# Structure and TBP binding of the Mediator head subcomplex Med8–Med18–Med20

Laurent Larivière, Sebastian Geiger, Sabine Hoeppner, Susanne Röther, Katja Sträßer & Patrick Cramer

**The Mediator head module stimulates basal RNA polymerase II (Pol II) transcription and enables transcriptional regulation. Here we show that the head subunits Med8, Med18 and Med20 form a subcomplex (Med8/18/20) with two submodules. The highly conserved N-terminal domain of Med8 forms one submodule that binds the TATA box–binding protein (TBP) *in vitro* and is essential *in vivo*. The second submodule consists of the C-terminal region of Med8 (Med8C), Med18 and Med20. X-ray analysis of this submodule reveals that Med18 and Med20 form related β-barrel folds. A conserved putative protein-interaction face on the Med8C/18/20 submodule includes sites altered by *srb* mutations, which counteract defects resulting from Pol II truncation. Our results and published data support a positive role of the Med8/18/20 subcomplex in initiation-complex formation and suggest that the Mediator head contains a multipartite TBP-binding site that can be modulated by transcriptional activators.**

Mediator is a multiprotein complex that transmits gene regulatory signals from transcription factors to RNA polymerase (Pol) II[1–4]. Yeast Mediator comprises 25 subunits, of which 11 are essential for viability and 22 are at least partially conserved among eukaryotes[5,6]. Nine Mediator subunits are products of the *SRB* genes, which were identified in a genetic screen for mutants that suppress phenotypes resulting from truncation of the C-terminal repeat domain (CTD) of Pol II[7,8]. Interactions between Mediator subunits have been mapped by various methods (ref. 9 and references therein). Mediator subunits can be allocated to the head, middle and tail modules observed by electron microscopy[10,11] and to an additional kinase module.

Mediator promotes initiation-complex assembly through contacts with transcriptional activators, Pol II and general transcription factors[12]. The Mediator head module has a prominent role during initiation-complex assembly, as it binds the CTD and the general transcription factors TBP (TATA box–binding protein) and TFIIB, and stimulates basal transcription[10,13]. The head module contains subunits Med6, Med8, Med11, Med17, Med18, Med19, Med20 and Med22, which are all conserved from yeast to human. The central importance of the head module is demonstrated by a mutation in Med17 that stops essentially all messenger RNA transcription *in vivo*[14] and abolishes the stimulation of basal transcription *in vitro*[15].

Previous results have suggested that the head subunits Med8, Med18 and Med20 are structurally and functionally connected. Med18 interacts with Med20 and with Med8 (ref. 9 and references therein). Med18 and Med20 are required for formation of a stable transcription initiation complex, for efficient basal transcription and for transcription activation *in vitro*[8,16]. Med20 was first characterized as a TBP-binding protein that is required for basal and activated transcription *in vivo*[17]. Med18 and Med20 are not essential for yeast viability, but Med8 has an essential function *in vivo*.

We report here structural information on the Mediator head module, in particular the Med8/18/20 subcomplex, and complementary biochemical and genetic data. Our results elucidate the role of the Med8/18/20 subcomplex in the Mediator mechanism.

## RESULTS

### Structure determination of the Med18/20 heterodimer

Bicistronic expression of the *Saccharomyces cerevisiae* Mediator subunits Med18 and Med20 enabled purification of a stoichiometric and highly soluble Med18/20 heterodimer. The pure heterodimer formed microcrystals, which we sought to improve by removal of flexible protein regions. Flexible regions were identified around Med18 residues 55 and 120 by limited proteolysis, secondary structure prediction and alignment of protein sequences from various species. Different Med18 variants that lacked portions of these flexible regions were coexpressed with full-length Med20 and formed stoichiometric and soluble heterodimers. The variant heterodimer Med18$_{\Delta 109–140}$–Med20 formed large crystals that allowed structure determination by selenomethionine labeling and SAD phasing. Four heterodimer models in the asymmetric unit were refined to a free *R*-factor of 26.4% at 2.4-Å resolution (**Supplementary Fig. 1** online).

### Structure extension to the Med8C/18/20 trimer

As the C-terminal residues 190–223 of Med8 suffice for interaction with Med18 in a yeast two-hybrid assay[9], we tested whether C-terminal fragments of Med8 stably bind the Med18/20 heterodimer *in vitro*. Different His$_6$-tagged C-terminal Med8 variants were

Gene Center Munich, Department of Chemistry and Biochemistry, Ludwig-Maximilians-Universität München, Feodor-Lynen-Str. 25, 81377 Munich, Germany. Correspondence should be addressed to P.C. (cramer@lmb.uni-muenchen.de).

