LETTER

doi:10.1038/nature10162

Architecture of the Mediator head module

Tsuyoshi Imasaki[1], Guillermo Calero[2]†, Gang Cai[3]†, Kuang-Lei Tsai[3], Kentaro Yamada[1], Francesco Cardelli[1], Hediye Erdjument-Bromage[4], Paul Tempst[4], Imre Berger[5], Guy Lorch Kornberg[2], Francisco J. Asturias[3], Roger D. Kornberg[2] & Yuichiro Takagi[1]

Mediator is a key regulator of eukaryotic transcription[1], connecting activators and repressors bound to regulatory DNA elements with RNA polymerase II[1–4] (Pol II). In the yeast *Saccharomyces cerevisiae*, Mediator comprises 25 subunits with a total mass of more than one megadalton (refs 5, 6) and is organized into three modules, called head, middle/arm and tail[7–9]. Our understanding of Mediator assembly and its role in regulating transcription has been impeded so far by limited structural information. Here we report the crystal structure of the essential Mediator head module (seven subunits, with a mass of 223 kilodaltons) at a resolution of 4.3 ångströms. Our structure reveals three distinct domains, with the integrity of the complex centred on a bundle of ten helices from five different head subunits. An intricate pattern of interactions within this helical bundle ensures the stable assembly of the head subunits and provides the binding sites for general transcription factors and Pol II. Our structural and functional data suggest that the head module juxtaposes transcription factor IIH and the carboxy-terminal domain of the largest subunit of Pol II, thereby facilitating phosphorylation of the carboxy-terminal domain of Pol II. Our results reveal architectural principles underlying the role of Mediator in the regulation of gene expression.

In the yeast *S. cerevisiae*, the Mediator head module is composed of seven subunits[10]: Med17 (also known as Srb4), Med11, Med22 (Srb6), Med6, Med8, Med18 (Srb5) and Med20 (Srb2). Four subunits are encoded by *SRB* genes, first identified through a genetic screen for mutations suppressing the Pol II carboxy-terminal domain (CTD) truncation[11,12]. The head module is essential for Mediator function because mutations in the head abolish messenger RNA synthesis *in vivo*[13,14] and *in vitro*[15], and eliminate Mediator interaction with promoters *in vivo*[10]. The head module is organized into three domains that can undergo significant conformational changes, and it interacts with the TATA-binding protein subunit of general transcription factor TFIID and the Rpb4 and Rpb7 subunits of Pol II (ref. 16). The head has also been shown to interact with TFIIH through the Med11 subunit[17]. Determining the architecture of the Mediator head module is therefore vital to understanding the mechanism by which Mediator controls gene expression.

We engineered the head module to obtain crystals of sufficient quality for structure determination (Supplementary Information, section 1). In our engineered Mediator head, Med18 loop regions and the amino-terminal 108 residues of Med17 were deleted, without apparent effect on the integrity of the complex (Supplementary Fig. 1). The modified head module was labelled with selenomethionine (SeMet) and purified as described previously[16]. By overcoming two major technical obstacles (Supplementary Information, section 2), we produced SeMet crystals that diffract to 4.3 Å (Supplementary Table 1). The electron density map was calculated to a resolution of 4.3 Å (Supplementary Fig. 2) by SeMet single anomalous dispersion (SAD) after initial phases had been obtained using Ta_6Br_{14} and $K_3Ir(NO_3)_6$ derivatives (Supplementary Information, section 3).

We began identification of the individual polypeptide constituents of the Mediator head module by docking the Med18–Med20–Med8 C-terminal helix (CTH) complex structure[18] (Protein Data Bank ID, 2HZS) into the electron density map and then performing rigid-body refinement. The polypeptide chains of the other subunits were identified on the basis of the SeMet positions and their juxtaposition with large amino-acid side chains within ordered regions of secondary structure (Methods). This approach permitted the unambiguous assignment of all discernible elements of secondary structure in the density map to individual head module subunits (Fig. 1 and Supplementary Figs 3–5).

Our crystal structure is consistent with the molecular envelope of the head module derived at a resolution of 30–35 Å by single-particle electron microscopy analysis (Supplementary Figs 6 and 7). The head can be described in terms of three major domains, a 'fixed jaw', a 'movable jaw' and a 'neck' (Fig. 1 and Supplementary Figs 4 and 5), with a 'central joint' connecting these domains. Our X-ray structure of the head module reveals the overall architecture of the module and the domain boundaries. The domains are connected through flexible loops and linkers at the central joint.

Our previous work on expression and purification of the head module suggested that Med17 has a central role in head assembly[10]. The work we report here extends those results through a comprehensive biochemical analysis in combination with electron microscopy, to determine the Med17 domain structure and elucidate its interactions with other head components (Supplementary Information, section 4). The results support our model of the architecture of the head module.

Assembly of the head module starts with formation of the 'mini-head' (Med17–Med11–Med22). Subsequently, Med8 and Med6 are added, followed by Med20–Med18 (ref. 10). Our structure shows that a four-helix bundle, built by α-helices from Med11 (BH1 and BH2) and Med22 (BH1 and BH2) interact with BH2 of Med17 to form the larger helical bundle (Figs 2 and 3 and Supplementary Fig. 4). This is consistent with the observation that omission of either Med11 or Med22 leads to disassembly of the head[10]. Med6 interacts with the mini-head through its BH1, and Med8 serves to stabilize the central α-helical bundle by surrounding the central helices. Finally, the Med18–Med20 heterodimer binds to the core-head, which is composed of five subunits (Med6, Med8, Med11, Med22 and Med17), primarily through the CTH of Med8 (Fig. 2).

