

Taf2 mediates DNA binding of Taf14

Brianna J. Klein¹, Jordan T. Feigerle^{2,3}, Jibo Zhang⁴, Christopher C. Ebmeier⁵, Lixin Fan⁶, Rohit K. Singh¹, Wesley W. Wang⁷, Lauren R. Schmitt⁸, Thomas Lee⁶, Kirk C. Hansen⁸, Wenshe R. Liu⁷, Yun-Xing Wang⁹, Brian D. Strahl⁴, P. Anthony Weil³ & Tatiana G. Kutateladze¹✉

The assembly and function of the yeast general transcription factor TFIID complex requires specific contacts between its Taf14 and Taf2 subunits, however, the mechanism underlying these contacts remains unclear. Here, we determined the molecular and structural basis by which the YEATS and ET domains of Taf14 bind to the C-terminal tail of Taf2 and identified a unique DNA-binding activity of the linker region connecting the two domains. We show that in the absence of ligands the linker region of Taf14 is occluded by the surrounding domains, and therefore the DNA binding function of Taf14 is autoinhibited. Binding of Taf2 promotes a conformational rearrangement in Taf14, resulting in a release of the linker for the engagement with DNA and the nucleosome. Genetic *in vivo* data indicate that the association of Taf14 with both Taf2 and DNA is essential for transcriptional regulation. Our findings provide a basis for deciphering the role of individual TFIID subunits in mediating gene transcription.

¹Department of Pharmacology, University of Colorado School of Medicine, Aurora, CO 80045, USA. ²Department of Structural Biology, Stanford University, Stanford, CA 94305, USA. ³Department of Molecular Physiology and Biophysics, Vanderbilt University School of Medicine, Nashville, TN 37232, USA. ⁴Department of Biochemistry & Biophysics, The University of North Carolina School of Medicine, Chapel Hill, NC 27599, USA. ⁵Department of Biochemistry, University of Colorado, Boulder, CO 80303, USA. ⁶Basic Science Program, Frederick National Laboratory for Cancer Research, SAXS Core Facility of the National Cancer Institute, Frederick, MD 21702, USA. ⁷Department of Chemistry, Texas A&M University, College Station, TX 77843, USA. ⁸Department of Biochemistry and Molecular Genetics, University of Colorado School of Medicine, Aurora, CO 80045, USA. ⁹Protein-Nucleic Acid Interaction Section, Structural Biophysics Laboratory, Center for Cancer Research, National Cancer Institute, National Institutes of Health, Frederick, MD 27102, USA. ✉email: tatiana.kutateladze@cuanschutz.edu

The initiation of gene transcription by RNA polymerase II (RNA Pol II) requires formation of the preinitiation complex (PIC) around the gene promoter region. In addition to RNA Pol II and a coactivator, also known as Mediator, the PIC contains six general transcription factor (TF) complexes, including TFIID and TFIIF. TFIIF is involved in the recruitment of RNA Pol II and the association with Mediator, whereas TFIID is responsible for binding to the promoter. Both yeast TFIID and TFIIF contain multiple subunits and share one subunit – Taf14 (transcription initiation factor subunit 14)^{1–5}. Additionally, Taf14 has been implicated in chromatin-remodeling and histone-modifying processes, being a component of the INO80, SWI/SNF and RSC complexes^{3,6,7} and the histone acetyltransferase complex NuA3⁸, and therefore plays numerous vital roles in chromatin biology. Taf14 is a 244-residue protein, which consists of the N-terminal YEATS (Yaf9, ENL, AF9, Taf14, Sas5) domain and the extra-terminal (ET) domain (Fig. 1a). The YEATS domain of Taf14 recognizes acylated lysine 9 of histone H3 (H3K9acyl)^{9–12}, and the ET domain is capable of associating with the hxxK motif found in several co-activators, such as Sth1 and Taf2¹³. Genetic studies have shown that the interaction between Taf14 and another subunit of TFIID complex, Taf2, could be essential in the TFIID assembly and function¹⁴, however the mechanism underlying this interaction remains unclear.

In this study, we describe the molecular mechanism by which the YEATS and ET domains of Taf14 bind to the C-terminal tail of Taf2 and characterize a DNA-binding activity of the linker connecting the two domains. We show that Taf14 undergoes a conformational rearrangement upon binding to Taf2 to release the linker for binding to DNA – a function that is important for the ability of Taf14 to regulate gene transcription.

Results and discussion

Structural insight into the Taf14-Taf2 complex formation. To define the molecular basis for the interaction between Taf14 and Taf2, we generated a set of truncated constructs of Taf14 and examined their association with the Taf2 peptide corresponding to the C-terminal hxxK motif-containing tail of Taf2 (Taf2_{CT}) by NMR and microscale thermophoresis (MST). Titration of Taf2_{CT} to the ¹⁵N-labeled ET domain of Taf14 (Taf14_{ET}) led to large chemical shift perturbations (CSPs) in the ¹H¹⁵N heteronuclear single quantum coherence (HSQC) NMR spectra of Taf14_{ET}, indicating formation of the complex (Fig. 1b). Many amide crosspeaks of Taf14_{ET} in the apo-state disappeared upon addition of the Taf2_{CT} peptide, and another set of resonances corresponding to the Taf2_{CT}-bound state of Taf14_{ET} appeared. The slow exchange regime on the NMR timescale suggested tight binding (Fig. 1c), which was confirmed by measuring binding affinity of Taf14_{ET} for Taf2_{CT} using MST (the dissociation constant (K_d) = 1.6 μ M) (Fig. 1c).

To gain insight into the binding mechanism, we crystallized the Taf14_{ET}-Taf2_{CT} complex and determined its structure to 1.7 Å by X-ray crystallography (Fig. 1d–f and Supplementary Table 1). Taf14_{ET} adopts a three-helix bundle fold with a two-stranded anti-parallel β -sheet connecting helices α 2 and α 3. The Taf2_{CT} peptide is in an extended conformation and occupies an elongated, primarily hydrophobic groove of Taf14_{ET} (Fig. 1d). Taf2_{CT} pairs with the second β -strand of Taf14_{ET}, forming the third anti-parallel β -strand in the β -sheet, and makes extensive intermolecular contacts with the domain (Fig. 1e). Characteristic β -sheet interactions are observed between the backbone amides of V1397, I1399 and T1401 of Taf2_{CT} and G218, F220 and I222 of Taf14_{ET}. The side chains of V1397 and I1399 of Taf2_{CT} are oriented toward and essentially buried in the hydrophobic groove

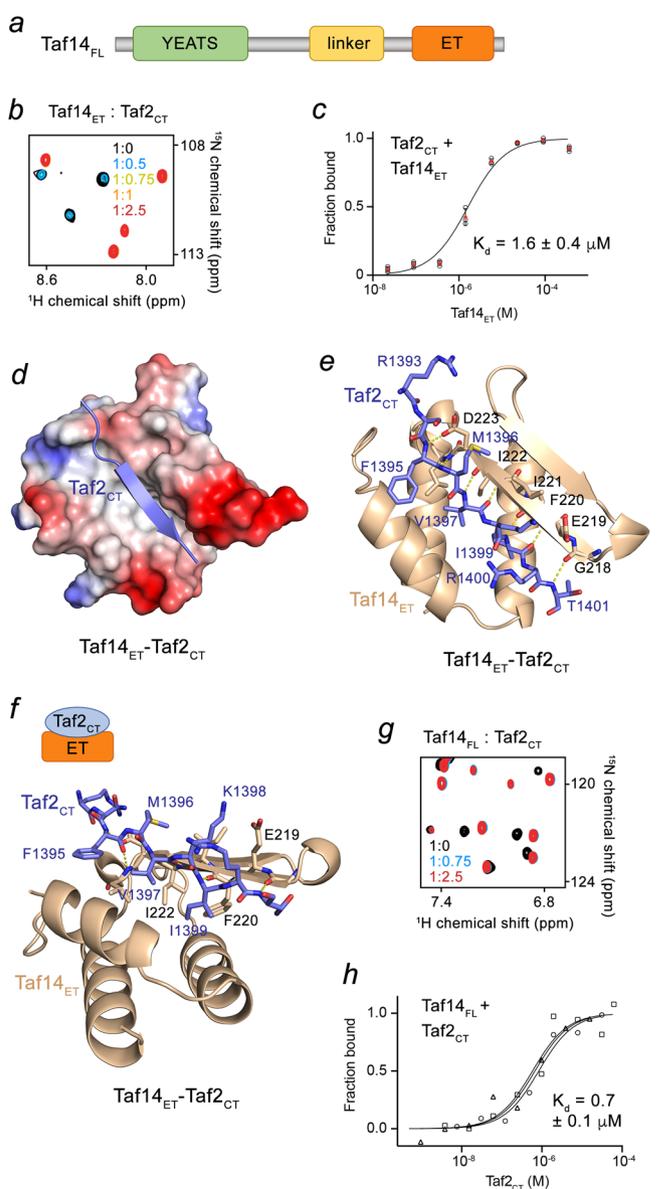


Fig. 1 Structural mechanism for the interaction of Taf14_{ET} with Taf2_{CT}.

a Architecture of Taf14_{FL}. **b** Superimposed ¹H¹⁵N HSQC spectra of Taf14_{ET} recorded in the presence of increasing amounts of the Taf2_{CT} peptide. The spectra are color coded according to the protein:peptide molar ratio. **c** MST binding curve for the FAM-Taf2_{CT} interaction with Taf14_{ET}. The K_d value represents average (shown as red diamonds) of three independent measurements (open circles), and error bars represent the standard deviation from average. $n = 3$ **d** Electrostatic surface potential of the crystal structure of the Taf14_{ET}:Taf2_{CT} complex, with blue and red colors representing positive and negative charges, respectively. The Taf2_{CT} peptide is shown as a purple ribbon. **e** Ribbon diagram of the Taf14_{ET}:Taf2_{CT} complex structure in the same orientation as in **(d)**. Taf14_{ET} domain is colored wheat and Taf2_{CT} peptide is purple. Residues involved in the interaction between Taf2_{CT} and Taf14_{ET} are labeled. Dashed lines represent hydrogen bonds. **f** Rotated view of the ribbon structure shown in **(e)**. Schematic of the Taf14_{ET}-Taf2_{CT} complex is shown above the structure. **g** Superimposed ¹H¹⁵N HSQC spectra of Taf14_{FL} recorded in the presence of increasing amounts of the Taf2_{CT} peptide. The spectra are color coded according to the protein:peptide molar ratio. **h** MST binding curves for the FAM-Taf2_{CT} interaction with Taf14_{FL}. The K_d value represents average of three independent measurements, and error represents the standard deviation. $n = 3$.

of Taf14_{ET}. The complex is further stabilized by a set of additional hydrogen bonds formed between the backbone amide group of F1395 of Taf2_{CT} and the side chain carboxyl group and the backbone amino group of D223 of Taf14_{ET}. The side chain amino group of K1398 of Taf2_{CT} donates a transient hydrogen bond to the side chain carboxyl moiety of E219 of Taf14_{ET}, whereas the guanidino moiety of R1400 of Taf2_{CT} is hydrogen bonded to the backbone carbonyl group of K1398.