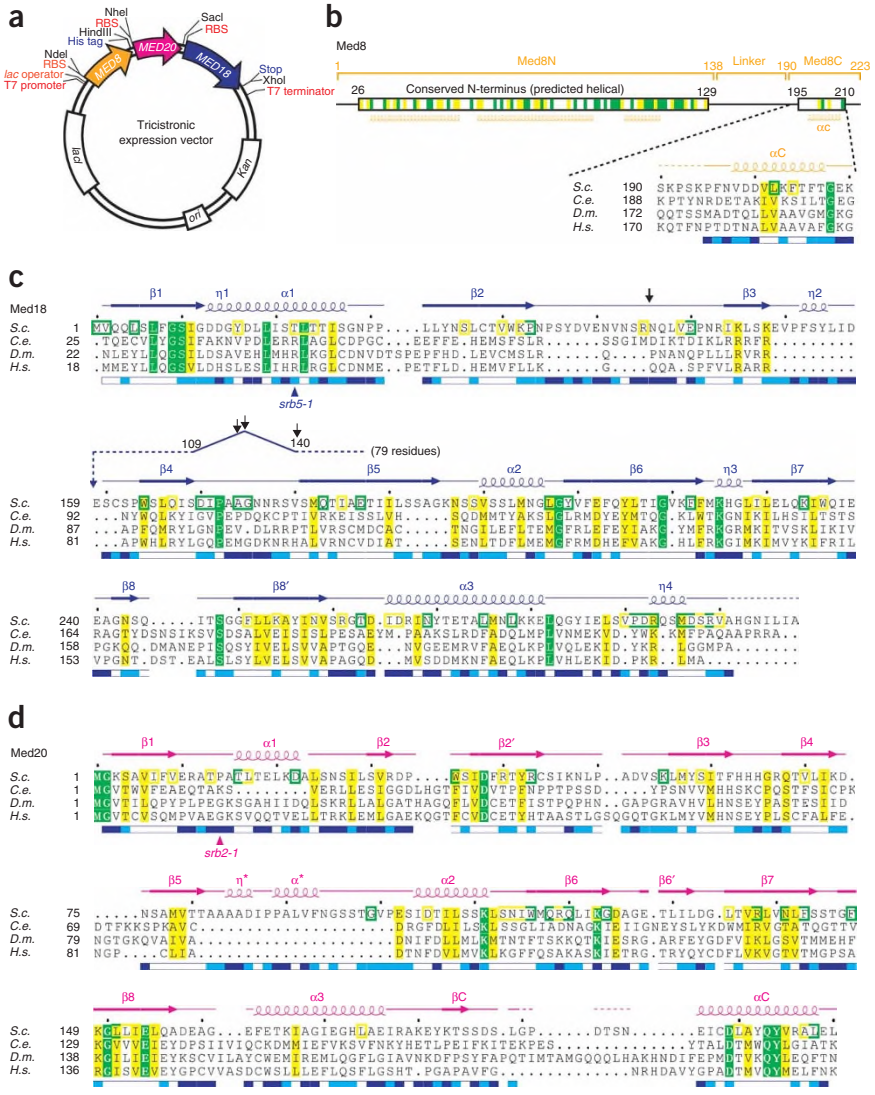


**Figure 1** Primary and secondary structure of Med8, Med18 and Med20. (**a**) Tricistronic pET24b-based vector used for heterologous expression of the trimeric Med8/18/20 Mediator submodule in *E. coli*. The ORFs are under the control of a single T7 promoter, but are each preceded by a ribosome-binding site (RBS). The trimeric complex is purified through a $His_6$ or streptactin tag located at the C terminus of Med8. Restriction sites are indicated. (**b**–**d**) Primary and secondary structure. Sequence alignments of Med8 (**b**), Med18 (**c**) and Med20 (**d**) from *S. cerevisiae* (*S.c.*), *Caenorhabditis elegans* (*C.e.*), *Drosophila melanogaster* (*D.m.*) and *Homo sapiens* (*H.s.*). For Med8, only the C-terminal Med18-binding region Med8C is shown in detail. Secondary structure elements are indicated above the sequences (spirals, α-helices and $3_{10}$(η)-helices; arrows, β-strands; lines, loops; dashed lines, disordered regions). Invariant and conserved residues are highlighted in green and yellow, respectively. Additionally, residues that are invariant or conserved among the yeast family Saccharomycotinae (*S. cerevisiae*, *Candida glabrata*, *Candida albicans*, *Ashbya gossypii*, *Kluyveromyces lactis* and *Debaryomyces hansenii*) are highlighted with green or yellow frames, respectively, on the *S.c.* sequence. Surface accessibility is indicated below the sequences (blue, exposed; white, buried). Cleavage sites revealed by limited proteolysis with trypsin, chymotrypsin or proteinase K are indicated with black arrows. Two suppressor mutations of CTD truncation are indicated by a triangle (*srb2-1*, ref. 17; *srb5-1*, ref. 8). The flexible regions in Med18 comprise residues 44–63 and 69–162. Sequence alignments were done with ClustalW[47] and figures were prepared with ESPript[48].

coexpressed with $Med18_{\Delta 109–140}$ and Med20 using a tricistronic vector in *Escherichia coli* (**Fig. 1**). A Med8 fragment comprising residues 190–210 was sufficient for copurification of the Med18/20 heterodimer. The resulting stoichiometric trimeric complex between $Med8_{190–210}$, $Med18_{\Delta 109–140}$ and Med20 (Med8C/18/20) crystallized under different conditions and in a different space group than did the Med18/20 heterodimer (Methods). The structure was solved by molecular replacement with the Med18/20 structure and was refined to a free *R*-factor of 27.2% at 2.7-Å resolution.

## Overall Med8C/18/20 structure

Med18 and Med20 form a heterodimer with pseudo two-fold symmetry (**Fig. 2**). Med18 consists of an eight-stranded β-barrel with a central pore and three flanking helices (**Fig 2a**). Med20 adopts a very similar fold but contains an additional helix at the C terminus, which fills the central pore of its barrel fold (helix αC, **Fig. 2a**). The Med20 barrel is open on one side, as if helix αC forms a wedge. The Med18–Med20 interaction is mainly hydrophobic and involves the α2 helices of both subunits and the β5 strands at the opening of the Med20 barrel (**Fig. 2a** and **Supplementary Table 1** online). The very extended heterodimer interface (3,900 Å$^2$ buried surface area) explains why the Med18-Med20 interaction was readily detected by various techniques[9,10,18,19]. The C-terminal region of Med8 forms an α-helix that binds across the top of the Med18 barrel, mainly through extensive hydrophobic contacts (**Fig. 2a**, orange, and **Supplementary Table 2** online). Binding of the Med8 helix causes local ordering of Med18 loops but otherwise does not change the free Med18/20 structure. The overall structure of the Med8C/18/20 complex is apparently conserved, as hydrophobic core residues are identical or similar from yeast to human (**Fig. 1**), and many residues in the subunit-subunit interfaces are also conserved (**Supplementary Fig. 2** online).

## Med18 resembles phosphoryl transfer enzymes

Comparison of the Med18 structure with known folds[20] identified Cet1, the triphosphatase in the yeast mRNA capping apparatus (PDB entry 1D8I)[21], and two proteins of unknown function from *Pyrococcus furiosus* (PDB entry 1YEM, **Fig. 3a**) and *Vibrio parahaemolyticus* (PDB entry 2ACA). A BLAST search with sequences of the two latter proteins revealed close similarity (45% and 67% for 1YEM and 2ACA, respectively) with adenylate cyclase-2 of *Aeromonas hydrophila*[22], identifying 1YEM and 2ACA as putative adenylate cyclases. Both putative cyclases contain a CYTH domain with a conserved N-terminal ExExK motif (where x is any residue)[23]. This motif is also found in Cet1 and forms a β-strand at the same location in all structures (**Fig. 3b**). The side chains of the two glutamates and the lysine in the ExExK motif protrude into the pore (**Fig. 3b**). The two

