Title
Structural and functional insight into TAF1–TAF7, a subcomplex of transcription factor II D

Journal
Proceedings of the National Academy of Sciences of the United States of America, 111(25)

ISSN
0027-8424

Authors
Bhattacharya, Suparna
Lou, Xiaohua
Hwang, Peter
et al.

Publication Date
2014-06-24

DOI
10.1073/pnas.1408293111

Peer reviewed

Structural and functional insight into TAF1–TAF7, a subcomplex of transcription factor II D

Suparna Bhattacharya[a], Xiaohua Lou[a,b], Peter Hwang[c], Kanagalaghatta R. Rajashankar[d], Xiaoping Wang[e], Jan-Åke Gustafsson[b], Robert J. Fletterick[c], Raymond H. Jacobson[e], and Paul Webb[a,1]

[a]Genomic Medicine Program, Houston Methodist Research Institute, Houston, TX 77030; [b]Center for Nuclear Receptors and Cell Signaling, University of Houston, Houston, TX 77204; [c]University of California Medical Center, San Francisco, CA 94158; [d]The Northeastern Collaborative Access Team and Department of Chemistry and Chemical Biology, Cornell University, Argonne National Laboratory, Argonne, IL 60439; and [e]Department of Molecular Biology and Biochemistry, MD Anderson Cancer Center, Houston, TX 77030

Contributed by Jan-Åke Gustafsson, May 13, 2014 (sent for review April 17, 2014; reviewed by Fraydoon Rastinejad and Stephen K. Burley)

Transcription factor II D (TFIID) is a multiprotein complex that nucleates formation of the basal transcription machinery. TATA binding protein-associated factors 1 and 7 (TAF1 and TAF7), two subunits of TFIID, are integral to the regulation of eukaryotic transcription initiation and play key roles in preinitiation complex (PIC) assembly. Current models suggest that TAF7 acts as a dissociable inhibitor of TAF1 histone acetyltransferase activity and that this event ensures appropriate assembly of the RNA polymerase II-mediated PIC before transcriptional initiation. Here, we report the 3D structure of a complex of yeast TAF1 with TAF7 at 2.9 Å resolution. The structure displays novel architecture and is characterized by a large predominantly hydrophobic heterodimer interface and extensive cofolding of TAF subunits. There are no obvious similarities between TAF1 and known histone acetyltransferases. Instead, the surface of the TAF1–TAF7 complex contains two prominent conserved surface pockets, one of which binds selectively to an inhibitory trimethylated histone H3 mark on Lys27 in a manner that is also regulated by phosphorylation at the neighboring H3 serine. Our findings could point toward novel roles for the TAF1–TAF7 complex in regulation of PIC assembly via reading epigenetic histone marks.

initiation complex | protein structure | protein–protein interaction | X-ray crystallography

The general transcription factor II D (TFIID) plays a central role in recognition of the core promoter element and mediates accurate transcription initiation by RNA Polymerase (Pol) II for a large class of genes. TFIID nucleates the formation of the preinitiation complex (PIC) at the transcriptional start site by recruiting other general transcription factors (TFII-A, -B, -E, -F, and -H), along with RNA Pol II. TFIID is assembled in a stepwise manner, with a symmetric TFIID core complex recruited first, followed by further recruitment of additional TATA binding protein (TBP)-associated factors (TAFs) to form the complete asymmetric holo–TFIID complex (1). This megadalton-sized multiprotein assembly, comprised of TBP and 13 evolutionary conserved TAFs (2), is organized into a trilobed structure (3) and undergoes striking rearrangements upon binding to TFIIA and DNA (4). TAF subunits serve multiple functions within the TFIID holocomplex. In addition to TBP, TAF1, TAF2, TAF6, and TAF9 are also involved in recognition of DNA initiator and promoter elements. Moreover, TFIID can behave as an epigenetic effector, capable of recognizing posttranslational histone modifications associated with activated transcription. Eukaryotic TAF1 contains a double bromodomain that recognizes acetylated histones, and TAF3 contains a plant homeo domain (PHD) that binds to histone H3 methylated at lysine 4 (5, 6). In addition to roles in basal transcription, TFIID is also associated with diseases. Overexpression of TAF1, which acts as a specific coactivator of androgen receptor, is related to the progression of human prostate cancer (7). Additionally, alterations of TAF gene copy numbers and mutations are reported in high-grade serous ovarian cancer (8).