The fixed jaw domain comprises the CTHs of Med11 and Med22 and the CTD of Med17. The Med11 and Med22 CTHs interact with the helical regions of the Med17 CTD. Med17 (residues 610–660) forms a β-sheet structure that lines the inner surface of the fixed jaw and faces the movable jaw (Fig. 3a). The Med17 CTD interacts with the loop region of Med18. The functional importance of the Med17 CTD correlates with the biochemical activity of the Head module *in vitro*, as well as phenotypic analysis *in vivo*, as loss of the Med17 CTD abolishes the transcription activity of the head module (Supplementary Fig. 12),

[1]Department of Biochemistry and Molecular Biology, Indiana University School of Medicine, 635 Barnhill Drive, Indianapolis, Indiana 46202, USA. [2]Department of Structural Biology, Stanford University School of Medicine, Stanford, California 94350, USA. [3]Department of Cell Biology, The Scripps Research Institute, 10550 North Torrey Pines Road, La Jolla, California 92037, USA. [4]Molecular Biology Program, Memorial Sloan-Kettering Cancer Center, New York, New York 10021, USA. [5]European Molecular Biology Laboratory, Grenoble Outstation, 6 rue Jules Horowitz, 38042 Grenoble Cedex 9, France. †Present addresses: Department of Structural Biology, University of Pittsburgh School of Medicine, Pittsburgh, Pennsylvania 15260, USA (G. Calero); School of Life Sciences, University of Science and Technology of China, Hefei, Anhui 230027, China (G. Cai).

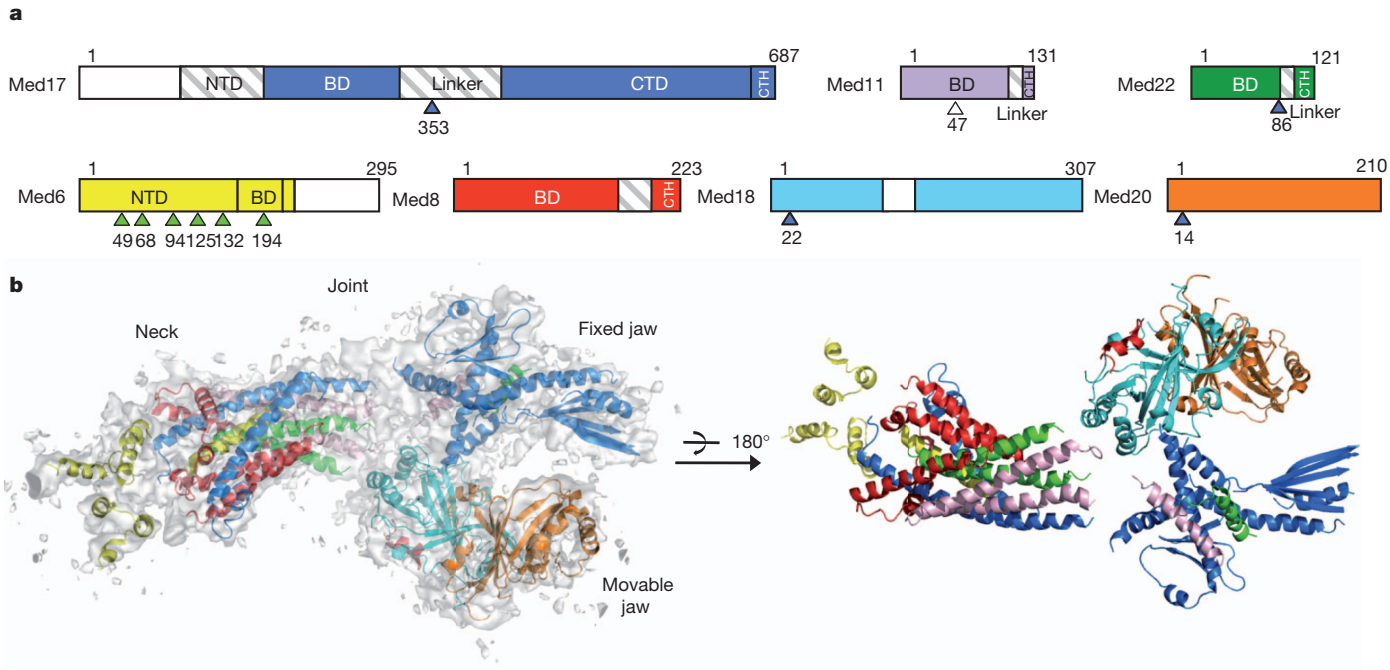

Figure 1 | Overall structure of the Mediator head module. **a,** Head module subunit domains. Med17 is shown in blue, Med11 in purple, Med22 in dark green, Med6 in yellow, Med8 in red, Med18 in cyan and Med20 in orange. The regions not modelled are hatched in grey and the regions not present in the crystal are shown in white. Positions of *med6*^{ts} mutations are marked by green arrows, *srb* suppressor mutations by blue arrows and Med11 residue 47 (Thr) by a white arrow. BD, bundle domain; CTD, C-terminal domain; NTD, N-terminal domain. **b,** A ribbon model of the Mediator head module superimposed on the experimental electron density map contoured at 1.5σ.

