-essential function(s), either in transcription or in post-transcription processes. Recall that during stress most Rpb4 molecules are found in the cytoplasm [18], suggesting a role for Rpb4 in the cytoplasm.

The current data do not rule out any of these possibilities. In future research, revealing the molecular

basis for the sensitivity of *RPB4Δ* cells to stress will be important for furthering our understanding of the functions of the Rpb4/7 heterodimer and of the capacity of the transcription and transport machineries to cope with stress conditions.

Concluding remarks and future perspectives

The Rpb4/7 heterodimer plays a part in transcription, mRNA transport and DNA repair. Hence, Rpb4/7 is found at two important junctions: (i) at the intersection of transcription factors, core pol II and transcript exit, and (ii) at the intersection between the mechanisms of transcription (including transcription-coupled DNA repair) and transport.

It is possible to view Rpb4/7 as one of three distinct structural constituents of pol II (Figure 1). At the same time, several features of Rpb4/7 resemble those of general transcription factors. Similar to general transcription factors, Rpb4/7 aids the assembly of the initiation complex in the promoter region by interacting with both transcription activators and general transcription factors, including pol II, TFIIF and TFIIB. Both views of Rpb4/7 suggest an important role for Rpb4/7 in transcription.

Rpb4 is involved in the appropriate responses of the cell to various stressful conditions, suggesting that Rpb4/7 helps link environmental signals to key processes in gene expression. However, the signaling pathway (or pathways) that controls the various functions of Rpb4/7 is still unknown. There are additional enigmas related to Rpb4/7 functions:

- Is there a common mechanistic denominator underlying the various different functions of Rpb4/7?
- Does Rpb4/7 have a role in the cytoplasm?

As Rpb4/7 functions impinge on transcription initiation, termination and mRNA transport, detailed understanding of its activities will advance our understanding of how the various steps of gene expression are intimately coupled.

Acknowledgements

I thank Silvia Onesti for critically reading the article and Roger Kornberg for providing the model of the 12-subunit complex. Work in my laboratory is supported mainly by grants from the Israel Science Foundation founded by the Israel Academy of Sciences and Humanities, and by the USA–Israel Binational Science Foundation. I apologize to colleagues whose work could not be included in this review owing to space constraints.