To explore whether the interaction with Taf2_{CT} is conserved in full-length Taf14 (Taf14_{FL}), we expressed Taf14_{FL} as an ¹⁵N-labeled protein and tested it in NMR titration experiments. The pattern of CSPs induced in Taf14_{FL} by Taf2_{CT} and fast saturation suggested a stronger interaction compared to the binding of Taf14_{ET} (Fig. 1g). In agreement, the *K_d* value for the interaction of Taf14_{FL} with Taf2_{CT} was found to be 0.7 μM (Fig. 1h), demonstrating that binding of the full-length protein is ~two-fold tighter than binding of the isolated Taf14_{ET}.

Both domains of Taf14, Taf14_{ET} and Taf14_{YEATS}, bind to Taf2.

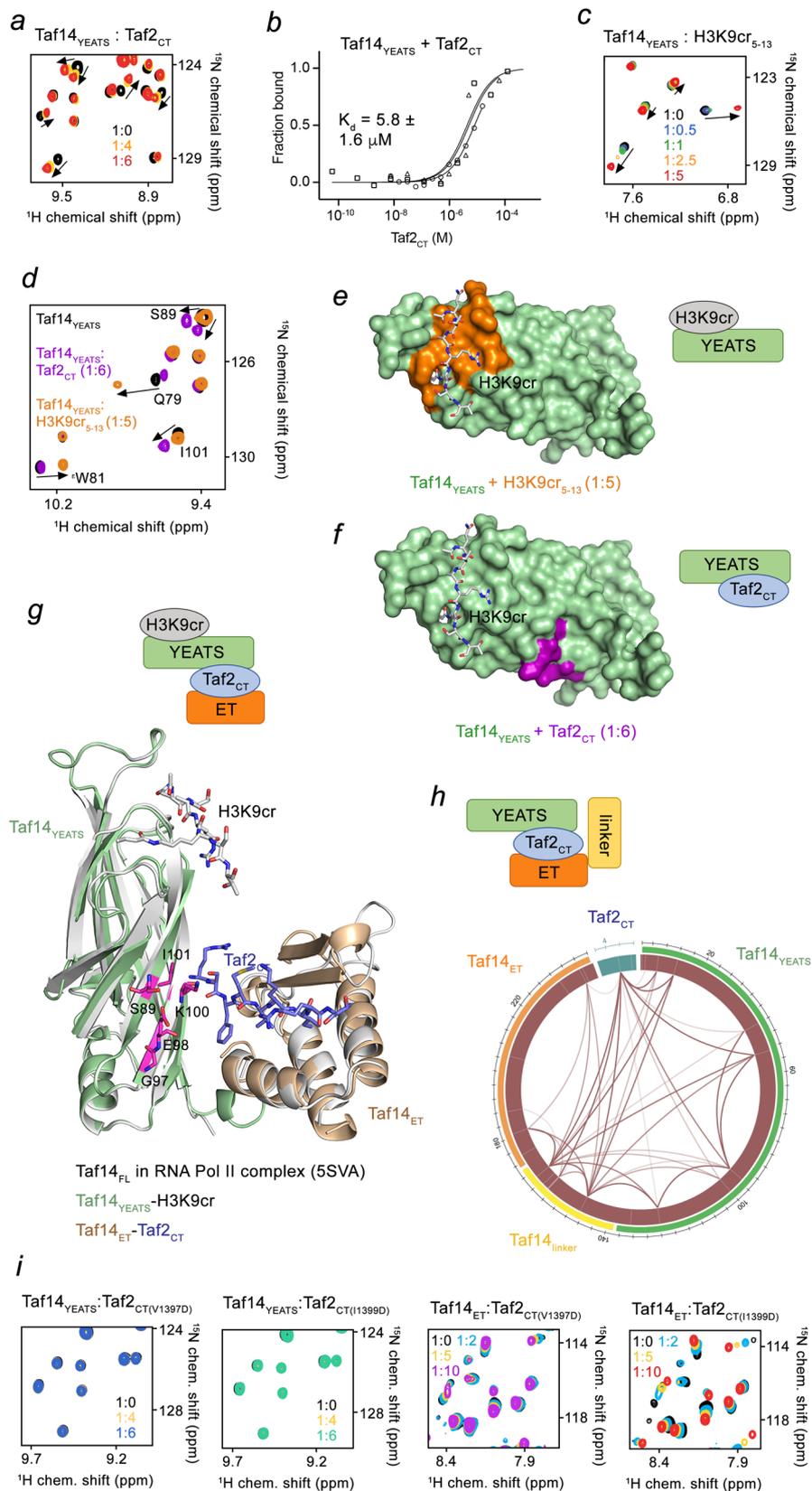
In effort to explain the increase in binding affinity of Taf14_{FL} toward Taf2_{CT}, we examined if other parts of Taf14 can contribute to the interaction with Taf2. Surprisingly, we found that the YEATS domain of Taf14 (Taf14_{YEATS}) is capable of binding to Taf2_{CT} as titration of the Taf2_{CT} peptide into the ¹⁵N-labeled Taf14_{YEATS} caused substantial CSPs in NMR spectra of Taf14_{YEATS} (Fig. 2a). MST measurements revealed that Taf14_{YEATS} interacts with Taf2_{CT} with a *K_d* of 6 μM (Fig. 2b). Taf14_{YEATS} is a well-established reader of H3K9acyl, particularly the crotonyl modification, H3K9cr^{9,11}. Titration of the H3K9cr peptide into the Taf14_{YEATS} NMR sample resulted in CSPs that were distinctly different from CSPs caused by titration of Taf2_{CT}, implying that Taf14_{YEATS} has separate binding sites for the two ligands, H3K9cr and Taf2_{CT} (Fig. 2c, d and Supplementary Fig. 1). Indeed, the structure of the Taf14_{YEATS}-H3K9cr complex and the mapping of the most perturbed residues in NMR titration experiments showed that the H3K9cr-binding site of Taf14_{YEATS} is located at one of the open ends of the Taf14_{YEATS} β-sandwich (Fig. 2e, orange), whereas residues located at the opposite side of the β-sandwich, in particular S89, G97, E98, K100 and I101, were perturbed upon binding of Taf2_{CT} (Fig. 2f, magenta).

Evaluation of the TFIIF complex, and particularly Taf14_{FL}, in the previously reported cryo-EM structure of the PIC⁴ showed that Taf14_{YEATS} and Taf14_{ET} are in close proximity (Fig. 2g). An overlay of the cryo-EM structure of Taf14_{FL} (grey) with the structure of the Taf14_{ET}(wheat)-Taf2_{CT}(blue) complex and the Taf14_{YEATS}(green)-H3K9cr complex (with the Taf2_{CT}-binding site colored magenta) suggests that Taf2_{CT} could be concurrently engaged with both domains, Taf14_{ET} and Taf14_{YEATS} (Fig. 2g and Supplementary Fig. 2). The possibility of such an engagement was corroborated by mass spectrometry analysis of DSSO cross-linked Taf14_{FL} in the presence of Taf2_{CT} (Fig. 2h, Supplementary Fig. 3 and Supplementary Data 1). In the Taf14_{ET}-Taf2_{CT} complex, the only lysine residue of Taf2_{CT}, K1398, is in the middle of the β-strand and thus its position is fixed. The side chain of K1398 is perpendicular to the Taf14_{ET} surface, pointing straight away from the domain (Supplementary Fig. 3), and cannot be easily crosslinked backward to Taf14_{ET}, as only a single weak contact was detected by MS (Fig. 2h). However, K1398 is readily crosslinked with several lysine residues of Taf14_{YEATS}, located in the same part of Taf14_{YEATS} that was perturbed in NMR experiments (Fig. 2h and Supplementary Fig. 3). The hydrophobic residues V1397 and I1399 of Taf2_{CT} were found to be required for this engagement, as V1397D and I1399D mutations in Taf2_{CT} substantially reduced binding of Taf14_{ET} and abrogated binding of Taf14_{YEATS} (Fig. 2i).