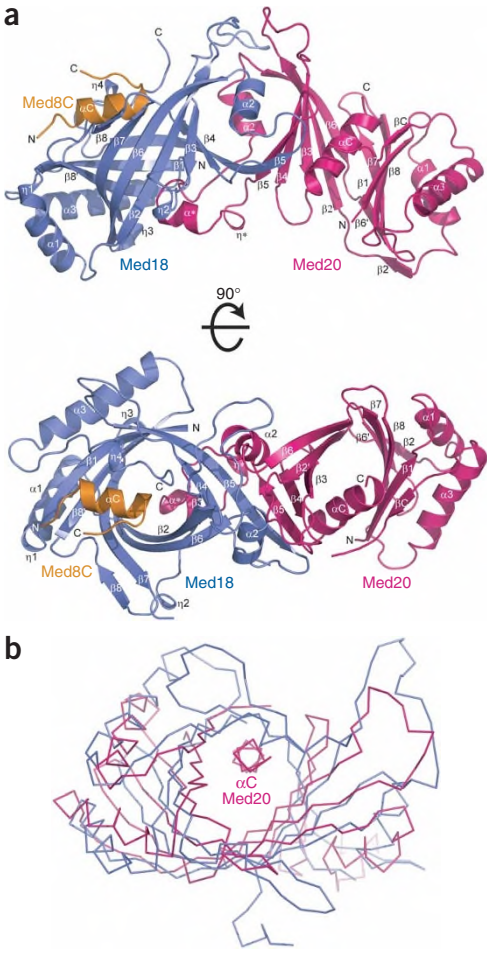


**Figure 2** Crystal structure of the Mediator head submodule Med8C/18/20. (**a**) Two views of a ribbon diagram of the trimeric complex, related by a 90° rotation around the horizontal axis. Med8C, Med18 and Med20 are in orange, blue and magenta, respectively. Secondary structure elements are labeled as in **Figure 1**. Vertical axis corresponds to a pseudo dyad of the heterodimer that relates Med18 and Med20. (**b**) Superposition of Cα traces of Med18 and Med20 reveals a similar fold. Indicated C-terminal helix of Med20, which occupies the central pore, does not exist in Med18. This figure and other structure figures were prepared with PyMOL[49].

glutamates bind a catalytic metal ion in Cet1 (ref. 21), and the lysine binds a phosphate ion in the 2ACA structure. A second phosphate ion is bound in the 2ACA structure by conserved residues outside the ExExK motif, and equivalent residues in Cet1 bind a sulfate (**Fig. 3b**). The phosphate and sulfate ions presumably mimic substrate phosphate groups. Together with previous findings[21], these structures show that the pore of Cet1 and the CYTH proteins forms a conserved active site for phosphoryl transfer reactions. In contrast, the pore of Med18 cannot constitute an active site, as residues in the pore are not conserved and the ExExK motif is absent (**Fig. 3c**). Furthermore, the Med18 pore is not accessible, as it is closed on one side by the C-terminal helix of Med8 (**Figs. 2** and **3**) and on the other side by the region including helix α1 of Med20 (**Fig. 2a**).

### Modularity of Med8 *in vitro* and *in vivo*

In an attempt to extend the Med8C/18/20 structure to a trimer containing full-length Med8, we coexpressed full-length Med8 with $Med18_{\Delta 109–140}$ and Med20 from a tricistronic vector in *E. coli*. A stoichiometric and soluble trimeric complex was obtained but could not be crystallized, indicating flexibility in Med8. Partial proteolysis and sequence analysis suggest that Med8 contains a well-conserved N-terminal helical domain, termed Med8N (**Fig. 1b**). Limited proteolysis of individually expressed Med8 followed by Edman sequencing identified a stable 14-kDa fragment, supporting the presence of an independently folded Med8N domain (data not shown). Indeed, Med8N could be purified in soluble form after its overexpression in *E. coli* (not shown). Med8N is apparently connected to Med8C by a nonconserved, proteolytically sensitive linker (**Fig. 1b**). Consistent with an exposed flexible linker, the mammalian Med8 linker sequence contains an interaction element for binding the elongin B–elongin C complex[24].

To investigate which Med8 regions are required for cell viability, we carried out complementation studies in yeast. We generated plasmids expressing, under the control of the *MED8* promoter, full-length Med8 and three Med8 variants with different C-terminal truncations (**Fig. 4**). Plasmids were introduced into a *med8Δ* strain rescued by a *MED8*-encoding *URA3* plasmid. Complementation would allow for loss of the *URA3* plasmid and growth on media containing 5-fluorotic acid (5-FOA). Complementation was observed with full-length *MED8*, with $med8_{1–189}$, which lacks Med8C, and with $med8_{1–138}$, which corresponds to Med8N (**Fig. 4a**). The plasmid encoding $med8_{1–94}$, predicted to express a disrupted Med8N domain, did not rescue cell growth (**Fig. 4a**). Thus, the conserved N-terminal domain Med8N is essential for cell viability, whereas the C-terminal Med18-binding region Med8C and the linker are not. The Med8 linker seems to have an important function *in vivo*, as the shuffle strain with $med8_{1–138}$, which lacks the linker, showed severely compromised growth (**Fig. 4a**).

### The essential Med8 N-terminal domain binds TBP

Candidate interaction partners of the Med8/18/20 subcomplex include Pol II and general transcription factors, as reconstituted recombinant Mediator head module binds TBP and TFIIB[10]. We therefore tested Med8/18/20 for its ability to bind recombinant $His_6$-tagged TBP and TFIIB immobilized on nickel–nitrilotriacetic acid (Ni-NTA) beads (**Fig. 5a**). In these assays, we detected clear binding of Med8/18/20 to TBP, but not to TFIIB, the Pol II core or the Pol II subcomplex Rpb4–Rpb7 (**Fig. 5a** and not shown), showing that the TBP interaction is specific. To investigate which region of Med8/18/20 is responsible for TBP binding, we tested Med8 and Med8C/18/20 in pull-down assays. Med8C/18/20 did not show detectable TBP binding (data not shown), providing a negative control. In contrast, free Med8 bound strongly to TBP (**Fig. 5b**). Further, recombinant Med8N was sufficient for TBP binding (**Fig. 5c**), establishing this domain as a TBP-binding submodule of the Mediator head. The protein interactions revealed by our assay are robust and specific, as we used highly soluble and pure recombinant proteins, analyzed bound proteins by direct gel staining rather than by western blotting and used a stringent washing protocol, and because all controls are free of background (Methods and **Fig. 5**).

## DISCUSSION

Following our crystal structures of the Med7/21 heterodimer in the Mediator middle module[25] and cyclin C in the kinase module[26], we provide here structural information on the head module, in particular the Med18/20 heterodimer in free form and bound to the C-terminal region of Med8. The work shows that coexpression of Mediator subunits in *E. coli* can generate crystallization-grade dimeric and trimeric subcomplexes, and probably higher-order assemblies, that are amenable to high-resolution structure determination by X-ray