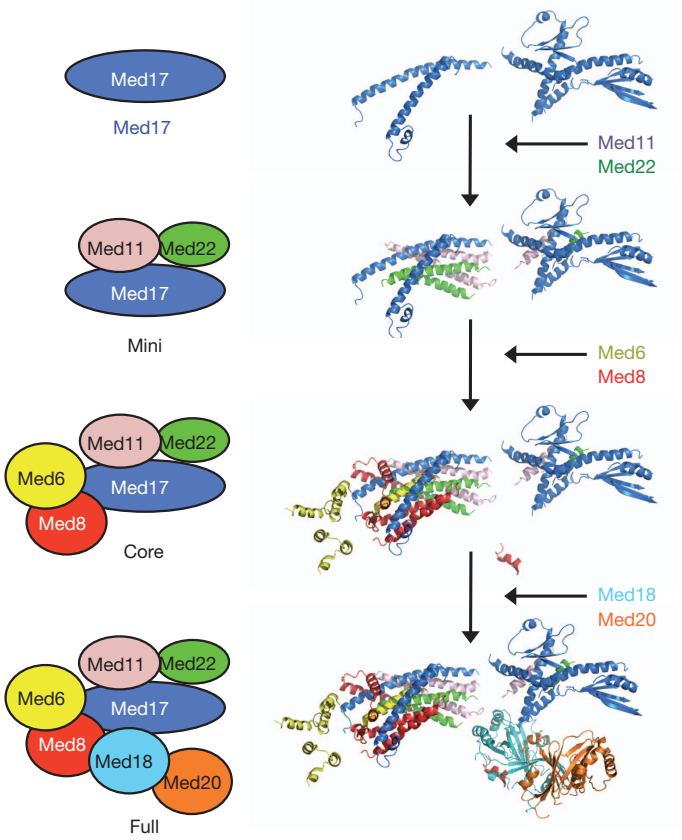

Figure 2 | Mechanism of Mediator head module complex assembly. Models of the mini-head (Med17, Med11 and Med22) and core-head (mini-head with Med6 and Med8) modules as derived from our crystal structure of the full head module (core-head with Med18 and Med20). Diagrams of head module components (left) and corresponding structures (right) are shown.

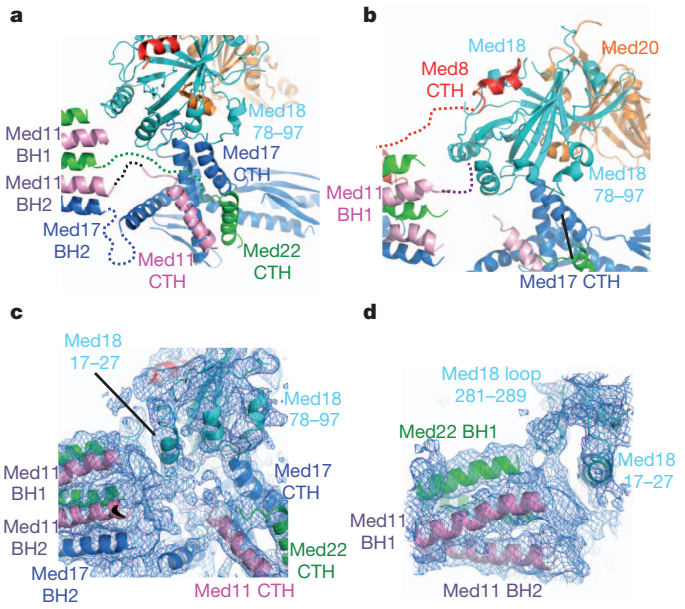

Figure 3 | Structures of fixed and movable jaw domains. **a**, Fixed jaw domain interactions. The linker regions of Med17 (residues 320–420), Med11 (94–110) and Med22 (87–100) are drawn as dotted lines. **b**, Movable jaw domain interactions. Linker regions of Med8 and Med11 are drawn as dotted lines. **c**, Electron density map at the central joint region, showing density corresponding to the linker regions of Med11, Med22 and Med17. **d**, Electron density map at the junction of the Med11, Med22 and Med18 subunits. The models of Med22 BH1, Med11 BH1 and BH2, and the Med18 loop region (residues 17–27 and 281–289) are superimposed.

and all Med17 CTD deletion mutants as well as internal deletion mutants result in lethality (Supplementary Fig. 13).

The movable jaw, so called because previous electron microscopy studies[16] demonstrated multiple orientations of this domain with respect to the rest of the head module, is formed by the Med18–Med20–Med8 CTH complex. As for the interaction with the Med8 CTH[18], our complete head module structure has revealed additional interactions with the fixed jaw and neck domains. First, the Med18 loop region formed by residues 78–97 interacts with the Med17 CTH of the fixed jaw domain (Fig. 3a, b). Second, the electron density corresponding to the N terminus of the Med11 subunit (residues 1–20) indicates an interaction with Med18 (residues 17–27 and 281–289; Fig. 3c, d). The assignment of Med11 residues 1–20 was complicated by the substitution of Ser 17 for Met 17 (Methods), and an unambiguous sequence marker is therefore lacking. However, our biochemical data (Supplementary Information, section 5) support our architectural model, in which a stable association between Med18–Med20 and the head module requires binding to Med8 and at least one additional interaction (with Med11 or Med17). The interactions with the CTH of Med17 and the NTD of Med11 are likely to be critical for the functional positioning and flexibility of the movable jaw[16] (Fig. 3b–d and Supplementary Fig. 6).

The neck domain has an unusual structure: a total of ten helices from five different subunits associate through the formation of a large helical bundle. The NTD of Med6 is located adjacent to the large helical bundle and consists of four α-helices (Figs 1b and 4a). The helical bundle of the neck domain can be divided into two parts, a short bundle composed of four short α-helices and a long bundle composed of six long α-helices. Three helices of the Med8 subunit (BH3, BH4 and BH5) seem to stabilize the assembly of both short and long bundles, and, thus, the entire neck domain structure. TATA-binding protein was reported to bind to the N-terminal 138 residues of Med8 (ref. 18), which corresponds to helices BH1 to BH5, all of which are located on the surface of the neck domain.