TAF1, the largest subunit of TFIID, plays a particularly important role in TFIID complex activity (9, 10). Unlike other TAFs, it spans two lobes (from C through A) within the trilobed structure of TFIID, as indicated by EM and immunomapping studies, with the TAF1 C terminus located in the A lobe (3, 11). TAF1 is comprised of multiple separable domains that variously harbor HAT and kinase activities, a TAND domain that binds TBP, DNA binding functions, and the aforementioned double bromodomain that contacts acetylated histone tails (5, 9, 10). Yeast TAF1 (yTAF1) retains the HAT, TAND, and C-terminal DNA recognition domains (12), and ∼90% of yeast gene expression is TAF1 dependent (13). Current reports suggest that TAF7 exerts transcriptional checkpoint control activity by behaving as a dissociable inhibitor of TAF1 HAT activity; this event serves to ensure correct PIC assembly for effective Pol II-mediated transcription initiation and is essential for progression through the cell cycle (14, 15).

Although crystal structures of small fragments of the TAF1 double bromodomain (5) and TAND domain bound to TBP (16) are available, the structure of the large putative TAF1 HAT domain and the basis of its molecular interaction with TAF7 remain unknown. To gain insight into TAF1–TAF7

Significance

Transcription factor II D (TFIID) is a multiprotein complex that is essential for gene transcription. Together, TATA binding protein-associated factor 1 (TAF1), the biggest TFIID subunit, and TAF7 form an important control point for transcriptional initiation. Although current models suggest that TAF7 binds TAF1 to block its intrinsic histone acetyltransferase (HAT) activity, almost nothing is known about the molecular basis of TAF1–TAF7 (TAF1/7) interaction and TAF1 activity. Here, we report the atomic structure of the yeast TAF1/7 heterodimer and probe its function using biochemical techniques. Our structure suggests that yeast TAF1 is not a HAT and instead reveals that TAF1/7 displays the unexpected capacity to bind a specific repressive histone mark. This raises the possibility that TFIID binds repressive chromatin marks to control gene expression.

Author contributions: S.B., R.H.J., and P.W. designed research; S.B., X.L., P.H., K.R.R., and X.W. performed research; S.B., X.L., and K.R.R. performed X-ray data collection; S.B. solved the structure; X.W. performed limited proteolysis of TFIID; S.B., X.L., and P.H. analyzed data; P.W. supervised the project; and S.B., J.-Å.G., R.J.F., and P.W. wrote the paper.

Reviewers: F.R., Sanford-Burnham Medical Research Institute; and S.K.B, Rutgers, The State University of New Jersey.

The authors declare no conflict of interest.

Data deposition: The atomic coordinates have been deposited in the Protein Data Bank, www.pdb.org (PDB ID code 4OY2).

[1]To whom correspondence should be addressed. E-mail: pwebb@tmhs.org.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1408293111/-/DCSupplemental.