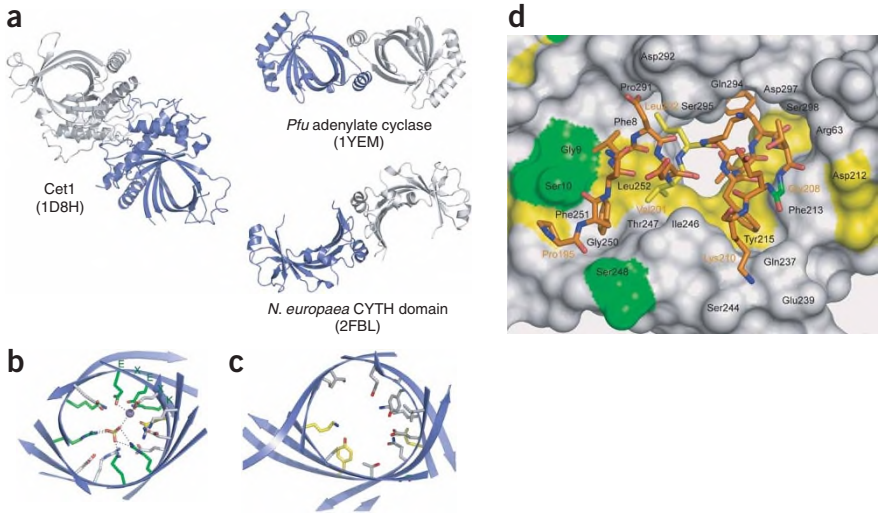


**Figure 3** Comparison of the Med18/20 heterodimer with similar known structures. (**a**) Structures with folds similar to Med18. Shown are Cet1, the triphosphatase subunit of the yeast mRNA-capping apparatus[21]; a putative adenylate cyclase of *P. furiosus* (ORF 838710-001, identity/similarity to adenylate cyclase CyaB from *A. hydrophila* 25%/46%); and the CYTH domain of a putative adenylate cyclase of *N. europaea* (NE1496). The two latter proteins share the dimerization mode of Med18/20, whereas Cet1 does not. (**b**) Chemical nature and conservation of the catalytic pore in Cet1 and CYTH domains. Charged, functionally relevant Cet1 residues are shown as sticks[21]. Green and yellow respectively indicate residues that are invariant or conserved among Cet1, the putative adenylate cyclases of *P. furiosus* and *V. parahaemolyticus*, and the CYTH domain of *N. europaea*. A manganese ion is shown as a purple sphere and a sulfate ion as a stick model. The ExExK motif is indicated. (**c**) Chemical nature and lack of conservation of the pore in Med18. Color code for sequence conservation is as in **b**. (**d**) Interaction of Med8C with the Med18 barrel. The Med8C helix is shown as orange sticks, with terminal, conserved or invariant residues labeled in orange. Med18 is shown as a white surface representation, with residues in contact with Med8C labeled in black. Green and yellow respectively indicate residues that are invariant and conserved among *S. cerevisiae*, *C. elegans*, *D. melanogaster* and *H. sapiens* (compare **Fig. 1b–d**).

analysis. The structures can later be used for the interpretation of medium-resolution X-ray crystallographic maps or cryo-EM reconstructions of larger Mediator subcomplexes and eventually Mediator–Pol II complexes. The results also enabled structure-guided studies *in vitro* and *in vivo*, thereby providing insights into the mechanism of Mediator function.

Our results indicate that the Med8/18/20 complex consists of two structurally independent submodules, Med8N and Med8C/18/20 (**Fig. 6a**). Whereas Med8N is essential for yeast viability, all three parts of the Med8C/18/20 submodule are nonessential, and deletion of any part leads to a similar growth phenotype. The $med8_{1-189}$ strain and knockout strains that lack either *MED18* or *MED20* (also called *SRB5* and *SRB2*, respectively) all showed slow growth at 30 °C and did not grow at 37 °C (**Fig. 4b**). The Med8C/18/20 submodule is apparently involved in protein interactions. Analysis of the surface conservation of Med8C/18/20 reveals that one face is well conserved, whereas the other is not (**Fig. 6b**). Both Med18 and Med20 contribute to a continuous conserved surface, consistent with a functional Med8C/18/20 submodule. Published data suggest that Med8C/18/20 is involved in interactions that have a positive effect on transcription. Med18 and Med20 are required for stable initiation complex formation[8,13,17], and Mediator–Pol II complex lacking Med18 and Med20 is defective in basal transcription[16]. Expression profiles of yeast strains lacking either *MED18* or *MED20* are nearly identical and show mainly decreased transcript levels[27], consistent with a functional Med8C/18/20 submodule that has a positive role in transcription *in vivo*.

The *srb* mutations *srb2-1* and *srb5-1* alter Med20 and Med18, respectively, at sites on the putative protein-interaction face of the Med8C/18/20 submodule; these sites are not involved in stabilizing subunit interfaces or hydrophobic cores (**Fig. 6b**). Thus, these two *srb* mutations probably affect protein interactions of the Med8C/18/20 submodule and do not have structural roles. Med18 and Med20 are products of the *SRB5* (or *MED18*) and *SRB2* (or *MED20*) genes, which were originally identified in a screen for suppressors of the negative effects of CTD truncations[7,8]. Cells lacking *MED20* require a Pol II with 20 of its 26 CTD repeats for viability[17]. However, cells grow normally with only ten CTD repeats in the presence of the *srb2-1* suppressor mutation, which changes the Med20 residue Pro14 to a histidine[17]. The CTD truncation

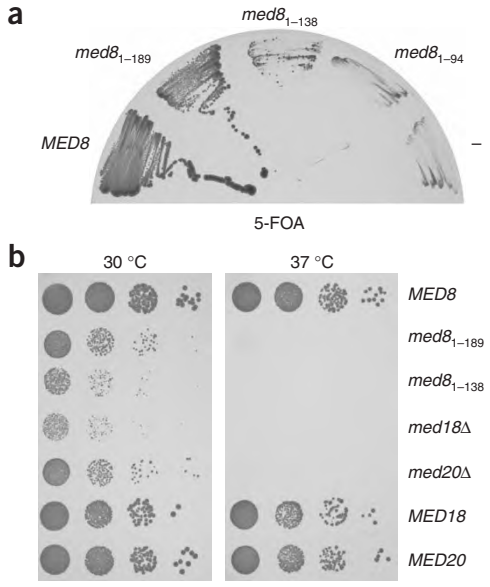


**Figure 4** Modularity of Med8 *in vivo*. (**a**) Complementation studies reveal that Med8N is essential for cell viability. Plasmids encoding full-length *MED8* or C-terminal truncations of Med8 expressed from the *MED8* promoter were transformed into the *MED8* shuffle strain and streaked onto 5-FOA–containing plates to shuffle out the *MED8*-encoding *URA3* plasmid. Growth indicates expression of a functional Med8 protein. Yeast cells carrying only Med8N and the Med8 linker (residues 1–189, Med8C deleted) or only Med8N (residues 1–138) are viable, whereas a disrupted Med8N domain (residues 1–94) cannot rescue cell growth. (**b**) Comparison of growth phenotypes of the strains in **a**. Ten-fold serial dilutions of the strains were spotted onto SDC (–Leu) plates and incubated at the indicated temperatures. Deletion of Med8C or Med8C and the Med8 linker leads to a slow-growth phenotype and causes temperature sensitivity. Cells lacking *MED18* or *MED20* show a similar growth phenotype.