The organization of the helical bundle in the neck domain may produce a relatively rigid structure that could mechanically convey regulatory signals. Several observations suggest that Med6 may function as an interface between the Mediator head and middle modules, and transduce a mechanical signal from the tail or middle to the head and onto Pol II (Supplementary Information, section 6).

Mediator stimulates the phosphorylation of Ser 5 in the Pol II CTD by TFIIH (ref. 19), which promotes dissociation of Mediator from Pol II (refs 20, 21), an important step in the transition from initiation to elongation of transcription[22]. Our structural and biochemical data, along with relevant previous observations[12,17,23,24], suggest an interaction of the Pol II CTD, the Mediator head module and TFIIH. First, mutation of Thr 47 to Ala in Med11 affects the interaction of TFIIH with the head module *in vivo*, resulting in a reduction of Pol II CTD Ser 5 phosphorylation[17]. Thr 47 of Med11 is located near the centre of the two symmetrical, long helical bundles of the neck, which thus could constitute the docking surface for TFIIH (Fig. 4b and Supplementary Fig. 16c). Second, three of four suppressor mutations of Pol II CTD truncation—Med17 (Gly 353 to Cys), Med22 (Asn 86 to Lys) and Med18 (Thr 22 to Ile)—map to the central joint region[12,24] (Supplementary Fig. 16a), suggesting that there is a functional interaction between the CTD and this portion of the head, consistent with previous observations[23]. Third, the head module within the Mediator/Pol II structure (Fig. 4b and Supplementary Fig. 17) is located near the base of the CTD. Finally, our biochemical data show that the head module stimulates phosphorylation of the Pol II CTD by TFIIH (Supplementary Information, section 7, and Supplementary Fig. 18). Therefore, we suggest that the head module may function as a scaffold that juxtaposes TFIIH and the Pol II CTD, thereby facilitating CTD phosphorylation (Fig. 4b). Our Mediator head structure reveals intricate interaction networks, notably the striking multi-helical bundle in the neck domain, engaging five Mediator subunits in a single structure unit. Such interactions could not have been determined from structures of individual subunits alone, nor from analysing pairwise small domain–domain interactions, but only by study of the multi-protein complex in its entirety.

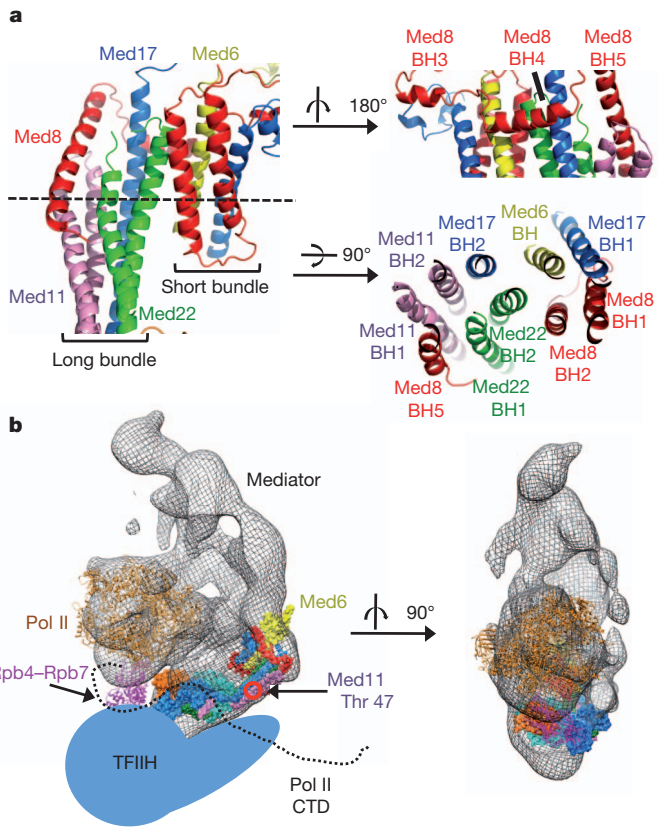

Figure 4 | Structure of the neck domain, and model of the Pol II–Mediator–TFIIH complex. **a**, The neck domain is depicted in front (left), back (top right) and top (bottom right) views. **b**, Model of the Pol II–Mediator–TFIIH complex. Pol II and the head structures were docked into the electron microscopy map of Mediator–Pol II, shown as a mesh[8]. The head module is coloured as in Fig. 1b. Core Pol II is in brown, Rpb4–Rpb7 is in purple, the Pol II CTD is drawn as a black dotted line and TFIIH is shown schematically (light blue). The location of Med11 residue 47 (Thr) is indicated.

METHODS SUMMARY

Structure determination. Modified head module was expressed with the MultiBac system[25] in insect cells and purified by nickel affinity chromatography. Crystals were obtained by the hanging-drop vapour diffusion method. The structure was determined by SeMet SAD after a sufficient number (98) of SeMet sites had been identified from a combination of initial phases obtained using Ta_6Br_{14} and iridium derivatives and partial-model SAD phases.

Biochemical and electron microscopy analysis. The Mediator head and its mutants were expressed in insect cells and purified by nickel affinity chromatography. The electron microscope images of the head module and the mutants were collected and class averages were calculated.

***In vitro* assays and yeast genetics.** The *in vitro* transcription assay to assess activity of the recombinant head module and its mutant form using *srb4*^{ts} mutant crude extract, the assay for phosphorylation of the CTD of Pol II by TFIIH, and the yeast phenotypic analysis were all done as described previously[15].

Full Methods and any associated references are available in the online version of the paper at www.nature.com/nature.

Received 24 January; accepted 28 April 2011.
Published online 3 July 2011.