References

1 Young, R.A. (1991) RNA polymerase II. *Annu. Rev. Biochem.* 60, 689–715
2 Levine, M. and Tjian, R. (2003) Transcription regulation and animal diversity. *Nature* 424, 147–151
3 Armache, K.J. *et al*. (2003) Architecture of initiation-competent 12-subunit RNA polymerase II. *Proc. Natl. Acad. Sci. U. S. A.* 100, 6964–6968
4 Bushnell, D.A. and Kornberg, R.D. (2003) Complete, 12-subunit RNA polymerase II at 4.1-Å resolution: implications for the initiation of transcription. *Proc. Natl. Acad. Sci. U. S. A.* 100, 6969–6973
5 Palancade, B. and Bensaude, O. (2003) Investigating RNA polymerase II carboxyl-terminal domain (CTD) phosphorylation. *Eur. J. Biochem.* 270, 3859–3870
6 Chung, W.H. *et al*. (2003) RNA polymerase II/TFIIF structure and conserved organization of the initiation complex. *Mol. Cell* 12, 1003–1013
7 Kamenski, T. *et al*. (2004) Structure and mechanism of RNA polymerase II CTD phosphatases. *Mol. Cell* 15, 399–407
8 Kimura, M. *et al*. (2002) Formation of a carboxy-terminal domain phosphatase (Fcp1)/TFIIF/RNA polymerase II (pol II) complex in *Schizosaccharomyces pombe* involves direct interaction between Fcp1 and the Rpb4 subunit of pol II. *Mol. Cell. Biol.* 22, 1577–1588
9 Ferri, M.L. *et al*. (2000) A novel subunit of yeast RNA polymerase III interacts with the TFIIB-related domain of TFIIIB70. *Mol. Cell. Biol.* 20, 488–495
10 Yuan, X. *et al*. (2002) Multiple interactions between RNA polymerase I, TIF-IA and TAF(I) subunits regulate preinitiation complex assembly at the ribosomal gene promoter. *EMBO Rep.* 3, 1082–1087
11 Peyroche, G. *et al*. (2002) The A14–A43 heterodimer subunit in yeast RNA pol I and their relationship to Rpb4-Rpb7 pol II subunits. *Proc. Natl. Acad. Sci. U. S. A.* 99, 14670–14675
12 Woychik, N.A. and Young, R.A. (1989) RNA polymerase II subunit RPB4 is essential for high- and low-temperature yeast cell growth. *Mol. Cell. Biol.* 9, 2854–2859
13 McKune, K. *et al*. (1993) RPB7, one of two dissociable subunits of yeast RNA polymerase II, is essential for cell viability. *Yeast* 9, 295–299
14 Sheffer, A. *et al*. (1999) Rpb7 can interact with RNA polymerase II and support transcription during some stresses independently of Rpb4. *Mol. Cell. Biol.* 19, 2672–2680
15 Edwards, A.M. *et al*. (1991) Two dissociable subunits of yeast RNA polymerase II stimulate the initiation of transcription at a promoter *in vitro*. *J. Biol. Chem.* 266, 71–75
16 Mitsuzawa, H. *et al*. (2003) Rpb7 subunit of RNA polymerase II interacts with an RNA-binding protein involved in processing of transcripts. *Nucleic Acids Res.* 31, 4696–4701
17 Choder, M. and Young, R.A. (1993) A portion of RNA polymerase II molecules has a component essential for stress responses and stress survival. *Mol. Cell. Biol.* 13, 6984–6991
18 Farago, M. *et al*. (2003) Rpb4p, a subunit of RNA polymerase II, mediates mRNA export during stress. *Mol. Biol. Cell* 14, 2744–2755
19 Miyao, T. *et al*. (2001) Deletion of the RNA polymerase subunit RPB4 acts as a global, not stress-specific, shut-off switch for RNA polymerase II transcription at high temperatures. *J. Biol. Chem.* 276, 46408–46413
20 Tan, Q. *et al*. (2000) Multiple mechanisms of suppression circumvent transcription defects in an RNA polymerase mutant. *Mol. Cell. Biol.* 20, 8124–8133
21 Khazak, V. *et al*. (1995) Human RNA polymerase II subunit hsRPB7 functions in yeast and influences stress survival and cell morphology. *Mol. Biol. Cell* 6, 759–775
22 Kolodziej, P.A. *et al*. (1990) RNA polymerase II subunit composition, stoichiometry, and phosphorylation. *Mol. Cell Biol.* 10, 1915–1920
23 Ruet, A. *et al*. (1980) A mutation of the B220 subunit gene affects the structural and functional properties of yeast RNA polymerase B *in vitro*. *J. Biol. Chem.* 255, 6450–6455
24 Jensen, G.J. *et al*. (1998) Structure of wild-type yeast RNA polymerase II and location of Rpb4 and Rpb7. *EMBO J.* 17, 2353–2358
25 Rosenheck, S. and Choder, M. (1998) Rpb4, a subunit of RNA polymerase II, enables the enzyme to transcribe at temperature extremes *in vitro*. *J. Bacteriol.* 180, 6187–6492
26 Na, X. *et al*. (2003) Identification of the RNA polymerase II subunit hsRPB7 as a novel target of the von Hippel–Lindau protein. *EMBO J.* 22, 4249–4259
27 Petermann, R. *et al*. (1998) Oncogenic EWS-Fli1 interacts with hsRPB7, a subunit of human RNA polymerase II. *Oncogene* 17, 603–610
28 Sakurai, H. *et al*. (1999) The Rpb4 subunit of fission yeast *Schizosaccharomyces pombe* RNA polymerase II is essential for cell viability and similar in structure to the corresponding subunits of higher eukaryotes. *Mol. Cell. Biol.* 19, 7511–7518
29 Khazak, V. *et al*. (1998) Analysis of the interaction of the novel RNA polymerase II (pol II) subunit hsRPB4 with its partner hsRPB7 and with pol II. *Mol. Cell Biol.* 18, 1935–1945
30 Kimura, M. *et al*. (2001) Intracellular contents and assembly states of all 12 subunits of the RNA polymerase II in the fission yeast *Schizosaccharomyces pombe*. *Eur. J. Biochem.* 268, 612–619
31 Maillet, I. *et al*. (1999) Rpb4p is necessary for RNA polymerase II activity at high temperature. *J. Biol. Chem.* 274, 22586–22590

32 Singh, S.R. *et al*. (2004) Domainal organization of the lower eukaryotic homologs of the yeast RNA polymerase II core subunit Rpb7 reflects functional conservation. *Nucleic Acids Res.* 32, 201–210

33 Orlicky, S.M. *et al*. (2001) Dissociable Rpb4-Rpb7 subassembly of RNA polymerase II binds to single-strand nucleic acid and mediates a post-recruitment step in transcription initiation. *J. Biol. Chem.* 276, 10097–100102

34 Pillai, B. *et al*. (2001) Rpb4, a non-essential subunit of core RNA polymerase II of *Saccharomyces cerevisiae* is important for activated transcription of a subset of genes. *J. Biol. Chem.* 276, 30641–30647

35 Todone, F. *et al*. (2001) Structure of an archaeal homolog of the eukaryotic RNA polymerase II RPB4/RPB7 complex. *Mol. Cell* 8, 1137–1143

36 Meka, H. *et al*. (2003) Structural and functional homology between the RNAP(I) subunits A14/A43 and the archaeal RNAP subunits E/F. *Nucleic Acids Res.* 31, 4391–4400