Taf14_{ET} and Taf14_{YEATS} regulate DNA binding activity of the linker (Taf14_{Linker}). Because TFIID plays an essential role in binding to DNA and bridging the PIC to gene promoters, we tested whether Taf14 has a DNA binding capability by NMR and electrophoretic mobility shift assays (EMSA). The direct interaction of Taf14_{FL} with 601 Widom DNA was initially detected by CSPs in ¹H¹⁵N HSQC spectra of Taf14_{FL} (Fig. 3a) and then confirmed in EMSA experiments (Fig. 3b). In EMSA, 601 DNA was incubated with increasing amounts of Taf14_{FL}, and the reaction mixtures were resolved on a native polyacrylamide gel. A shift of the free 601 DNA indicated the formation of Taf14_{FL}:DNA complex (Fig. 3b). To identify the DNA binding region, we generated a set of truncated and mutated Taf14 constructs and tested them in EMSA (Fig. 3c–e and Supplementary Fig. 4a–d). We found that Taf14_{YEATS} does not bind to 601 DNA, supporting previous findings¹⁵, and Taf14_{ET} shows a very weak DNA binding activity (Supplementary Fig. 4a, b). The W81A mutation in Taf14_{FL} that impairs binding of Taf14_{YEATS} to H3K9cr had no effect on the DNA binding (Supplementary Fig. 4c), whereas separate K149E/R150E/R151E (3K/Rmut) and K161E/K163E/R164E/K166E (4K/Rmut) mutations in the linker (Taf14_{Linker}) decreased binding of Taf14_{FL} to DNA (Fig. 3c and Supplementary Fig. 4d). We concluded that these two positively charged motifs in Taf14_{Linker} are responsible for DNA binding of Taf14. Interestingly, when either ET domain or the YEATS domain were removed (Taf14_{ΔET} or Taf14_{ΔYEATS}) but Taf14_{Linker} remained intact, the binding of these truncated constructs to DNA was over 6-fold tighter than the binding of Taf14_{FL} as seen in EMSA and NMR experiments (Fig. 3d–f). To explore the DNA sequence selectivity of Taf14, we tested binding of Taf14 to 36 bp AT rich dsDNA and 36 bp GC rich dsDNA by EMSA (Supplementary Fig. 5). As shown in Supplementary Figures 5b, c, Taf14 displayed a weak preference for the AT rich DNA sequence. Overall, these data indicate that Taf14_{Linker} binds to DNA via the two positively charged motifs, however the DNA binding activity is reduced when both Taf14_{YEATS} and Taf14_{ET} are present – likely due to the occlusion of the linker region in Taf14_{FL}.

Interaction of Taf14_{FL} with DNA leads to a conformational rearrangement.

How can Taf14_{Linker} be occluded in Taf14_{FL}? We investigated the ternary structure of Taf14_{FL} by NMR, mass spectrometry (MS), MST and SAXS. We found that Taf14_{ET} directly binds to Taf14_{Linker} because CSPs were induced in NMR spectra of ¹⁵N-labeled Taf14_{ET} by the peptide corresponding to the linker region (Fig. 3g). The binding affinity of Taf14_{ET} to the Taf14_{Linker} peptide was found to be 640 μM, as measured by MST (Fig. 3h). We note that due to an increase in apparent concentration of the ligand this interaction must be stronger when Taf14_{Linker} and Taf14_{ET} are linked in Taf14_{FL}. MS analysis revealed that both Taf14_{YEATS} and Taf14_{ET} are cross-linked to Taf14_{Linker} and therefore are in close proximity (Fig. 3i, Supplementary Fig. 6a and Supplementary Data 1), however the presence of a short 37 bp DNA in the MS sample substantially reduced the number of cross-linked peptides identified between Taf14_{YEATS} and the rest of Taf14 (Fig. 3j and Supplementary Fig. 6b). The SAXS data were in agreement and further suggested that Taf14_{FL} in the apo state is flexible, as indicated by the Kratky plot, and may exist in several conformations in solution, with a ‘close’ conformation being in dynamic equilibrium with a more ‘open’ conformation (Fig. 3k and Supplementary Figs. 7 and 8). When a 37 bp DNA was added to the SAXS sample, the complex became rigid (Fig. 3k), and equilibrium shifted toward the predominantly open, DNA-bound conformation (Fig. 3l). Lastly, the incomplete overlap between ¹H¹⁵N HSQC spectra of Taf14_{FL} and the isolated Taf14_{YEATS} and Taf14_{ET} suggested



that the domains are not fully independent and interactions or conformational changes occur within $Taf14_{FL}$ (Supplementary Fig. 9). We note that the conformational rearrangement and the ternary structure of $Taf14_{FL}$ did not reduce binding to H3K9cr, indicating that the H3K9cr binding site in $Taf14$ remains accessible (Fig. 3m).

Taf2_{CT} promotes binding of $Taf14_{FL}$ to the nucleosome. Unexpectedly, we found that neither wild type $Taf14_{FL}$ nor the W81A mutant of $Taf14_{FL}$ or $Taf14_{ET}$ appreciably interact with the H3K9cr-containing nucleosome (H3K9cr-NCP) in EMSA experiments (Fig. 4a and Supplementary Fig. 10a–c). In contrast to its binding to the H3K9cr peptide (Fig. 2c)¹¹, the individual

Fig. 2 Taf14_{ET} and Taf14_{YEATS} both interact with Taf2_{CT}. **a** Superimposed ¹H¹⁵N HSQC spectra of Taf14_{YEATS} collected upon titration with Taf2_{CT}. The spectra are color coded according to the protein:peptide molar ratio. **b** MST binding curves for the Taf2_{CT} interaction with Taf14_{YEATS}. The K_d value represents average of three independent measurements, and error represents the standard deviation. *n* = 3 **c** Superimposed ¹H¹⁵N HSQC spectra of Taf14_{YEATS} collected upon titration with H3K9_{CR5-13} (residues 5–13 of H3) peptide. The spectra are color coded according to the protein:peptide molar ratio. **d** Overlay of ¹H¹⁵N HSQC spectra of Taf14_{YEATS} in the apo state (black) and bound to Taf2_{CT} (purple) or H3K9_{CR5-13} (orange). **e, f** Identification of the H3K9_{CR}- and Taf2_{CT}-binding sites of Taf14_{YEATS}. Residues that exhibit substantial CSPs in Taf14_{YEATS} upon binding of H3K9_{CR} (**e**; orange) or Taf2_{CT} (**f**; magenta) are mapped on the surface of Taf14_{YEATS} (green) with the H3K9_{CR} peptide shown in grey sticks. **g** Overlay of the crystal structures of the Taf14_{YEATS}:H3K9_{CR} complex (light green:grey), Taf14_{ET}:Taf2_{CT} complex (wheat:purple), and the cryo-EM structure of Taf14_{FL} in the RNA Pol II TFIIF complex (5SVA, grey)⁴. The residues of the Taf2_{CT}-binding site of Taf14_{YEATS} are labeled and colored magenta. **h** Mass spectrometry analysis wheel diagram showing positions of DSSO-generated cross-links identified in the Taf14_{FL}:Taf2_{CT} complex. Line thickness correlates with the number of cross-link identifications. Taf14 domains and Taf2_{CT} within the circle are colored as in the diagram above and in Fig. 1a. **i** Superimposed ¹H¹⁵N HSQC spectra of Taf14_{YEATS} (two left panels) or Taf14_{ET} (two right panels) collected upon titration with the indicated mutant Taf2_{CT} peptides. Spectra are color coded according to the protein:peptide molar ratio. Schematics of the complexes are shown in (**e, f, g** and **h**).