1. Kornberg, R. D. Mediator and the mechanism of transcriptional activation. *Trends Biochem. Sci.* **30,** 235–239 (2005).

2. Conaway, R. C., Sato, S., Tomomori-Sato, C., Yao, T. & Conaway, J. W. The mammalian Mediator complex and its role in transcriptional regulation. *Trends Biochem. Sci.* **30,** 250–255 (2005).
3. 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).
4. Malik, S. & Roeder, R. G. The metazoan Mediator co-activator complex as an integrative hub for transcriptional regulation. *Nature Rev. Genet.* **11,** 761–772 (2010).
5. Bjorklund, S. & Gustafsson, C. M. The yeast Mediator complex and its regulation. *Trends Biochem. Sci.* **30,** 240–244 (2005).
6. Guglielmi, B. *et al.* A high resolution protein interaction map of the yeast Mediator complex. *Nucleic Acids Res.* **32,** 5379–5391 (2004).
7. Asturias, F. J., Jiang, Y. W., Myers, L. C., Gustafsson, C. M. & Kornberg, R. D. Conserved structures of mediator and RNA polymerase II holoenzyme. *Science* **283,** 985–987 (1999).
8. Davis, J. A., Takagi, Y., Kornberg, R. D. & Asturias, F. A. Structure of the yeast RNA polymerase II holoenzyme: Mediator conformation and polymerase interaction. *Mol. Cell* **10,** 409–415 (2002).
9. Cai, G., Imasaki, T., Takagi, Y. & Asturias, F. J. Mediator structural conservation and implications for the regulation mechanism. *Structure* **17,** 559–567 (2009).
10. Takagi, Y. *et al.* Head module control of mediator interactions. *Mol. Cell* **23,** 355–364 (2006).
11. 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,** 715–724 (1989).
12. 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).
13. Thompson, C. M. & Young, R. A. General requirement for RNA polymerase II holoenzymes *in vivo*. *Proc. Natl Acad. Sci. USA* **92,** 4587–4590 (1995).
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. Cai, G. *et al.* Mediator head module structure and functional interactions. *Nature Struct. Mol. Biol.* **17,** 273–279 (2010).
17. Esnault, C. *et al.* Mediator-dependent recruitment of TFIIH modules in preinitiation complex. *Mol. Cell* **31,** 337–346 (2008).
18. Larivière, L. *et al.* Structure and TBP binding of the Mediator head subcomplex Med8–Med18–Med20. *Nature Struct. Mol. Biol.* **13,** 895–901 (2006).
19. Kim, Y. J., Bjorklund, S., Li, Y., Sayre, M. H. & Kornberg, R. D. A multiprotein mediator of transcriptional activation and its interaction with the C-terminal repeat domain of RNA polymerase II. *Cell* **77,** 599–608 (1994).
20. Svejstrup, J. Q. *et al.* Evidence for a mediator cycle at the initiation of transcription. *Proc. Natl Acad. Sci. USA* **94,** 6075–6078 (1997).
21. Max, T., Sogaard, M. & Svejstrup, J. Q. Hyperphosphorylation of the C-terminal repeat domain of RNA polymerase II facilitates dissociation of its complex with mediator. *J. Biol. Chem.* **282,** 14113–14120 (2007).
22. Payne, J. M., Laybourn, P. J. & Dahmus, M. E. The transition of RNA polymerase II from initiation to elongation is associated with phosphorylation of the carboxyl-terminal domain of subunit IIa. *J. Biol. Chem.* **264,** 19621–19629 (1989).
23. Kang, J. S. *et al.* The structural and functional organization of the yeast mediator complex. *J. Biol. Chem.* **276,** 42003–42010 (2001).
24. 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).
25. Fitzgerald, D. J. *et al.* Protein complex expression by using multigene baculoviral vectors. *Nature Methods* **3,** 1021–1032 (2006).

Supplementary Information is linked to the online version of the paper at www.nature.com/nature.

Acknowledgements We thank L. Messerle for providing the Ta_6Br_{14} metal cluster, T. Hurley for reading the manuscript, M. Georgiadis for discussions, C. Kaplan for his advice on yeast genetics, T. Earnest for giving us beam time and L. Fabrizio for assisting with the N-terminal sequence analysis. We thank the CCP4 summer school, funded by the NCI (Y1-CO-1020), and NIGMS (Y1-GM-1104) for their assistance with the twinning data analysis. Y.T. thanks the instructors on 'The X-Ray Methods Course' at Cold Spring Harbor Laboratory. This work was supported by US National Science Foundation grant MCB 0843026 (Y.T.); the American Heart Association 0735395N (Y.T.); a Human Frontier Science Program long-term fellowship (T.I.); NIH grants R01 GM67167 (F.J.A.) and GM36659 (R.D.K.); NCI Cancer Center Support Grant P30 CA08748 (to the MSKCC Microchemistry and Proteomics Core Laboratory); and the European Commission Framework Program 7 projects INSTRUCT and P-CUBE (I.B.). X-ray data were collected at the GM/CA-CAT at the Advanced Photon Source, Argonne National Laboratory. GM/CA-CAT is funded by the NIH (Y1-CO-1020 and Y1-GM-1104) and the Advanced Photon Source is supported by the DOE (DE-AC02-06CH11357). Portions of this research were carried out at the Stanford Synchrotron Radiation Lightsource, supported by the DOE and the NIH, and at the Advanced Light Source, supported by the DOE (DE-AC02-05CH11231).