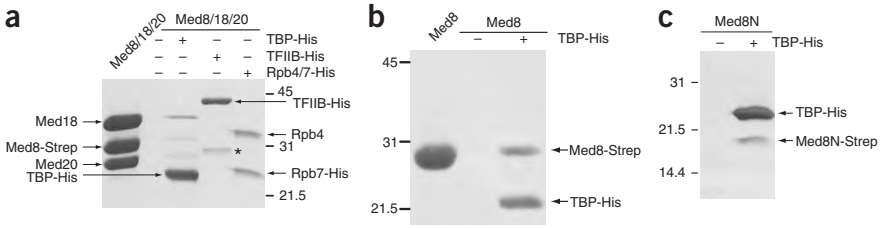


**Figure 5** TBP binding of Med8/18/20, Med8 and Med8N. (**a**) Binding of Med8/18/20 to $His_6$-tagged TBP, TFIIB and Rpb4–Rpb7 (Rpb4/7) in pull-down assays. Binding is observed only to TBP. Gels were stained with Coomassie blue. Left lane contains the total input. Asterisk marks a known C-terminal TFIIB fragment. See Methods for details. (**b**) Med8 binds $His_6$-tagged TBP in pull-down assays. The left lane corresponds to the total input. (**c**) The N-terminal domain of Med8, Med8N (residues 1–138), is sufficient for TBP binding. This gel was stained with silver to enhance the weaker Coomassie staining of the short Med8N variant.

phenotypes are also suppressed by the *srb5-1* mutation, which changes the Med18 residue Thr22 to isoleucine[8]. Although the physical basis for rescue of CTD truncation defects by *srb* mutations remains yet to be understood, our data indicate that *srb* mutations can influence the strength of protein interactions within the initiation complex.

The second submodule in the Med8/18/20 subcomplex, Med8N, binds TBP *in vitro* and has an essential function *in vivo*. Sequence analysis reveals that Med8N, together with Med6, forms the most highly conserved region of the Mediator head module (37% sequence similarity between yeast residues 26–129 and human). Protein solubility studies, limited proteolysis and secondary structure prediction suggest that Med8N forms a structured helical domain. Thus, Med8N is a conserved, essential TBP-binding submodule. As Med20 has previously been shown to interact directly with TBP[17], the TBP-binding site of the Mediator head is apparently multipartite and includes Med8N, Med20 and possibly additional parts of the head module. We did not detect an interaction between TBP and the Med8C/18/20 submodule, which is apparently inconsistent with the published Med20-TBP interaction. However, the sensitivity of our assay is much lower than that of the published protocol, which used western blotting to reveal the protein interaction[17]. Thus, our results do not exclude a Med20-TBP interaction, but they suggest that it is weaker than the Med8-TBP interaction. A direct Mediator head–TBP interaction could partially account for maintenance of a stable scaffold complex that remains at the promoter for multiple rounds of transcription[28]. The Mediator head–TBP interaction is also consistent with a defect in initiation-complex assembly at the promoter in extracts of yeast strains with deletions of *MED18* or *MED20* (ref. 13) and with genome-wide Mediator interactions with promoters that correlate with increased transcription activity[29,30].

Published results show that the Mediator-TBP interaction depends on an activator. Binding of the yeast Pol II–Mediator complex to TBP requires that the activator Gal4VP16 is prebound[16]. Gal4VP16 also stabilizes assembly of human Mediator and TBP on DNA[31], and human Mediator enhances TBP-DNA binding in an activator-dependent manner[32]. Although these results can be explained by simultaneous binding of activator to TBP and Mediator, we suggest an alternative model, namely that activator binding to Mediator results in a modulation and exposure of Mediator's TBP-binding site. The idea of activator-induced TBP site exposure on Mediator is consistent with the small peptide–like size of many activation domains, with recent genome-wide Mediator localization studies that query recruitment models of regulation[30] and with the electron-microscopic observation of activator-induced large-scale conformational changes in Mediator[33]. Structural changes in Mediator could rely on hinge motions[25] or on a relative repositioning of submodules connected by flexible linkers (**Fig. 6b**). In addition, some folds of Mediator subunits may undergo structural changes. For example, the Med20 barrel may open, as observed for the barrel in ORF1496 of *Nitrosomonas europaea* (PDB entry 2FBL; **Fig. 3a**), as it shows some flexibility in our crystals (**Supplementary Fig. 3** online).

The Med8/18/20 subcomplex is anchored to the remainder of the head module by contacts of Med8 and Med20 to subunit Med17 (refs. 9,10,18,19). The strength of the Med20-Med17 interaction is, however, insufficient to retain Med20 within Mediator, as Mediator–Pol II complex purified from yeast *MED18* knockout cells lacks both Med18 and Med20 (ref. 16). Accordingly, *in vitro* transcription with an extract from a *med18*Δ strain requires addition of both Med18 and Med20 to restore basal transcription[8]. In contrast, transcription with an extract from a *med20*Δ strain requires addition only of Med20 (ref. 17,34), showing that the Med8-Med17 interaction is sufficient to retain Med18 in the Mediator complex. Taken together, the Med8-Med17 contact seems to be mainly responsible for tethering

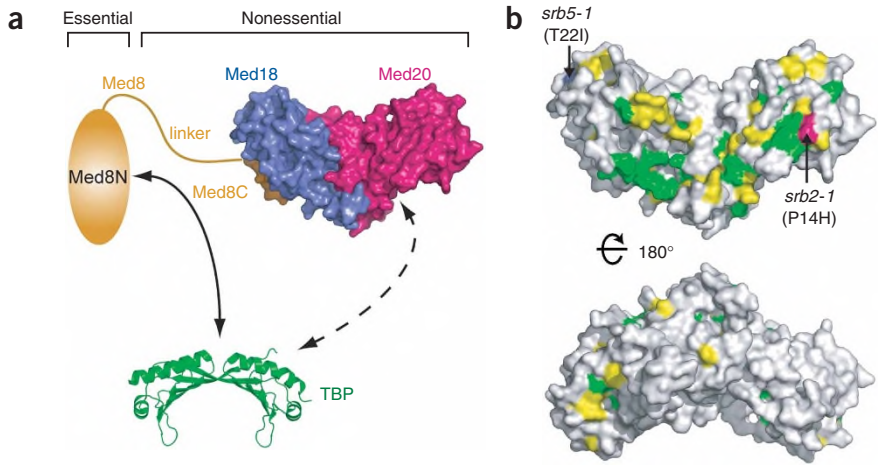


**Figure 6** Med8/18/20 architecture and surface conservation. (**a**) Modular architecture of the Med8/18/20 subcomplex. Two submodules can be distinguished that correspond to essential and nonessential regions *in vivo*. The Med8N-TBP interaction is indicated with a double-headed arrow. A previously described[17] and apparently weaker Med20-TBP interaction is indicated with a dashed double-headed arrow. (**b**) Two views of a surface representation of the Med8C/18/20 submodule, related by a 180° rotation around the horizontal axis. Residues invariant and conserved among Saccharomycotinae are highlighted in green and yellow, respectively (compare **Fig. 1b–d**). Residues affected by *srb2-1* and *srb5-1* mutations[8,17] are in magenta and blue, respectively.

the Med8/18/20 subcomplex to the remainder of the head module. The transient Med20-Med17 interaction could be relevant for a modulation of Mediator's TBP site. For example, the *srb2-1* mutation affecting the Med20 surface could influence the strength of the Med20-Med17 interaction and change TBP-site accessibility (**Fig. 6b**).