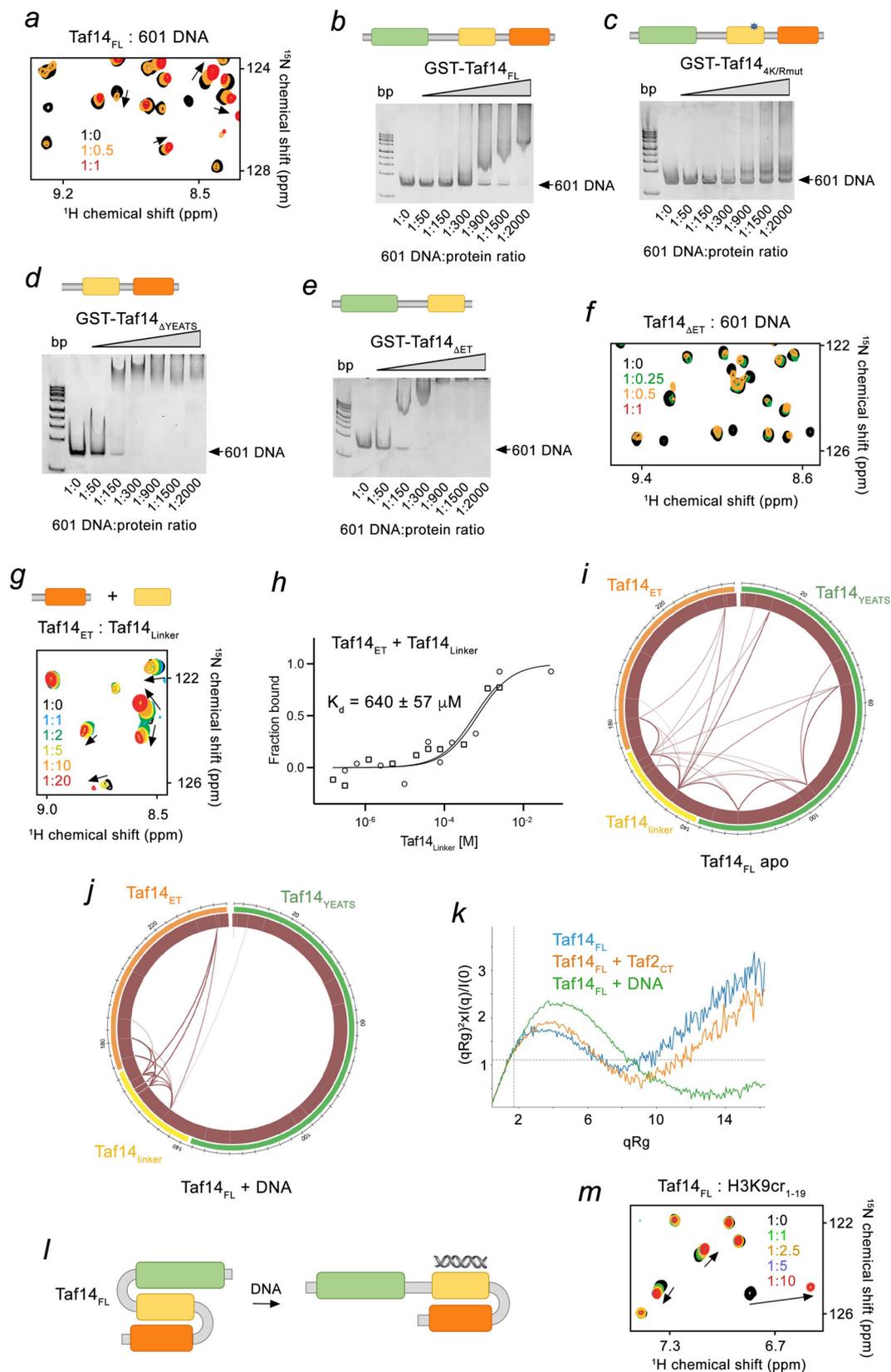
Taf14_{YEATS} was also incapable of binding to H3K9_{CR}-NCP, reflecting the previous findings that histone tails within the nucleosome are not as freely available for the interactions with readers as peptides (Supplementary Fig. 10b)^{16,17}. The association with the nucleosome was detected after the ET domain was removed in Taf14_{ΔET} (Fig. 4b) or the YEATS domain was removed in Taf14_{ΔYEATS} (Supplementary Fig. 10d), indicating that Taf14_{ET} and Taf14_{YEATS} impede the ability of Taf14_{FL} to be engaged with the nucleosome and therefore chromatin. The notable increase in binding to the nucleosome was also observed when Taf2_{CT} was added to the Taf14 samples (Fig. 4c and Supplementary Fig. 10e). The binding of Taf2_{CT} to Taf14_{FL}, Taf14_{ΔYEATS} and Taf14_{W81A}, the Taf14 constructs containing the linker region, enhanced the formation of the complexes with the nucleosome. Likewise, binding of Taf14_{FL} to double stranded 37 bp DNA was augmented in the presence of Taf2_{CT} (compare lanes labeled red in Fig. 4d). However, the Taf2-mediated effect in binding to the nucleosome was reduced in the case of Taf14_{4K/Rmut}, in which one of the DNA-binding motifs (4K/R) is impaired while another DNA-binding motif (3K/R) remains (Supplementary Fig. 10f, g). Together, these data demonstrate that the linker region Taf14 is necessary for binding to the nucleosome. The linker region is occluded by Taf14_{ET} and Taf14_{YEATS} in the absence of ligands, whereas binding of the ligand, i.e. Taf2, releases the linker for the binding to DNA and NCP (Fig. 4e). Interestingly, despite the isolated Taf14_{YEATS} did not associate with H3K9_{CR}-NCP (Supplementary Fig. 10b) but bound well to the H3K9_{CR} peptide (Fig. 2c), the interaction of Taf14_{FL} with the unmodified NCP in the presence of Taf2 was enhanced when the crotonylated H3K9 modification was present in the nucleosome (Fig. 4c and Supplementary Fig. 10h, i).

Association of Taf14 with Taf2 and DNA is essential for transcriptional regulation. To define the role of the Taf14-Taf2 interaction in Taf14-dependent gene regulation, we performed reverse transcriptase (RT) quantitative PCR (RT-qPCR) analysis on a set of genes regulated by Taf14 (*YPR145*, *YER145*, *YKR039W*, and *ACS1*)¹⁸. *Taf14Δ* cells were transformed with either full-length wild type *TAF14*, a vector only control (to monitor the impact of Taf14 loss), or *taf14* mutants harboring either ET domain truncation (Taf14_{ΔET}), K149D/R150D/R151D/K161D/K163D/R164D/K166D mutations in the linker region that disrupt binding to DNA (Taf14_{7K/Rmut}) or W81A mutation in the YEATS domain that abolishes binding to H3K9_{acyl} (Taf14_{W81A}) (Supplementary Fig. 11). As shown in Fig. 4f, deletion of the ET domain resulted in a significant loss of Taf14-mediated suppression of all examined genes. The Taf14_{7K/Rmut} linker mutant also showed a loss of Taf14-mediated suppression similar or equal to the loss of Taf14_{ET} for *YPR145* and *YER145*. The loss of the H3K9_{acyl} binding activity of Taf14 greatly impacted one gene

(*YER145*) but had a minimal or no effect on the others. Taken together, these findings demonstrate a critical role of the Taf2-binding ET domain and the effect of the DNA-binding linker in transcriptional regulation by Taf14.

We next determined the impact of the *taf14* mutations on the recruitment of Taf14 to genes. As shown in Fig. 4g, ChIP-qPCR analysis of the same set of genes examined by RT-qPCR revealed that the loss of the ET domain and the linker region of Taf14 dramatically impaired its recruitment. These results were consistent with the effects imparted by these mutations on gene expression shown in Fig. 4f. In contrast, loss of Taf14 acyl-reading greatly increased the ability of Taf14 to associate with *YER145*, *YKR039W* and *ACS1*. Furthermore, a deletion of *TAF14* (*taf14Δ*) in cellular growth assays resulted in a slow growth phenotype which became more severe at lower temperatures (16 °C and 25 °C) (Fig. 4h). As expected, addition of wild type *TAF14* to *taf14Δ* rescued the observed growth defects, however the rescue effect at lower temperatures was diminished in *taf14Δ* cells expressing Taf14 mutants. Among the mutants, Taf14_{7K/Rmut} and Taf14_{W81A} rescued the growth defects of the *taf14Δ* cells to the highest degree but loss of the ET domain (Taf14_{ΔET}) led to only a partial rescue of the growth defects, supporting previous findings¹⁹ and pointing to the essential role of this domain in multifaceted functions of Taf14. While the DNA-binding activity of Taf14 was largely dispensable for the yeast growth in the spot assays (Fig. 4h), the Taf14_{7K/Rmut} was deficient in repression of Taf14-mediated suppression of a set of genes (Fig. 4f). The observed differential effects indicate that distinctive functions or a different degree of these functions of Taf14 are necessary for these processes. As Taf14 is a component of multiple complexes, differential effects can also arise depending on which and to what extent the Taf14-containing complexes dominate in these processes.

To assess the importance of Taf2_{CT} for biological activities of Taf2, we examined growth and viability of the Taf2_{ΔC} mutant in which the C-terminus (residues 1261–1407 of Taf2) was deleted in a plasmid shuffle assay. The Taf2_{ΔC} variant was shown to disrupt the ability of Taf14 to stably associate with TFIID¹⁴, and it displayed a mild slow growth phenotype at all temperatures tested (Fig. 4i). Notably, additionally mutating residues 775–780 (RRLRKR) of Taf2 to alanine, which were proposed to play a role in DNA binding of Taf2¹⁴, in the Taf2_{mutΔC} variant led to a strong slow growth phenotype at 25 °C and 30 °C and a temperature sensitive phenotype at 37 °C (Fig. 4i). Furthermore, steady-state mRNA levels of the TFIID-dominated ribosomal protein genes *RPS5*, *RPS3*, *RPS8a* and *RPS9b* in Taf2_{mutΔC} cells were 6- to 10-fold lower than mRNA levels of these genes in the wild type cells, while steady-state mRNA levels of the glycolytic SAGA dominated genes *PGK1* and *PYK1* were 2-fold lower than mRNA levels of these genes in the wild type cells (Fig. 4j). Finally, Taf2_{mutΔC} efficiently precipitated TFIID subunits, including



Taf3, Taf7 and Taf8 but not Taf14, in co-IP assays (Fig. 4k). Collectively, these results suggest that the Taf14-dependent and Taf2-dependent interactions with promoter DNA provide a synergistic effect substantially impacting both growth and steady state RNA levels and therefore are important for the regulation of gene transcription *in vivo*.

In conclusion, our findings demonstrate that Taf14 exists in an autoinhibited state in the absence of binding partners and is unable to bind to the nucleosome, and hence chromatin, even when H3K9acetyl is present. Binding of Taf2 to the ET domain and the YEATS domain of Taf14 eliminates the autoinhibition: it causes a substantial rearrangement of Taf14, resulting in

Fig. 3 DNA binding activity of Taf14_{Linker}. **a** Overlay of ¹H¹⁵,N HSQC spectra of Taf14_{FL} in the presence of increasing amount of 147 bp 601 DNA. Spectra are color coded according to the protein:DNA molar ratio. **b–e** EMSAs of 147 bp 601 DNA in the presence of increasing amounts of the indicated GST-Taf14 proteins. DNA:protein ratio is shown below gel images. Images are from single experiment. Architecture of each Taf14 construct is depicted above the gels, and mutations are indicated with a blue asterisk. **f, g** Superimposed ¹H¹⁵,N HSQC spectra of the indicated Taf14 proteins collected upon titration with 147 bp 601 DNA (**f**) or Taf2_{Linker} (**g**). Spectra are colored according to the protein:ligand molar ratio. **h** MST binding curves for the Taf2_{Linker} interaction with Taf14_{ET}. The K_d value represents average of two independent measurements, and error represents the standard deviation. *n* = 2. **i, j** Mass spectrometry analysis wheel diagram showing positions of DSSO-generated cross-links identified in the Taf14_{FL} (**i**) or Taf14_{FL} + 37 bp dsDNA (**j**). Line thickness relates to the number of cross-link identifications. Taf14 domains within the circle are colored as in Fig. 1a. **k** Dimensionless Kratky plots of SAXS analysis for the indicated samples: Taf14_{FL} (blue), Taf14_{FL} + Taf2_{CT} (orange), or Taf14_{FL} + 37 bp dsDNA (green). **l** A model of the conformational rearrangement in Taf14_{FL} upon binding to DNA. **m** Overlay of ¹H¹⁵,N HSQC spectra of Taf14_{FL} in the presence of increasing amounts of H3K9cr_{1–19} peptide. Spectra are colored according to the protein:peptide molar ratio. Source data are provided in a Source Data file.

the release of the linker region which binds to DNA via the positively charged motifs. Our data further suggest that Taf14 may play a role in TFIID promoter recognition in *S. cerevisiae* by concurrently targeting the specific chromatin state (H3K9acyl) and nucleosomal DNA. The two functions of Taf14 may serve to position TFIID for Taf-DNA interactions, and this positioning effect can lead to appropriate deposition of TATA-binding protein (TBP) at the TATA-box or TATA-like element. Indeed, ChIP-exo studies characterizing the positioning of TFIID on promoters genome wide, and independently ChEC-seq studies, have shown that TFIID/TBP positioning in TFIID dominated genes depends on the engagement with the nucleosome^{20,21}.