Author Contributions T.I., I.B. and Y.T. implemented the MultiBac system. T.I. was mainly responsible for protein complex preparation, crystallization, data collection, data analysis and model building in collaboration with Y.T. T.I., H.E.-B. and P.T. carried out mass spectroscopy analysis. G. Calero, G.L.K. and Y.T. carried out the initial crystallization and data collection, supervised by R.D.K.. Y.T., T.I. and F.C. designed and carried out expression of the mutant head modules and their biochemical characterization. Y.T. and K.Y. designed and carried out the yeast genetic experiment. Y.T. carried out the *in vitro* transcription assay and the CTD kinase assay; G. Cai, K.-L.T. and F.J.A. carried out the electron microscopy study on the head module and its mutants. T.I., F.J.A. and Y.T. discussed and interpreted all results. Y.T. supervised the X-ray, biochemical and yeast genetic work, and wrote the manuscript in collaboration with T.I., I.B., F.J.A. and R.D.K.

Author Information Coordinates and structure factors have been deposited in the Protein Data Bank under accession code 3RJ1. Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of this article at www.nature.com/nature. Correspondence and requests for materials should be addressed to Y.T. (ytakagi@iupui.edu).

METHODS

Construction of vectors. All the vectors used in this study are summarized in Supplementary Tables 3 and 4. For expression of the modified head module for crystallization, DNA sequences corresponding to residues 1–108 of Med17 were removed from vector pFL-10xHis-Med17 (ref. 16; pYT49) by the SLIC method[26], yielding pFL-10xHis-Med17 (109–687) (pYT171). DNA sequences corresponding to residues 109–140 and 71–156 of Med18, respectively, were removed from vector pSPL-Med18-Med20 (ref. 16; pYT75), yielding pSPL-Med20-Med18 (Δ109–140) (pYT115) and pSPL-Med20-Med18 (Δ71–156) (pYT114). To eliminate an alternative translation start site, the Met (residue 17) of Med11 was mutated to Ser (pYT311). Finally, the transfer vector for the modified head module was generated by fusing three vectors, pYT171, pYT114 and pUCDM-Med6-Med22-Med11-Med8 (pYT120) by Cre/*LoxP* recombination as previously described[25].

DNA sequences corresponding to residues 1–16 of Med11 were removed from the vector pYT111 by SLIC, yielding the vector pUCDM-Med22-Med11 (Δ1–16) (pYT147). Fusion of pYT171, pYT147 and pYT120 with either pYT114 or, alternatively, pYT115 generated the expression vectors for a series of double Med18–Med11 partial head module deletion mutants.

The constructs for Med17 mutagenesis were generated as follows. BamHI and HindIII fragments corresponding to the C-terminal deletion mutants of Med17 were generated by first introducing a stop codon and a HindIII site (TAAAAGCTT) into pBacPAK9-10His-SRB4 (MED17) vector[10] adjacent to the sequences corresponding to residues 108, 200, 300, 400, 500 and 600 of Med17, by using the QuickChange mutagenesis kit (Stratagene), followed by BamHI and HindIII digestion and gel purification. The respective purified fragments were cloned into the BamHI and HindIII sites of pFL vector[25], yielding vectors pYT165 to pYT170 (Supplementary Table 3). The N-terminal deletion, as well as the internal deletion mutant constructs of Med17, were generated by removing DNA sequences corresponding to residues 1–108, 1–201, 1–302, 1–400, 101–200, 201–300 and 301–400 from pFL-10xHis-Med17 by the SLIC method, yielding the respective vectors pYT183 to pYT186 and pYT289 to pYT291 (Supplementary Table 3). These vectors were fused with pUCDM-Med6-Med22-Med18-Med20-Med11-Med8 (pYT151), yielding vectors encoding for head modules comprising Med17 mutant forms. The vector pYT151 was created by two rounds of sequential cloning of PmeI and the AvrII fragments containing Med18–Med20 and Med22–Med11 into SpeI and NruI sites of pUCDM-Med6-Med8 (pYT110).

Introduction of the deletion mutations into the yeast shuttle vector pCT127, carrying the wild-type *MED17* gene, was also carried out by SLIC. The yeast shuttle vectors used in this study are listed in Supplementary Table 4.

Expression and purification of the head module and mutants. Expression and purification of the recombinant head module, the mutant forms and the subcomplex was carried out in insect cells using the MultiBac system[25]. Production of high-titer viruses in Sf9 cells, and expression and purification of recombinant head modules of Mediator and its mutant forms was carried out as described previously[16].

Preparation of SeMet labelled Head module will be described elsewhere (T.I. *et al.*, manuscript in preparation). Briefly, the insect cells were cultured in Met-free medium (Expression Systems) overnight before baculovirus infection. L-selenomethionine (20 mg l^{-1}; Sigma-Aldrich) was added at sequential 24-h intervals. Cells were collected 96 h after infection. SeMet-labelled complex was purified as described above.

Limited proteolysis and identification of the peptide fragments. A total of 135 μg of the recombinant head module was incubated at 37 °C with chymotrypsin (Sigma-Aldrich) at a final concentration of 0.01 mg ml^{-1} in a volume of 150 μl in buffer A (20 mM Tris-HCl (pH 8.0), 100 mM NaCl and 1 mM DTT). Aliquots (20 μl) were taken at 0, 5, 10, 30 and 60 min, and 15 μl of PMSF stock solution (100 mg ml^{-1}) was added to stop the reaction by inhibiting the protease. Aliquots were applied to 12.5% SDS–PAGE and transferred onto a Sequi-Blot PVDF membrane (Bio-Rad). Protein bands were stained by Coomassie blue (R-250). Protein bands resulting from proteolysis during the time course were identified, excised and subjected to Edman degradation using a Procise 494 instrument from Applied Biosystems as previously described[27]. Stepwise-liberated PTH-amino acids were identified using an 'on-line' HPLC system (Applied Biosystems) equipped with a PTH C18 (2.1 × 220 mm; 5-μm particle size) column (Applied Biosystems).