## METHODS

**Cloning.** We modified pET-24b vectors for bicistronic expression of Med18/20 or tricistronic expression of Med8/18/20 under the control of one T7 promoter (**Fig. 1a**). Genes encoding the Mediator subunits were inserted into the vector in the order Med20-Med18 or Med8-Med20-Med18, and ribosomal binding sites and new restriction sites were added (**Fig. 1a**). For expression of the Med18/20 heterodimer, $Med18_{\Delta 109–140}$ was fused to a C-terminal $His_6$ tag. For expression of the Med8C/18/20 for crystallization, $Med8_{190–210}$ was fused to a C-terminal $His_6$ tag and a stop codon was added at the end of the $Med18_{\Delta 109–140}$ gene sequence. For expression of Med8/18/20 or Med8C/18/20 for pull-down experiments, the $His_6$ tag was replaced by a Strep tag II streptactin tag and full-length Med18 was used. For expression of Med8 or Med8N, the corresponding genes were fused to a C-terminal Strep tag II and cloned into a pET-21b vector using the restriction sites NdeI and HindIII.

**Protein expression and purification.** For heterologous expression, transformed *E. coli* BL21 (DE3) RIL cells (Stratagene) were grown in LB medium at 37 °C to an $A_{600}$ of 0.5. Expression was induced with 0.5 mM IPTG for 16 h at 18 °C. For selenomethionine labeling of Med18/20, expression was carried out as described[35,36]. For protein purification, cells were lysed by sonication in buffer A (20 mM Tris (pH 8.0), 150 mM NaCl, 10 mM β-mercaptoethanol). After centrifugation, the supernatant was loaded onto a Ni-NTA column (Qiagen) equilibrated with buffer A. The column was washed with 40 ml of buffer A and, for the trimeric complex, was further washed sequentially with 40 ml buffer A containing 20 mM and 50 mM imidazole. Bound proteins were eluted with buffer A containing 300 mM imidazole. For purification of Strep-tagged proteins, cells were lysed by sonication in buffer B (100 mM Tris (pH 8.0), 150 mM NaCl, 10 mM β-mercaptoethanol). After centrifugation, the supernatant was loaded onto a 1-ml Strep-Tactin column (IBA Tools) equilibrated with buffer B. The column was washed six times with one column volume of buffer B and bound proteins were eluted with buffer B containing 2.5 mM desthiobiotin. All complexes were further purified by anion exchange chromatography (Mono Q, Amersham Biosciences). The column was equilibrated with buffer C (20 mM Tris (pH 8.0), 100 mM NaCl, 5 mM dithiothreitol) and proteins were eluted with a linear gradient of 10 column volumes from 100 mM to 1 M NaCl in buffer C. After concentration, the sample was applied to a Superose-12 size-exclusion column (Amersham) equilibrated with buffer C containing 150 mM NaCl. For crystallization, the complexes were concentrated to 20 mg $ml^{-1}$. Rpb4–Rpb7 and TBP were purified as described[37,38] and TFIIB was purified by Ni-NTA and ion-exchange chromatography (Mono S, Amersham) (K. Armache & P.C., unpublished data).

**Crystallization and X-ray structure determinations.** Crystals of the Med18/20 heterodimer were grown at 20 °C in hanging drops over reservoirs containing 100 mM Tris (pH 8.5) and a 1.8 M mixture of sodium phosphate and potassium phosphate. Crystals of the trimeric complex Med8C/18/20 were grown at 20 °C in sitting drops over a reservoir containing 100 mM sodium acetate (pH 4.6) and 0.6 M sodium fluoride. Crystals were harvested by gradually adding glycerol to the mother solution to a final concentration of 25% (v/v). Crystals were flash-cooled in liquid nitrogen. Diffraction data were collected at 100 K on a MARCCD 225 detector at the Swiss Light Source, Villigen, Switzerland.

For the labeled dimeric complex Med18/20, diffraction data were processed with MOSFLM and SCALA[39]. SOLVE[40] identified 31 selenium sites, which were used for phasing. Solvent flattening, noncrystallographic symmetry averaging and building of an initial protein model were done with RESOLVE[40]. The resulting electron density map allowed for building most of Med18 and Med20 using TURBO-FRODO[41]. The model was refined using the conjugate gradient in the program CNS[42].

For the trimeric complex Med8C/18/20, diffraction data were processed with DENZO and SCALEPACK[43]. The structure was phased by molecular replacement with PHASER[44], which positioned four copies of the refined Med18/20 structure in the asymmetric unit. Model bias was removed with high-temperature simulated annealing in CNS. A model-phased difference Fourier map revealed density for the Med8 C-terminal region and some changes in the Med18/20 structure. After rebuilding, the new model was refined with CNS. Owing to the limited resolution and high *B*-factors in the Med20 structure, noncrystallographic symmetry restraints were maintained during the refinement process, except in clearly divergent areas. Refinement using TLS in REFMAC[45] slightly improved statistics but did not substantially change the coordinates of the structures. For consistency, the CNS-refined models were kept. Both refined structures have low *R*-factors and good stereochemistry (**Table 1**). Ramachandran plots of the Med18/20 and Med8C/18/20 models show 86.3% and 85.5% of the residues, respectively, in the most favored region and none in disallowed regions[46].

**Table 1 Data collection and refinement statistics for Med18/20 and Med8C/18/20**

| | Med18/20 | Med8C/18/20 |
|---|---|---|
| **Data collection** | | |
| Space group | $P2_12_12_1$ | $P2_1$ |
| Cell dimensions | | |
| *a*, *b*, *c* (Å) | 72.9, 129.4, 241.7 | 75.4, 115.8, 129.2 |
| α, β, γ (°) | 90, 90, 90 | 90, 98.5, 90 |
| Resolution (Å) | 20–2.4 (2.53–2.4) | 20–2.7 (2.8–2.7) |
| $R_{sym}$ | 7.3 (37.5) | 7.5 (46.3) |
| $I / \sigma I$ | 25.2 (5.6) | 19.2 (2.1) |
| Completeness (%) | 98.8 (95.1) | 99.9 (99.1) |
| Redundancy | 13.2 (10.4) | 6.4 (5.5) |
| **Refinement** | | |
| Resolution (Å) | 20–2.4 | 20–2.7 |
| No. reflections | 89,181 | 61,548 |
| $R_{work}$ / $R_{free}$ | 22.8 / 26.4 | 23.6 / 27.2 |
| No. atoms | | |
| Protein | 13,358 | 14,013 |
| Ligand/ion | 4 | – |
| Water | 230 | 85 |
| *B*-factors | | |
| Protein | 58.9 | 64.2 |
| Ligand/ion | 45.9 | – |
| Water | 42.1 | 43.4 |
| R.m.s. deviations | | |
| Bond lengths (Å) | 0.008 | 0.008 |
| Bond angles (°) | 1.371 | 1.261 |

Highest-resolution shell is shown in parentheses.