This work provides a model for promoter recognition by yeast TFIID at least in part through the histone mark-binding and nucleosomal DNA-binding activities of the Taf14 subunit, which in turn is mediated by another TFIID subunit, Taf2. The identification of the DNA-binding function of Taf14 is important. It has been shown that the Taf1 and Taf2 subunits in metazoan TFIID interact with the initiator (INR) element and the downstream promoter element (DPE)²², but the DPE and INR have not been identified or tested experimentally in *S. cerevisiae*. Although both metazoan and yeast TFIID contain TBP and can recognize the TATA-box, this sequence is often quite degenerate in promoters and requires additional contacts afforded by Taf subunits. These contacts could involve binding of bromodomains of Taf1 to acetylated histones and/or a PHD finger of Taf3 to H3K4me3 in metazoan TFIID and binding of the YEATS domain of Taf14 to H3K9acyl in yeast TFIID^{9,11,23–26}. Lastly, because the Taf2_{mut_ΔC} variant led to a strong slow growth phenotype, the functional importance of the region encompassing residues 775–780 of Taf2¹⁴ needs to be examined in future studies. Our study serves as benchmark for further analysis of the yeast TFIID engagement with chromatin.

Methods

Protein purification. The *S. cerevisiae* Taf14_{FL} (aa 1–244), Taf14_{YEATS} (aa 1–137), Taf14_{ΔET} (aa 1–174), Taf14_{ΔYEATS} (aa 148–243) and Taf14_{ET} (aa 162–243, aa 168–243) constructs were cloned into pGEX-6P1, pET22b, pDEST15 or pDEST17 vectors. The Taf14_{3K/R} (K149E/R150E/R151E), Taf14_{4K/R} (K161E/K163E/R164E/K166E), and Taf14_{ET} construct for NMR (K163N/R164Q/K166N) were generated using the Stratagene QuikChange Lightning Site-Directed Mutagenesis Kit. Sequences were confirmed by DNA sequencing. Taf14_{ET} (aa 168–242) construct was used for all experiments except NMR, where the Taf14_{ET} (aa 162–242 K163N/R164Q/K166N) construct was used. All proteins were expressed in *Escherichia coli* Rosetta-2 (DE3) pLysS cells grown in either Terrific Broth, Luria Broth or in minimal media supplemented with ¹⁵NH₄Cl (Sigma). Protein production was induced with 0.3 mM IPTG for 18 h at 16 °C. Bacteria were harvested by centrifugation and lysed by sonication. GST-fusion proteins were purified on glutathione agarose 4B beads (Thermo Fisher Sci). The GST-tag was cleaved with either PreScission or tobacco etch virus (TEV) protease, or the fusion proteins were eluted from the resin with reduced L-glutathione (Fisher). His-fusion proteins were incubated with Ni-NTA resin (ThermoFisher), washed and eluted with a gradient of imidazole. When necessary, removal of the His-tag was completed during an overnight dialysis, while incubating the protein with TEV protease. Proteins were further purified by column chromatography and concentrated in

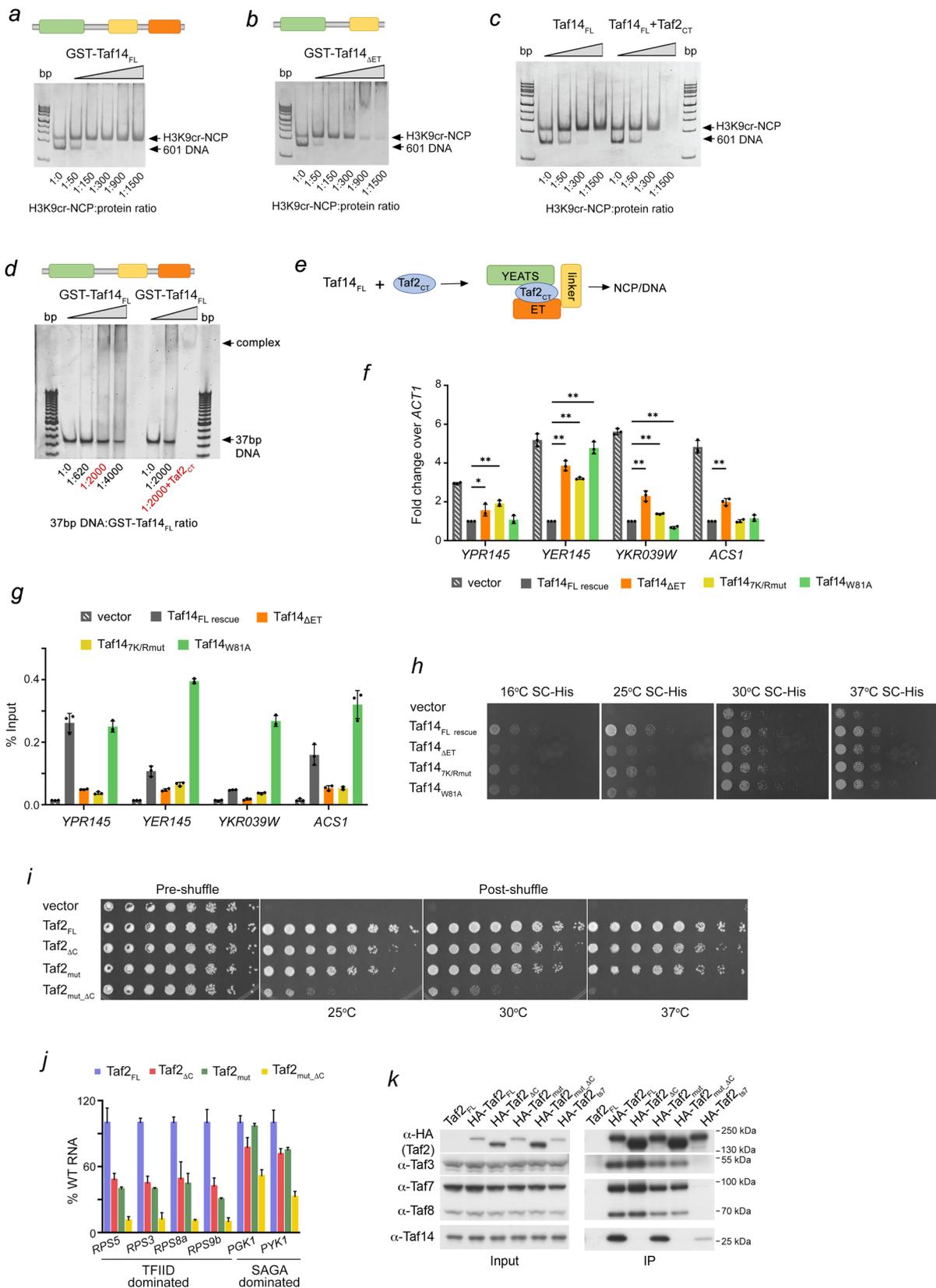
Millipore concentrators (Millipore). Proteins used in this study are summarized in the source data file.

DNA purification and preparation. Double stranded DNA containing the 147 bp 601 Widom DNA sequence in the pJ201 plasmid (32x) was transformed into DH5α cells. The plasmids containing the 601 sequence were amplified in TB media and purified using the PureLink HiPure Expi Plasmid Gigaprep Kit (Invitrogen K210009XP). Separation of the individual sequences was completed by digestion of the plasmid with EcoRV (NEB) followed by PEG and ethanol precipitation. After the final purification step, the DNA was suspended in water and the 601 DNA sequences were evaluated for purity by native polyacrylamide gel electrophoresis.

The double stranded DNA for EMSA and mass spectrometry were purchased as single stranded DNA (IDT) and annealed. DNA sequences used in this study are listed in the source data file. Complimentary DNA strands were combined in a 1:1 molar ratio in water in PCR tubes. Using a thermocycler to regulate temperature, the samples were brought to 95 °C for 20 min and then to 16 °C at a rate of 0.1 °C/second. All samples of annealed DNA were evaluated for purity by native polyacrylamide gel electrophoresis.