Crystallization and data collection. Crystals were obtained at 293 K by hanging-drop vapour diffusion against a reservoir solution of 0.1 M Tris-HCl (pH 7.6) containing 10–12.5% (w/v) PEG-6K and 0.4 M $(NH_4)_2SO_4$. Crystals were transferred into the reservoir solution containing 25% triethylene glycol (TEG). The crystals were flash-frozen for data collection at 100 K. SDS–PAGE analysis of the dissolved crystals confirmed the presence of all seven subunits. However, *in situ* proteolysis resulted in about 10% of the Med17 subunits being shortened at the N terminus by 76 residues and almost 100% of the Med6 subunits being shortened at the C terminus by 80 residues (Supplementary Fig. 1). Diffraction data were collected at beamline 23ID at the Advanced Photon Source (APS) at Argonne National Laboratory. All diffraction data were processed with HKL2000[28]. Twinning rates of the data sets were analysed using program PHENIX XTRIAGE[29].

Structure determination of the Mediator head module. Initial phases were determined by two approaches: Ta_6Br_{14} single isomorphous replacement with anomalous scattering (SIRAS) and iridium single anomalous dispersion (SAD). Ta_6Br_{14} derivative crystals were prepared by soaking the native head module crystals in reservoir solution containing 1 mM Ta_6Br_{14}. The initial phase was determined by SIRAS at a resolution of 7.5 Å. Density modification using the program PARROT extended the phase resolution to 4.3 Å using the SeMet data set. Iridium derivatives of the crystal of Mediator head module were prepared by soaking the crystals in crystallization reservoir solution containing 10 mM $K_3Ir(NO_3)_6$. The initial iridium phase was obtained by SAD using the programs SHELXD and PHASER[30,31]. The phase was extended followed by density modification by program PARROT[32] with the SeMet data set. However, the maps obtained at this stage were not yet interpretable.

To improve the maps, we used them together and applied the following methods: (i) location of SeMet sites in the crystal; (ii) non-crystallographic symmetry (NCS), averaging between three molecules in the NCS using the program DM[33]; (iii) partial model building into the clearly discernible rod-like electron density from α-helices, followed by rigid-body refinement using the programs COOT and REFMAC5[34]; and (iv) re-calculating phases by SeMet SAD phasing with PHASER, using the partial model and SeMet positions. Iterative rounds combining these procedures were performed until the model covered all interpretable secondary structure elements. Eventually, we could identify 98 SeMet sites. To minimize model bias, phases were re-calculated by SeMet SAD with PHASER, using only the positions of these 98 selenium sites, and these improved SAD phases guided the final model-building steps.

Model building and refinement. Assignment of polypeptide identities was carried out as follows. The published structure[18] of Med18–Med20–Med8 (CTH) (PDB ID, 2HZS) was manually docked into the electron density map, followed by rigid-body refinement by using COOT. Then the α-helices for the further polypeptides of the head module were manually built, and connected. Next, β-sheets were manually built into the unassigned structured regions in the electron density, which corresponded to the neck and fixed jaw domains. Subsequently, we began assigning the polypeptide identities at the neck domain. We tracked specific SeMet labelling patterns dictated by the presence of Met in the primary sequences of the polypeptides, and the presence of bulky regions corresponding to aromatic residue positions, as markers. We used secondary structure predictions for additional guidance. First, Med8 BH1, Med17 BH1 and Med22 BH2 were identified in α-helical bundle regions in the neck domain from their primary-sequence-specific, unique SeMet labelling pattern: these regions all contain more than two SeMet peaks and the spacing of SeMet peaks was consistent with the corresponding amino-acid sequences in the subunits. This assignment was consistent with the secondary structure predictions indicating α-helical structure. The remaining Med8 residues (60–170), as well as Med22 BH1, were assigned by tracing from the Med8 BH1 helix back to the Med8 C terminus, and by tracing from the Med22 BH2 helix back to the Med22 N terminus. This assignment was validated by the fact that their Met locations aligned with anomalous peaks on the experimental map. Next, we identified the NTD of Med6 based on its unique SeMet positions, and also identified Med6 BH based on a specific location of SeMet (Met 48), a bulky aromatic ring (Phe 52) (Supplementary Fig. 19a) and continuity from the NTD of Med6, consistent with secondary structure predictions. The C-terminal 80 residues of Med6 were proteolyzed in the crystals (Supplementary Fig. 1). Consequently, no density was found corresponding to the C terminus of Med6. We traced Med17 BH1, and identified the longest helix in the neck domain as the BH2 helix of Med17 on the basis of a single SeMet (Met 313) and the aromatic side chain of Tyr 269 (Supplementary Fig. 19b); this assignment was also consistent with secondary structure predictions. Finally, the two remaining continuous α-helices in the neck domain were identified as Med11 BH1 and Med11 BH2 because of one unique SeMet position of Med11. This assignment matches perfectly to the secondary structure prediction as well.

Next we focused on the fixed jaw domain. By subtracting the polypeptides already assigned to the neck and the movable jaw (see above), the fixed jaw should only contain the C-terminal regions of subunits Med11, Med17 and Med22. First, on the basis of continuity, SeMet position (Met 422), aromatic ring position (Tyr 423) (Supplementary Fig. 19c), and secondary structure prediction, we identified helix 420–455, β-sheet 456–480 and helices 496–523 and 540–570 of Med17. The remainder of the electron density in this region was continuous, and thus enabled us to trace Med17 completely to its C terminus. We identified the Med17 CTH and β-sheet with α-helix 600–608 on the basis of the SeMet positions and α-helix length from secondary structure prediction. Finally, we assigned two remaining helices: Med11 CTH was identified from the presence of one SeMet peak, and we assigned the last helix to Med22 CTH, which entirely lacks SeMet.