**Yeast manipulations and growth assays.** Plasmids pRS316-*MED8*, pRS316-*MED20* and pRS316-*MED18* were constructed by amplifying the respective ORF plus 500 base pairs (bp) upstream and 300 bp downstream by PCR, generating BamHI and XhoI sites for *MED8* and *MED20* and BamHI and HindIII for *MED18*, and cloning the fragments into the same sites of pRS316. pRS315-*MED8*, pRS315-$med8_{1–189}$, pRS315-$med8_{1–138}$ and pRS315-$med8_{1–94}$ were constructed by amplifying the *MED8* sequence encoding amino acid residues 1–223, 1–189, 1–138 and 1–94, respectively, plus 500 bp upstream, generating BamHI and XhoI sites, and cloning the fragments into the same sites of pRS315-$T_{ADH1}$ (S.R. and K.S., unpublished data). *MED8*/*med8*Δ, *MED18*/*med18*Δ and *MED20*/*med20*Δ heterozygous knockout strains were obtained from Euroscarf and transformed with pRS316-*MED8*, pRS316-*MED18* and pRS316-*MED20*, respectively. Diploids were sporulated and tetrads dissected on YPD plates. To assess the functionality of pRS315-*MED8*, pRS315-$med8_{1–189}$, pRS315-$med8_{1–134}$ and pRS315-$med8_{1–94}$, plasmids were transformed into the *MED8* shuffle strain. Transformants were restreaked onto 5-FOA. After 6 d on 5-FOA, grown colonies were restreaked on SDC (–Leu). Equal amounts of

freshly grown yeast cells on SDC (–Leu) plates were resuspended in water and ten-fold dilutions were spotted on SDC (–Leu) plates.

**Protein interaction assays.** Recombinant Med8/18/20, Med8C/18/20, Med8 or Med8N carrying a Strep tag II (40 μg) was mixed with a stoichiometric amount of $His_6$-tagged TBP, TFIIB or Rpb4–Rbp7 in 50 μl of binding buffer (20 mM Tris (pH 8.0), 150 mM NaCl, 50 mM ammonium acetate, 10 mM β-mercaptoethanol). The mixtures were incubated for 1 h at room temperature before 15 μl of Ni-NTA beads (Qiagen) were added. The beads were washed four times with 500 μl of binding buffer containing 0.2% (v/v) Nonidet P40 and 10% (v/v) glycerol and were boiled in SDS sample buffer. Bound proteins were analyzed by SDS-PAGE and stained with Coomassie blue (or with silver in the case of Med8N, which stains weakly with Coomassie).

**Accession codes.** Protein Data Bank: Coordinates have been deposited with accession codes 2HZM (Med18/20) and 2HZS (Med8C/18/20).

*Note: Supplementary information is available on the Nature Structural & Molecular Biology website.*


ACKNOWLEDGMENTS

We thank S. Baumli for discussions, K. Armache for help with TFIIB preparation and other members of the Cramer laboratory for help. This work was supported by a European Molecular Biology Organization long-term fellowship to L.L., by grants of the Deutsche Forschungsgemeinschaft, the Sonderforschungsbereich SFB646 and the Fonds der chemischen Industrie to P.C. and K.S. and by the EU-grant 3D repertoire, contract no. LSHG-CT-2005-512028. Part of this work was performed at the Swiss Light Source at the Paul Scherrer Institute, Villigen, Switzerland. We thank C. Schulze-Briese and his team at the Swiss Light Source for help.


COMPETING INTERESTS STATEMENT

The authors declare that they have no competing financial interests.