H3K9cr-nucleosome preparation and assembly. Recombinant Histone His-TEV-H2A, His-TEV-H2B, His-SUMO-TEV-H4 and H3K9cr or H3 were purified as described previously¹⁵. All histone pellets were dissolved in 6 M GuHCl buffer (6 M guanidinium hydrochloride, 20 mM Tris, 500 mM NaCl, pH 7.5), and concentration was measured by UV absorption at 280 nm (Biotek synergy H1 plate reader). To prepare H2A/H2B dimer, His-TEV-H2A and His-TEV-H2B were mixed in the molar ratio of 1:1, and 6 M GuHCl buffer was added to adjust total protein concentration to 4 μg/μl. Denatured His-TEV-H2A/His-TEV-H2B solution was dialyzed sequentially with stirring at 4 °C against 2 M TE buffer (2 M NaCl, 20 mM Tris, 1 mM EDTA, pH 7.8), 1 M TE buffer (1 M NaCl, 20 mM Tris, 1 mM EDTA), 0.5 M TE buffer (0.5 M NaCl, 20 mM Tris, 1 mM EDTA). The resulting dimer solution was centrifuged for 5 min at 4 °C to remove precipitates, and the concentration of dimer was determined by UV absorption at 280 nm. His-TEV-H3/His-SUMO-TEV-H4 tetramer was refolded as His-TEV-H2A/His-TEV-H2B above except without stirring and the total protein concentration was adjusted to 2 μg/μl prior to dialysis. Then His-TEV-H2A/His-TEV-H2B dimers were mixed with His-TEV-H3/His-SUMO-TEV-H4 tetramers in a molar ratio of 1:1 to generate histone octamers, and NaCl solid was added to adjust NaCl concentration to 2 M. The 147 bp biotinylated 601 DNA was prepared by Polymerase chain reaction with biotinylated primers and purified with PCR cleanup kit (#2360250 Epoch Life Science). Purified 147 bp DNA was re-dissolved in 2 M TE buffer and added to histone octamer solution in the molar ratio of 0.85:1. 2 M TE buffer was added to adjust final 147 bp DNA concentration to 2–3 μM. The DNA histone mixture solution was then transferred to a dialysis bag and placed inside about 200 ml 2 M TE buffer, while stirring at room temperature. Tris buffer with no salt (20 mM Tris) was slowly added into the 2 M salt buffer through a liquid transfer pump (#23609-170 VWR®). Nucleosomes formed when salt concentration was reduced to around 150 mM (measured by EX170 salinity meter), and the DNA histone mixture solution was further dialyzed into low salt Tris buffer (20 mM Tris, 20 mM NaCl, 0.5 mM EDTA, pH 7.8). Precipitates were removed by centrifuge, and the nucleosome concentration was measured by the absorbance at 260 nm (A260) using the Biotek synergy H1 plate reader. His-TEV protease was added to nucleosome solutions (TEV:nucleosome 1:30, w-w) to remove all the histone tags after incubation for 1 h at 37 °C, and finally all the His tagged impurities were removed from the nucleosome solution by Ni²⁺-NTA resin (Thermo Fisher #88221).

EMSA. EMSAs were performed by mixing increasing amounts of Taf14 proteins with 0.25 pmol/lane of 601 DNA, dsDNA (36 bp AT rich, 36 bp GC rich, 37 bp non-specific), NCP, or H3K9cr-NCP in 20–25 mM Tris-HCl pH 7.5, 100–150 mM NaCl and 0.2 mM ethylenediaminetetraacetic acid (EDTA) (reactions with NCP or H3K9cr-NCP were supplemented with 20% glycerol) in a 10 μL reaction volume. For the samples containing the DNA ladder (O'RangeRuler 5 bp DNA Ladder,



ThermoSci) the buffer was 20 mM Tris-HCl pH 7.5, 150 mM NaCl and 5 mM DTT. The indicated Taf14 constructs were preincubated with Taf2_{CT} at a 1:2 (Fig. 4c, Supplementary Figs. 5d, e, 10e, g, h) or 1:10 (Fig. 4d) molar ratio for a minimum of 1 h on ice prior to addition of DNA or NCP. Reaction mixtures were incubated at 4°C for 10 min (2–2.5 μl of loading dye was added to each sample) and loaded onto a 5 or 10% native polyacrylamide gel. Electrophoresis was performed in 0.2x Tris-borate-EDTA (TBE) at 80–120 V on ice. The gels were stained

with SYBR Gold (Thermo Fisher Sci) and visualized by Blue LED (UltraThin LED Illuminator-GelCompany).

NMR experiments. Nuclear magnetic resonance (NMR) experiments were performed at 298 K on Varian INOVA 900 MHz and 600 MHz spectrometers equipped with cryogenic probes. The NMR samples contained 0.1–0.3 mM

Fig. 4 Binding of Taf14_{FL} to the nucleosome requires binding to Taf2_{CT}. **a–c** EMSAs of H3K9cr-NCP₁₄₇ in the presence of increasing amounts of the indicated GST-Taf14 or Taf14 proteins (and Taf2_{CT} in c). Architecture of each Taf14 construct is depicted above gels. Images are from single experiment. **d** EMSA of 37 bp dsDNA in the presence of increasing amounts of the indicated GST-Taf14 (and Taf2_{CT} in lane 9). Architecture of the Taf14 construct is depicted above the gel. GST-Taf14_{FL} and Taf2_{CT} were premixed at a 1:10 molar ratio prior to incubation with DNA. **e** A schematic of the mechanism for the engagement of Taf14_{FL} with NCP: following the interaction with Taf2_{CT}, the released Taf14_{Linker} binds DNA/NCP. **f** Reverse transcriptase qPCR analysis of various transcripts in the *TAF14* full-length rescued strain, vector only strain, *taf14* ET delete strain (Taf14_{ΔET}), and *taf14* mutant strains (Taf14_{7k/Rmut} and Taf14_{W81A}). The mean ± SD are calculated from three biological replicates. *ACT1* was used as a normalization control. Two-tailed multiple T tests were performed, **P* < 0.05; ***P* < 0.005. **g** Chromatin immunoprecipitation assays for promoter-bound Taf14 in the same strains and at the same set of genes examined in (f). The mean ± SD are calculated from three biological replicates. **h** Yeast spotting assays with wild type (BY4741) or *taf14Δ* strains transformed with the indicated plasmids and grown on SC-His medium at the indicated temperatures. **i** Plasmid shuffle assay. *Taf2* null cells transformed with wild type *TAF2* or LEU2-marked mutant *taf2*-containing variants plasmids were plated onto SC-Leu (pre-shuffle) or SC-Leu + 5-FOA (post-shuffle). Plate images are representative of at least two biological replicates. **j** Steady-state RNA analysis. qRT-PCR was performed on total yeast RNA. Data is presented as the ratio of indicated target gene/*U3* with wild type *TAF2* set as 100%. Each pRT-PCR reaction was performed in triplicate using three biological experiments, with bars representing mean and errors representing standard deviation. **k** Co-IP of the HA-Taf2 proteins with the indicated TFIID proteins was analyzed by immunoblotting using anti-HA and anti-Taf polyclonal antibodies. Representative immunoblots from two biological replicates are shown. Source data are provided in a Source Data file.

uniformly ¹⁵N-labeled Taf14 (Taf14_{FL}, Taf14_{YEATS}, Taf14_{ΔET}, or Taf14_{ET} (aa 162–243 K163N/R164Q/K166N)) in 50 mM Tris-HCl pH 6.6–7.5, 150 mM NaCl, 0–5 mM DTT or 0.2 mM TCEP or 1x PBS pH 6.6–7.5 (supplemented with up to an additional 100 mM NaCl), 0–5 mM DTT and 8–10% D₂O. Binding was characterized by monitoring chemical shift changes in ¹H¹⁵N HSQC spectra of the proteins induced by the addition of Taf2_{CT} (aa 1392–1402), Taf2_{CT(V1397D)} (aa 1392–1402), Taf2_{CT(I1399D)} (aa 1392–1402), Taf14_{Linker} (aa 145–169), or H3K9cr_{5–13} peptides (all peptides synthesized by Synpeptide) or 601 DNA. NMR data were processed and analyzed using NMRPipe as previously described²⁷. NMR assignments of Taf14_{YEATS} were taken from²⁸.

X-ray crystallography. Taf14_{ET} was concentrated to ~3.5 mM (in 50 mM Tris-HCl pH 7.5, 150 mM NaCl, 0.2 mM TCEP). The protein was incubated with the Taf2 peptide (Synpeptide, aa 1392–1402) at a molar ratio of 1:2.5 for 1 h. The protein-peptide solution was concentrated to ~6.6 mg/mL in buffer (20 mM Tris-HCl pH 7.5, 100 mM NaCl) for crystallization. Crystals were grown using sitting-drop vapor diffusion method at 23 °C by mixing 0.6 μL of protein with 0.6 μL of well solution composed of 2.4 M sodium malonate pH 7.0. Crystals were cryoprotected with 30% (v/v) glycerol. X-ray diffraction data were collected on a beamline 4.2.2 at Advance Light Source. Indexing and scaling were completed using HKL2000²⁹. The structure was determined using Phenix phaser with the Taf14_{ET} (PDH ID: 6LQZ) structure as a search model. Model building was carried out with Coot³⁰, and refinement was performed with Phenix refine³¹. Crystallographic statistics for the Taf14_{ET}:Taf2 structure are shown in Supplementary Table 1.

MST. Microscale thermophoresis (MST) experiments were performed using a Monolith NT.115 instrument (NanoTemper). All experiments were performed using SEC purified Taf14 proteins (His-Taf14_{FL}, His-Taf14_{ET}, His-Taf14_{YEATS}, or Taf14_{ET}) in a buffer containing 20 mM Tris-HCl (pH 7.5), 150 mM NaCl and 0–0.2 mM TCEP. For the Taf14_{ET}:Taf2_{CT} interaction, the concentration of fluorophore, FAM-labeled Taf2_{CT} peptide (Synpeptide, 1392–1703 aa), was 20 nM. To monitor binding between Taf2_{CT} and Taf14_{FL} or Taf14_{YEATS}, and Taf14_{Linker} and Taf14_{ET}, His-tagged proteins were labeled using a His-Tag Labeling Kit RED-tris-NTA (2nd Generation, Nanotemper), and the concentration of each labeled protein was kept at 20 nM. Dissociation constants were determined using a direct binding assay in which increasing amounts of unlabeled binding partner (Taf2_{CT} peptide or Taf14_{Linker} peptide) were added stepwise. The measurements were performed at 80% LED and medium MST power with 3 s steady state, up to 20 s laser on time and 1 s off time. The *K_d* values were calculated using MO Affinity Analysis software (NanoTemper) (*n* = 2 or 3). Figures were generated in GraphPad PRISM.