37 Tan, Q. *et al*. (2003) Loss of the Rpb4/Rpb7 subcomplex in a mutant form of the Rpb6 subunit shared by RNA polymerases I, II, and III. *Mol. Cell. Biol.* 23, 3329–3338

38 Larkin, R.M. and Guilfoyle, T.J. (1998) Two small subunits in *Arabidopsis* RNA polymerase II are related to yeast RPB4 and RPB7 and interact with one another. *J. Biol. Chem.* 273, 5631–5637

39 Perbal, B. (1999) Nuclear localisation of NOVH protein: a potential role for NOV in the regulation of gene expression. *Mol. Pathol.* 52, 84–91

40 Zhou, H. and Lee, K.A. (2001) An hsRPB4/7-dependent yeast assay for *trans*-activation by the EWS oncogene. *Oncogene* 20, 1519–1524

41 Wu, X. *et al*. (2003) The ESS1 prolyl isomerase and its suppressor BYE1 interact with RNA pol II to inhibit transcription elongation in *Saccharomyces cerevisiae*. *Genetics* 165, 1687–1702

42 Cramer, P. *et al*. (2001) Structural basis of transcription: RNA polymerase II at 2.8 Å resolution. *Science* 292, 1863–1876

43 Gnatt, A.L. *et al*. (2001) Structural basis of transcription: an RNA polymerase II elongation complex at 3.3 Å resolution. *Science* 292, 1876–1882

44 Asturias, F.J. and Craighead, J.L. (2003) RNA polymerase II at initiation. *Proc. Natl. Acad. Sci. U. S. A.* 100, 6893–6895

45 Craighead, J.L. *et al*. (2002) Structure of yeast RNA polymerase II in solution: implications for enzyme regulation and interaction with promoter DNA. *Structure (Camb)* 10, 1117–1125

46 Vassylyev, D.G. *et al*. (2002) Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6 Å resolution. *Nature* 417, 712–719

47 Murakami, K.S. *et al*. (2002) Structural basis of transcription initiation: an RNA polymerase holoenzyme–DNA complex. *Science* 296, 1285–1290

48 Bourbonnais, Y. *et al*. (2001) Multiple cellular processes affected by the absence of the Rpb4 subunit of RNA polymerase II contribute to the deficiency in the stress response of the yeast rpb4Δ mutant. *Mol. Gen. Genet.* 264, 763–772

49 Werner, F. *et al*. (2000) Archaeal RNA polymerase subunits F and P are *bona fide* homologs of eukaryotic RPB4 and RPB12. *Nucleic Acids Res.* 28, 4299–4305

50 Smid, A. *et al*. (1995) The association of three subunits with yeast RNA polymerase is stabilized by A14. *J. Biol. Chem.* 270, 13534–13540

51 Sadhale, P.P. and Woychik, N.A. (1994) C25, an essential RNA polymerase III subunit related to the RNA polymerase II subunit RPB7. *Mol. Cell. Biol.* 14, 6164–6170

52 Siaut, M. *et al*. (2003) An Rpb4/Rpb7-like complex in yeast RNA polymerase III contains the orthologue of mammalian CGRP-RCP. *Mol. Cell. Biol.* 23, 195–205

53 Sampath, V. *et al*. (2003) The conserved and non-conserved regions of Rpb4 are involved in multiple phenotypes in *Saccharomyces cerevisiae*. *J. Biol. Chem.* 278, 51566–51576

54 Hu, P. *et al*. (2002) Characterization of human RNA polymerase III identifies orthologues for *Saccharomyces cerevisiae* RNA polymerase III subunits. *Mol. Cell. Biol.* 22, 8044–8055

55 Li, S. and Smerdon, M.J. (2002) Rpb4 and Rpb9 mediate subpathways of transcription-coupled DNA repair in *Saccharomyces cerevisiae*. *EMBO J.* 21, 5921–5929

56 He, C.H. and Ramotar, D. (1999) An allele of the yeast RPB7 gene, encoding an essential subunit of RNA polymerase II, reduces cellular resistance to the antitumor drug bleomycin. *Biochem. Cell Biol.* 77, 375–382

57 Cramer, P. *et al*. (2000) Architecture of RNA polymerase II and implications for the transcription mechanism. *Science* 288, 640–649

58 Arcus, V. (2002) OB-fold domains: a snapshot of the evolution of sequence, structure and function. *Curr. Opin. Struct. Biol.* 12, 794–801

59 Draper, D.E. and Reynaldo, L.P. (1999) RNA binding strategies of ribosomal proteins. *Nucleic Acids Res.* 27, 381–388

60 Glaser, F. *et al*. (2003) ConSurf: identification of functional regions in proteins by surface-mapping of phylogenetic information. *Bioinformatics* 19, 163–164