Initially, all models were refined using the program CNS DEN[35], refinement with strong NCS restraints between the three independent complexes in the asymmetric unit, and twinning refinement. Then the model was refined using PHENIX with NCS restraints and a single refined group isotropic temperature factor for each subunit, Ramachandran restraint, TLS refinement and twinning refinement. The geometry of the final model is good, with 91.3%, 8.0%, 0.7% of the amino-acid residues in the most favoured, allowed, and disallowed regions of the Ramachandran plot, respectively. All structural illustrations and electron density maps were prepared with PYMOL (http://www.pymol.org/) and COOT. PSIPRED was used for secondary structure prediction[36].

Docking of the X-ray structure into the electron microscopy map. The model of 12-subunit Pol II was docked into the Mediator–Pol II holoenzyme structure[8] followed by docking of the X-ray model of the head module into the density corresponding to the Mediator head module, using the program CHIMERA[37].

Electron microscopy sample preparation, data collection and image analysis. We diluted purified head module deletion mutants in buffer containing 25 mM KCl, 25 mM Tris-HCl (pH 7.8) and 5 mM DTT. For preparation of all electron microscopy samples, about 3 µl of protein solution was applied to a carbon-coated Maxtaform, 300-mesh Cu/Rh EM specimen grid (Ted Pella) freshly glow-discharged in the presence of amyl amine. The particles were then preserved by staining with a 2.0% (w/w) uranyl acetate solution using the sandwich carbon layer technique[38,39]. The images were recorded under low-dose conditions using a Tecnai Spirit (Philips/FEI) microscope equipped with a LaB6 filament and operating at an accelerating voltage of 120 kV. Images were recorded on a Tietz (TVIPS) CCD camera at ×42,000 magnification and approximately 1-µm underfocus, resulting in a final pixel size corresponding to 5.06 Å.

The images were initially analysed using the ml_align2d program, a multi-reference, two-dimensional alignment routine with a maximum-likelihood target function[40] implemented in the XMIPP package[41]. Averages derived from the ml_align2d program were used to run iterative alternating rounds of supervised multi-reference alignment/classification and reference-free alignment as described previously[42] to improve the homogeneity of the image classes.

***In vitro* transcription and the CTD phosphorylation assays.** The *in vitro* transcription assay to assess activity of the recombinant head module and its mutant form using *srb4*^{ts} mutant crude extract was performed as described previously[10]. Quantification of transcripts on an absolute scale was performed using a FLA-5100 FUJIFILM fluorescent image analyser and the MultiGauge software package after addition of 1 nCi of -^{32}P UTP to the gel 5 min before the end of the run. The CTD phosphorylation assay was performed as previously described[15].

Yeast phenotypic analysis. The shuttle vectors carrying the MED17 mutations are described in Supplementary Table 4. The shuttle vectors were introduced into yeast strain Z572 by plasmid shuffling, and grown on SC medium containing 5-FOA at 30 °C as previously described[15].

26. Li, M. Z. & Elledge, S. J. Harnessing homologous recombination in vitro to generate recombinant DNA via SLIC. *Nature Methods* **4,** 251–256 (2007).
27. Tempst, P., Geromanos, S., Elicone, C. & Erdjument-Bromage, H. Improvements in microsequencer performance for low picomole sequence analysis. *Methods* **6,** 248–261 (1994).
28. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276,** 307–326 (1997).
29. Adams, P. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D* **66,** 213–221 (2010).
30. Sheldrick, G. A short history of SHELX. *Acta Crystallogr. A* **64,** 112–122 (2008).
31. McCoy, A. *et al.* Phaser crystallographic software. *J. Appl. Crystallogr.* **40,** 658–674 (2007).
32. Cowtan, K. Recent developments in classical density modification. *Acta Crystallogr. D* **66,** 470–478 (2010).
33. Cowtan, K. Modified phased translation functions and their application to molecular-fragment location. *Acta Crystallogr. D* **54,** 750–756 (1998).
34. Emsley, P., Lohkamp, B., Scott, W. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. D* **66,** 486–501 (2010).
35. Schröder, G. F., Levitt, M. & Brunger, A. T. Super-resolution biomolecular crystallography with low-resolution data. *Nature* **464,** 1218–1222 (2010).
36. Bryson, K. *et al.* Protein structure prediction servers at University College London. *Nucleic Acids Res.* **33,** W36–W38 (2005).
37. Pettersen, E. *et al.* UCSF Chimera–a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25,** 1605–1612 (2004).
38. Stoffler, G. & Stoffler-Meilicke, M. *The Ultrastructure of Macromolecular Complexes Studied with Antibodies* 409–455 (De Gruyter, 1983).
39. Tischendorf, G. W., Zeichhardt, H. & Stoffler, G. Determination of the location of proteins L14, L17, L18, L19, L22, L23 on the surface of the 50S ribosomal subunit of *Escherichia coli* by immune electron microscopy. *Mol. Gen. Genet.* **134,** 187–208 (1974).
40. Scheres, S. H. *et al.* Maximum-likelihood multi-reference refinement for electron microscopy images. *J. Mol. Biol.* **348,** 139–149 (2005).
41. Sorzano, C. O. *et al.* XMIPP: a new generation of an open-source image processing package for electron microscopy. *J. Struct. Biol.* **148,** 194–204 (2004).
42. Brignole, E. J., Smith, S. & Asturias, F. J. Conformational flexibility of metazoan fatty acid synthase enables catalysis. *Nature Struct. Mol. Biol.* **16,** 190–197 (2009).