Sample preparation for cross-linking mass spectrometry. Each sample contained 1 mM SEC purified Taf14_{FL}, without ligands or with 37 bp double stranded DNA or Taf2_{CT} peptide (aa 1392–1403) (9.6 and 10.8 mM respectively) in 25 mM Hepes pH 7.5, 150 mM NaCl for 10 min prior to the addition of 2.5 mM DSSO (or DMSO for control reaction). The samples were incubated for 1 h at 23 °C, followed by the addition of 20 mM Tris-HCl pH 7.5 for 10 min to stop the reaction. DSSO cross-linked Taf14_{FL} samples were denatured, reduced and alkylated with 5% (w/v) SDS, 10 mM tris(2-carboxyethylphosphine) (TCEP), 40 mM chloroacetamide, 50 mM Tris-HCl, pH 8.5 boiling 10 min. The lysates were cleared via centrifugation at 13,000 × *g* for 10 min at ambient temperature. Cross-linked Taf14_{FL} samples were digested using the SP3 method³². Carboxylate-functionalized speedbeads (Cytiva Life Sciences) were added. Addition of acetonitrile to 80% (v/v) caused the proteins to bind to the beads. The beads were washed twice with 80% (v/v) ethanol

and twice with 100% acetonitrile. Samples were digested in 50 mM Tris buffer, pH 8.5, with 1 μg Lys-C/Trypsin (Promega) incubated at 37 °C overnight rotating to mix. Tryptic peptides were desalted using the carboxylate-functionalized speedbeads with the addition of 95% (v/v) acetonitrile twice and then eluted with 3% (v/v) acetonitrile, 1% (v/v) trifluoroacetic acid and dried in a speedvac vacuum centrifuge.

MS analysis of DSSO cross-linked proteins. Samples were suspended in 3% (v/v) acetonitrile/0.1% (v/v) trifluoroacetic acid and approximately 2 picomoles of tryptic peptides were directly injected onto a reversed-phase C18 1.7 μm, 130 Å, 75 mm × 250 mm M-class column (Waters), using an Ultimate 3000 UPLC (Thermo Scientific). Peptides were eluted at 300 nL/minute using a gradient from 2% to 20% acetonitrile over 40 min and detected using a Q-Exactive HF-X mass spectrometer (Thermo Scientific). Precursor mass spectra (MS1) were acquired at a resolution of 120,000 from 350 to 1600 *m/z* with 25 eV in-source collision induced dissociation, an AGC target of 3E6 and a maximum injection time of 50 ms. Dynamic exclusion was set for 5 seconds with a mass tolerance of ±10 ppm. Precursor peptide ion isolation width for MS2 fragment scans was 1.4 Da, and the top 15 most intense ions were sequenced. All MS2 sequencing was performed using higher energy collision dissociation (HCD) at 30% normalized collision energy. An AGC target of 1E5 and 100 ms maximum injection time was used. Raw files were searched against the Uniprot yeast sequences for Taf14 (accession P35189) using pLink2 with the enzyme specificity set to trypsin with two missed cleavages. The peptide mass range and length were 500–6000 and 5–60, respectively. Precursor and fragment tolerance were both set to ±20 ppm. Final results were filtered with the MS1 tolerance set to ±8 ppm. FDR was set to ≤1% at spectral level and E-values were calculated. Using a proXL conversion script (<https://github.com/yeastrc/proxl-import-link2>), xml files were generated and submitted to the proXL website. The crosslinking MS circle diagrams were generated using proXL and the data has been deposited and is available through the proXL web interface at <https://proxl-ms.org/>.

SAXS. Small angle X-ray scattering (SAXS) experiments were performed at beamline 12-ID-B, the Advanced Photon Sources facility (Argonne National Laboratory, Lemont, IL). Photon energy was 13.3-keV (wavelength λ was 0.932 Å). Scattered x-ray intensities were measured using a Pilatus 2M detector. A sample-to-detector distance of 1.9 m was used to obtain *q* range of 0.005 < *q* < 0.88 Å⁻¹. Here *q* is magnitude of the scattering vector, *q* = (4π/λ)sinθ, where 2θ is the scattering angle and λ is the wavelength of the radiation. The concentration series measurements were carried out to remove the scattering contribution due to interparticle interactions and to extrapolate the data to infinite dilution. The concentrations of Taf14_{FL} were 2.7, 3.6, 7.2 μM in a buffer consisting of 50 mM HEPES, 150 mM NaCl, 0.2 mM TCEP pH 7.5. For the solution of the complex, the concentration of Taf14_{FL} is the same, with the molar ratio of Taf14_{FL} to ligand (Taf2_{CT} peptide or DNA) of 1:4. The ligand (concentrations of 8.1, 10.8, 21.6 μM) in buffer were used as background for the complex. The solutions were loaded in a flow cell made of a cylindrical quartz capillary (1.5-mm diameter and 10-μm wall-thickness) for the measurements. To minimize radiation damage and obtain a good signal-to-noise ratio, 45 images were taken for each sample and buffer with an exposure time of 0.5 s. The two-dimensional scattering images were converted to one-dimensional SAXS curves through radial averaging after solid angle correction and then normalizing with the intensity of the transmitted x-ray beam, then were averaged after elimination of outliers using the software package developed at beamline 12-ID-B.

The buffer background subtraction and intensity extrapolation to infinite dilution were carried out using MatLab script. The radius of gyration *R_g* and

intensity at zero angle $I(0)$ were generated from Guinier plot in the range of $qR_g < 1.3$. For comparison, R_g and $I(0)$ were also calculated in real and reciprocal spaces using program GNOM in q range up to 0.30 \AA^{-1} . The pair-distance distribution function $P(r)$ and maximum dimension (D_{\max}) were also calculated using GNOM. In the process of producing $P(r)$, the D_{\max} was chosen in the way that the $P(r)$ curves approach zero at D_{\max} in a smooth concave manner. The molecular weights were estimated based on Bayesian statistics method. The obtained structural parameters and molecular weights are shown in Supplementary Fig. 4a, and Guinier plots are shown in Supplementary Figs. 4c and 5c. The R_g based dimensionless Kratky plots presented as $(qR_g)2I(q)/I(0)$ vs. qR_g was used for analysis of flexibility and/or degree of unfolding in samples³³.

Spotting assays. Spotting assays for wild type and mutated Taf14 were performed with serial dilutions of saturated overnight cultures of the indicated yeast strains, and growth was assessed after 2–3 d (dilution ratio of 1:5). All yeast strains used in this study are described in Supplementary Table 2, and plasmids used in this study are presented in Supplementary Table 3.

qPCR. For Taf14: reverse transcriptase qPCR was performed as described previously¹⁵. *taf14Δ* cells transformed with either vector only, *TAF14* full-length or the indicated *taf14* mutants were cultured in rich media YPD overnight at 30 °C. Ten OD₆₀₀ of cell pellets were collected from each strain and phenol/chloroform extraction was used to extract RNA. Residual DNA was digested by DNase I. Complementary DNA (cDNA) was produced from 2 mg of extracted RNA using IScript kit (BioRad). cDNA was then diluted to 0.1 mg/ml for reverse transcriptase qPCR using SYBR green. *ACT1* was used as the normalization control.

For Taf2: For RNA analysis, total RNA was subjected to reverse transcription with oligo(dT) and a gene specific U3 primer. qPCR was performed with SYBR green supermix (Bio-Rad). Samples were quantified using the relative standard curve method, normalized to U3, and expressed as a percentage of the average of the *HA-TAF2* strain. All qPCR reactions were performed in triplicate. Two biological replicates were used for each sample.

Chromatin immunoprecipitation (ChIP). ChIP assays were performed as previously described³⁴. *taf14Δ* cells transformed with either vector only, *TAF14* full-length or the indicated *taf14* mutants were cultured in rich media YPD overnight at 30 °C. Fifty OD₆₀₀ of cell pellets were collected from each strain and fixed in 1% formaldehyde for 15 min prior to processing for chromatin immunoprecipitation using a rabbit antibody directed against Taf14 (gift from Joseph Reese), 5 μl per 500 μl reaction. Primers directed against the gene regions are shown in the source data file.

Plasmid shuffle assay. The assay was performed essentially as described¹⁴. Briefly, a *taf2* null strain (JFTAF2del) was used for all experiments. For plasmid shuffle analyses, JFTAF2del was transformed an empty p415ADH plasmid, a p415ADH-*TAF2* plasmid, a p415ADH-*HA_{x3}NLS-TAF2* plasmid, or a p415ADH-*HA_{x3}NLS-TAF2* mutant plasmid, and transformants were grown on SC – Leu plates. Colonies were grown to saturation in SC – Leu at 30 °C prior to spotting and grown at various temperatures between 20 °C and 37 °C (cell dilution ratio of 1:4).

Co-IP. Co-immunoprecipitations were performed using pseudodiploids containing both a wild type *URA3*-marked *TAF2* gene and a WT or mutant *LEU2*-marked *TAF2* gene. All variants indicated with HA include a 3xHA-NLS N-terminal tag. Two mg of soluble protein extract were incubated with 2.5 μg of anti-HA 12CA5 mAb and 50 ng/μl ethidium bromide in a total volume of 412 μl at 4 °C for 2 h. Immune complexes were captured with 10 μl of protein A-Sepharose beads (Life Technologies), separated via SDS-PAGE and electrotransferred. Proteins were detected with anti-HA and anti-Taf polyclonal antibodies, followed by anti-rabbit Fc HRP-conjugated secondary antibodies (Abcam, 1:10,000 dilution) for chemiluminescent detection. Antibodies: anti-HA blot 3F10 (Sigma-Millipore, cat # 12013819001, 1:2000 dilution), anti-HA IP 12CA5 (made in house, 2.5 μg per reaction), anti-Taf3 (made in house, 1:1000 dilution), anti-Taf7 (made in house, 1:2000 dilution), anti-Taf8 (made in house, 1:1000 dilution), anti-Taf14 (made in house, 1:2000 dilution).