1. Malik, S. & Roeder, R.G. Transcriptional regulation through Mediator-like coactivators in yeast and metazoan cells. *Trends Biochem. Sci.* **25**, 277–283 (2000).
2. Bjorklund, S. & Gustafsson, C.M. The mediator complex. *Adv. Protein Chem.* **67**, 43–65 (2004).
3. Naar, A.M., Lemon, B.D. & Tjian, R. Transcriptional coactivator complexes. *Annu. Rev. Biochem.* **70**, 475–501 (2001).
4. Kornberg, R.D. Mediator and the mechanism of transcriptional activation. *Trends Biochem. Sci.* **30**, 235–239 (2005).
5. Bourbon, H.M. *et al.* A unified nomenclature for protein subunits of mediator complexes linking transcriptional regulators to RNA polymerase II. *Mol. Cell* **14**, 553–557 (2004).
6. Boube, M., Joulia, L., Cribbs, D.L. & Bourbon, H.M. Evidence for a mediator of RNA polymerase II transcriptional regulation conserved from yeast to man. *Cell* **110**, 143–151 (2002).
7. Nonet, M.L. & Young, R.A. Intragenic and extragenic suppressors of mutations in the heptapeptide repeat domain of *Saccharomyces cerevisiae* RNA polymerase II. *Genetics* **123**, 715–725 (1989).
8. Thompson, C.M., Koleske, A.J., Chao, D.M. & Young, R.A. A multisubunit complex associated with the RNA polymerase II CTD and TATA-binding protein in yeast. *Cell* **73**, 1361–1375 (1993).
9. Guglielmi, B. *et al.* A high resolution protein interaction map of the yeast Mediator complex. *Nucleic Acids Res.* **32**, 5379–5391 (2004).
10. Kang, J.S. *et al.* The structural and functional organization of the yeast mediator complex. *J. Biol. Chem.* **276**, 42003–42010 (2001).
11. Dotson, M.R. *et al.* Structural organization of yeast and mammalian mediator complexes. *Proc. Natl. Acad. Sci. USA* **97**, 14307–14310 (2000).
12. Cantin, G.T., Stevens, J.L. & Berk, A.J. Activation domain-mediator interactions promote transcription preinitiation complex assembly on promoter DNA. *Proc. Natl. Acad. Sci. USA* **100**, 12003–12008 (2003).
13. Ranish, J.A., Yudkovsky, N. & Hahn, S. Intermediates in formation and activity of the RNA polymerase II preinitiation complex: holoenzyme recruitment and a postrecruitment role for the TATA box and TFIIB. *Genes Dev.* **13**, 49–63 (1999).
14. Holstege, F.C. *et al.* Dissecting the regulatory circuitry of a eukaryotic genome. *Cell* **95**, 717–728 (1998).
15. Takagi, Y. & Kornberg, R.D. Mediator as a general transcription factor. *J. Biol. Chem.* **281**, 80–89 (2006).
16. Lee, Y.C., Park, J.M., Min, S., Han, S.J. & Kim, Y.J. An activator binding module of yeast RNA polymerase II holoenzyme. *Mol. Cell. Biol.* **19**, 2967–2976 (1999).
17. Koleske, A.J., Buratowski, S., Nonet, M. & Young, R.A. A novel transcription factor reveals a functional link between the RNA polymerase II CTD and TFIID. *Cell* **69**, 883–894 (1992).
18. Koh, S.S., Ansari, A.Z., Ptashne, M. & Young, R.A. An activator target in the RNA polymerase II holoenzyme. *Mol. Cell* **1**, 895–904 (1998).
19. Lee, T.I. *et al.* Interplay of positive and negative regulators in transcription initiation by RNA polymerase II holoenzyme. *Mol. Cell. Biol.* **18**, 4455–4462 (1998).
20. Holm, L. & Sander, C. Dali: a network tool for protein structure comparison. *Trends Biochem. Sci.* **20**, 478–480 (1995).
21. Lima, C.D., Wang, L.K. & Shuman, S. Structure and mechanism of yeast RNA triphosphatase: an essential component of the mRNA capping apparatus. *Cell* **99**, 533–543 (1999).
22. Sismeiro, O., Trotot, P., Biville, F., Vivares, C. & Danchin, A. *Aeromonas hydrophila* adenylyl cyclase 2: a new class of adenylyl cyclases with thermophilic properties and sequence similarities to proteins from hyperthermophilic archaebacteria. *J. Bacteriol.* **180**, 3339–3344 (1998).
23. Iyer, L.M. & Aravind, L. The catalytic domains of thiamine triphosphatase and CyaB-like adenylyl cyclase define a novel superfamily of domains that bind organic phosphates. *BMC Genomics* **3**, 33 (2002).
24. Brower, C.S. *et al.* Mammalian mediator subunit mMED8 is an Elongin BC-interacting protein that can assemble with Cul2 and Rbx1 to reconstitute a ubiquitin ligase. *Proc. Natl. Acad. Sci. USA* **99**, 10353–10358 (2002).
25. Baumli, S., Hoeppner, S. & Cramer, P. A conserved mediator hinge revealed in the structure of the MED7/MED21 (Med7/Srb7) heterodimer. *J. Biol. Chem.* **280**, 18171–18178 (2005).
26. Hoeppner, S., Baumli, S. & Cramer, P. Structure of the mediator subunit cyclin C and its implications for CDK8 function. *J. Mol. Biol.* **350**, 833–842 (2005).
27. van de Peppel, J. *et al.* Mediator expression profiling epistasis reveals a signal transduction pathway with antagonistic submodules and highly specific downstream targets. *Mol. Cell* **19**, 511–522 (2005).
28. Yudkovsky, N., Ranish, J.A. & Hahn, S. A transcription reinitiation intermediate that is stabilized by activator. *Nature* **408**, 225–229 (2000).
29. Zhu, X. *et al.* Genome-wide occupancy profile of mediator and the Srb8–11 module reveals interactions with coding regions. *Mol. Cell* **22**, 169–178 (2006).
30. Andrau, J.C. *et al.* Genome-wide location of the coactivator mediator: Binding without activation and transient Cdk8 interaction on DNA. *Mol. Cell* **22**, 179–192 (2006).
31. Johnson, K.M. & Carey, M. Assembly of a mediator/TFIID/TFIIA complex bypasses the need for an activator. *Curr. Biol.* **13**, 772–777 (2003).
32. Wu, S.Y., Zhou, T. & Chiang, C.M. Human mediator enhances activator-facilitated recruitment of RNA polymerase II and promoter recognition by TATA-binding protein (TBP) independently of TBP-associated factors. *Mol. Cell. Biol.* **23**, 6229–6242 (2003).
33. Taatjes, D.J., Naar, A.M., Andel, F., III, Nogales, E. & Tjian, R. Structure, function, and activator-induced conformations of the CRSP coactivator. *Science* **295**, 1058–1062 (2002).
34. Liu, Y., Ranish, J.A., Aebersold, R. & Hahn, S. Yeast nuclear extract contains two major forms of RNA polymerase II mediator complexes. *J. Biol. Chem.* **276**, 7169–7175 (2001).
35. Budisa, N. *et al.* High-level biosynthetic substitution of methionine in proteins by its analogs 2-aminohexanoic acid, selenomethionine, telluromethionine and ethionine in *Escherichia coli*. *Eur. J. Biochem.* **230**, 788–796 (1995).
36. Meinhart, A., Blobel, J. & Cramer, P. An extended winged helix domain in general transcription factor E/IIEalpha. *J. Biol. Chem.* **278**, 48267–48274 (2003).
37. Juo, Z.S., Kassavetis, G.A., Wang, J., Geiduschek, E.P. & Sigler, P.B. Crystal structure of a transcription factor IIIB core interface ternary complex. *Nature* **422**, 534–539 (2003).
38. Armache, K.-J., Kettenberger, H. & Cramer, P. Architecture of the initiation-competent 12-subunit RNA polymerase II. *Proc. Natl. Acad. Sci. USA* **100**, 6964–6968 (2003).
39. Leslie, A. in *Joint CCP4 and ESF-EACMB Newsletter on Protein Crystallography* No. 26 (Daresbury Laboratory, Warrington, UK, 1992).
40. Terwilliger, T.C. Automated structure solution, density modification and model building. *Acta Crystallogr. D Biol. Crystallogr.* **58**, 1937–1940 (2002).
41. Roussel, A. & Cambillau, C. Turbo-FRODO. in *Silicon Graphics Geometry, Partners Directory* 77–78 (Silicon Graphics, Mountain View, California, USA, 1989).
42. Brunger, A.T. *et al.* Crystallography & NMR system: a new software suite for macromolecular structure determination. *Acta Crystallogr. D Biol. Crystallogr.* **54**, 905–921 (1998).
43. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Meth. Enzym.* **276**, 307–326 (1996).
44. McCoy, A.J., Grosse-Kunstleve, R.W., Storoni, L.C. & Read, R.J. Likelihood-enhanced fast translation functions. *Acta Crystallogr. D Biol. Crystallogr.* **61**, 458–464 (2005).
45. Collaborative Computational Project, Number 4. The CCP4 Suite: programs for protein crystallography. *Acta Crystallogr. D Biol. Crystallogr* **50**, 760–763 (1994).
46. Laskowski, R.A., MacArthur, M.W., Moss, D.S. & Thornton, J.M. PROCHECK: a program to check the stereochemical quality of protein structures. *J. Appl. Crystallogr.* **26**, 283–291 (1993).
47. Thompson, J.D., Higgins, D.G. & Gibson, T.J. CLUSTAL W: improving the sensibility of progressive multiple sequence alignment through sequence weighing, positions-specific gap penalties and weight matrix choice. *Nucleic Acids Res.* **22**, 4673–4680 (1994).
48. Gouet, P., Courcelle, E., Stuart, D.I. & Metoz, F. ESPript: analysis of multiple sequence alignments in PostScript. *Bioinformatics* **15**, 305–308 (1999).
49. DeLano, W.L. *The PyMOL Molecular Graphics System* (DeLano Scientific, San Carlos, California, USA, 2002).