Reporting summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The data that support this study are available from the corresponding author upon reasonable request. Coordinates and structure factors of the Taf14 ET-Taf2 CT complex have been deposited in the Protein Data Bank under the accession code 7UHE. The mass spec data have been deposited to the PRIDE database under accession number PXD033397. Source data are provided with this paper.

Received: 14 September 2021; Accepted: 20 May 2022;

Published online: 08 June 2022

References

- Poon, D. et al. Identification and characterization of a TFIID-like multiprotein complex from *Saccharomyces cerevisiae*. *Proc. Natl Acad. Sci. USA* **92**, 8224–8228 (1995).
- Henry, N. L. et al. TFIIF-TAF-RNA polymerase II connection. *Genes Dev.* **8**, 2868–2878 (1994).
- Kabani, M., Michot, K., Boschiero, C. & Werner, M. Anc1 interacts with the catalytic subunits of the general transcription factors TFIID and TFIIF, the chromatin remodeling complexes RSC and INO80, and the histone acetyltransferase complex NuA3. *Biochemical biophysical Res. Commun.* **332**, 398–403 (2005).
- Robinson, P. J. et al. Structure of a complete mediator-RNA polymerase II pre-initiation complex. *Cell* **166**, 1411–1422 e1416 (2016).
- Kolesnikova, O. et al. Molecular structure of promoter-bound yeast TFIID. *Nat. Commun.* **9**, 4666 (2018).
- Shen, X. Preparation and analysis of the INO80 complex. *Methods Enzymol.* **377**, 401–412 (2004).
- Cairns, B. R., Henry, N. L. & Kornberg, R. D. TFG/TAF30/ANC1, a component of the yeast SWI/SNF complex that is similar to the leukemogenic proteins ENL and AF-9. *Mol. Cell. Biol.* **16**, 3308–3316 (1996).
- John, S. et al. The something about silencing protein, Sas3, is the catalytic subunit of NuA3, a γ TAF(II)30-containing HAT complex that interacts with the Spt16 subunit of the yeast CP (Cdc68/Pob3)-FACT complex. *Genes Dev.* **14**, 1196–1208 (2000).
- Shanle, E. K. et al. Association of Taf14 with acetylated histone H3 directs gene transcription and the DNA damage response. *Genes Dev.* **29**, 1795–1800 (2015).
- Andrews, F. H., Shanle, E. K., Strahl, B. D. & Kutateladze, T. G. The essential role of acetyllysine binding by the YEATS domain in transcriptional regulation. *Transcription* **7**, 14–20 (2016).
- Andrews, F. H. et al. The Taf14 YEATS domain is a reader of histone crotonylation. *Nat. Chem. Biol.* **12**, 396–398 (2016).
- Klein, B. J. et al. Yaf9 subunit of the NuA4 and SWR1 complexes targets histone H3K27ac through its YEATS domain. *Nucleic acids Res.* **46**, 421–430 (2018).
- Chen, G. et al. Taf14 recognizes a common motif in transcriptional machineries and facilitates their clustering by phase separation. *Nat. Commun.* **11**, 4206 (2020).
- Feigler, J. T. & Weil, P. A. The C terminus of the RNA polymerase II transcription factor IID (TFIID) subunit Taf2 mediates stable association of subunit Taf14 into the yeast TFIID complex. *J. Biol. Chem.* **291**, 22721–22740 (2016).
- Klein, B. J. et al. Structural insights into the pi-pi stacking mechanism and DNA-binding activity of the YEATS domain. *Nat. Commun.* **9**, 4574 (2018).
- Gatchalian, J. et al. Accessibility of the histone H3 tail in the nucleosome for binding of paired readers. *Nat. Commun.* **8**, 1489 (2017).
- Weaver, T. M., Morrison, E. A. & Musselman, C. A. Reading more than histones: the prevalence of nucleic acid binding among reader domains. *Molecules* **23**, 2614 (2018).
- Gowans, G. J. et al. Recognition of histone crotonylation by Taf14 links metabolic state to gene expression. *Mol. cell* **76**, 909–921 e903 (2019).
- Schulze, J. M., Kane, C. M. & Ruiz-Manzano, A. The YEATS domain of Taf14 in *Saccharomyces cerevisiae* has a negative impact on cell growth. *Mol. Genet. Genomics: MGG* **283**, 365–380 (2010).
- Grunberg, S., Henikoff, S., Hahn, S. & Zentner, G. E. Mediator binding to UASs is broadly uncoupled from transcription and cooperative with TFIID recruitment to promoters. *EMBO J.* **35**, 2435–2446 (2016).
- Rhee, H. S. & Pugh, B. F. Genome-wide structure and organization of eukaryotic pre-initiation complexes. *Nature* **483**, 295–301 (2012).
- Louder, R. K. et al. Structure of promoter-bound TFIID and model of human pre-initiation complex assembly. *Nature* **531**, 604–609 (2016).
- Jacobson, R. H., Ladurner, A. G., King, D. S. & Tjian, R. Structure and function of a human TAFII250 double bromodomain module. [see comment]. *Science* **288**, 1422–1425 (2000).
- Matangkasombut, O., Buratowski, R. M., Swilling, N. W. & Buratowski, S. Bromodomain factor 1 corresponds to a missing piece of yeast TFIID. *Genes Dev.* **14**, 951–962 (2000).
- Vermeulen, M. et al. Selective anchoring of TFIID to nucleosomes by trimethylation of histone H3 lysine 4. *Cell* **131**, 58–69 (2007).
- Laubert, S. M. et al. H3K4me3 interactions with TAF3 regulate preinitiation complex assembly and selective gene activation. *Cell* **152**, 1021–1036 (2013).

27. Klein, B. J. et al. The histone-H3K4-specific demethylase KDM5B binds to its substrate and product through distinct PHD fingers. *Cell Rep.* **6**, 325–335 (2014).
28. Zhang, W., Zhang, J., Zhang, X., Xu, C. & Tu, X. Solution structure of the Taf14 YEATS domain and its roles in cell growth of *Saccharomyces cerevisiae*. *Biochemical J.* **436**, 83–90 (2011).
29. Otwinowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326 (1997).
30. Emsley, P., Lohkamp, B., Scott, W. G. & Cowtan, K. Features and development of Coot. *Acta Crystallogr. Sect. D, Biol. Crystallogr.* **66**, 486–501 (2010).
31. Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. Sect. D, Biol. Crystallogr.* **66**, 213–221 (2010).
32. Hughes, C. S. et al. Single-pot, solid-phase-enhanced sample preparation for proteomics experiments. *Nat. Protoc.* **14**, 68–85 (2019).
33. Durand, D. et al. NADPH oxidase activator p67(phox) behaves in solution as a multidomain protein with semi-flexible linkers. *J. Struct. Biol.* **169**, 45–53 (2010).
34. Zhang, J., Gundu, A. & Strahl, B. D. Recognition of acetylated histone by Yaf9 regulates metabolic cycling of transcription initiation and chromatin regulatory factors. *Genes Dev.* **35**, 1678–1692 (2021).

Acknowledgements

We thank Xiaobing Zuo for help with SAXS experiments. This work was supported in part by grants from the NIH: HL151334, GM125195, GM135671, CA252707 and AG067664 to T.G.K., GM49985 and DK121366 to J.T.F., GM126900 to B.D.S., GM127575 to W.R.L. and from Welch Foundation: A-1715 to W.R.L. We acknowledge the use of the SAXS Core facility of Center for Cancer Research (CCR), National Cancer Institute (NCI). This project has been funded in part with federal funds from the National Cancer Institute, NIH, under contract HHSN26120080001E. The content of this publication does not necessarily reflect the views or policies of the Department of Health and Human Services, nor does mention of trade names, commercial products, or organizations imply endorsement by the U.S. Government. This research was supported in part by the Intramural Research Program of the NIH, National Cancer Institute, Center for Cancer Research. This research used 12-ID-B beamline of the Advanced Photon Source, a U.S. Department of Energy (DOE) Office of Science User Facility operated for the DOE Office of Science by Argonne National Laboratory under Contract No. DE-AC02-06CH11357.

Author contributions

B.J.K., J.T.F., J.Z., C.C.E., L.F., R.K.S. and W.W.W. performed experiments and together with L.R.S., T.L., K.C.H., W.R.L., Y.-X.W., B.D.S., P.A.W. and T.G.K. analyzed the data. B.J.K., J.T.F. and T.G.K. wrote the manuscript with input from all authors.

Competing interests

The authors declare no competing interests.

Additional information

Supplementary information The online version contains supplementary material available at <https://doi.org/10.1038/s41467-022-30937-w>.

Correspondence and requests for materials should be addressed to Tatiana G. Kutateladze.

Peer review information *Nature Communications* thanks the anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available

Reprints and permission information is available at <http://www.nature.com/reprints>

Publisher's note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.



Open Access This article is licensed under a Creative Commons Attribution 4.0 International License, which permits use, sharing, adaptation, distribution and reproduction in any medium or format, as long as you give appropriate credit to the original author(s) and the source, provide a link to the Creative Commons license, and indicate if changes were made. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in a credit line to the material. If material is not included in the article's Creative Commons license and your intended use is not permitted by statutory regulation or exceeds the permitted use, you will need to obtain permission directly from the copyright holder. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>.

© The Author(s) 2022