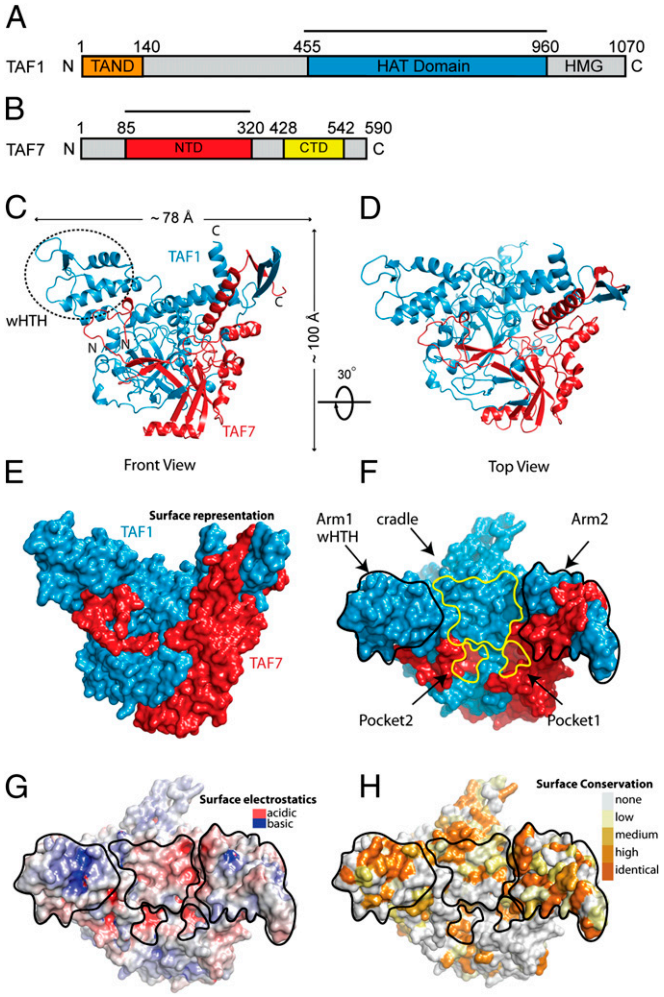

Fig. 1. Structural overview of the yTAF1/7 complex. (*A* and *B*) Domain structures of yTAF7 and yTAF1. The bars mark the crystallized fragments of the yTAF1/7 complex. (*C* and *D*) Ribbon representation of the yTAF1/7 heterodimer colored according to *A* and *B*. (*E* and *F*) Surface representation of the yTAF1/7 complex in front and top view. The location of the two pockets and arms and saddle position are outlined in top view. (*G*) Surface rendition of the yTAF1/7 complex, colored according to the electrostatic potential, from red (−10 k_BT/e) to blue (+10 k_BT/e). (*H*) Surface conservation score mapped on the complex based on multiple sequence alignment.

(TAF1/7) assembly and the molecular mechanism of transcriptional checkpoint control, we set out to obtain a structure of the TAF1/7 heterodimeric complex. Here, we report the structure of the yTAF1/7 complex at 2.9 Å resolution. Surprisingly, we find that yTAF1 displays no resemblance to known HATs. However, the complex does contain a prominent surface pocket that specifically recognizes an inhibitory chromatin mark, trimethylated histone H3 lysine 27 (H3K27me3), in a manner that is regulated by phosphorylation at neighboring H3 Serine28. The unexpected capacity of TAF1/7 to recognize H3K27me3 could point toward novel modes of regulation of TFIID activity in PIC formation.

Results

To map TAF1/7 interactions, we performed limited proteolysis on purified endogenous yeast TFIID with yTAF1 tagged at the C terminus (*Materials and Methods* and Fig. S1) and analyzed resulting fragments by liquid chromatograph mass spectrometry (LC-MSD SL model). Repeated cycles of proteolysis and mass spectrometry allowed us to define optimal yTAF1 (aa 455–960) and yTAF7 (aa 85–320) interaction domains suitable for crystallization trials (Fig. 1 *A* and *B* and Fig. S2 *A* and *B*). Both fragments appear conserved with equivalent regions of human, drosophila, zebrafish, and mouse TAFs (Figs. S3 and S4). The structure of the complex was determined by single anomalous dispersion using selenomethionine protein and refined to a 2.9 Å resolution (Table 1). All structural elements and loops were well resolved, except for part of the longest loop (aa 807–860) at the back of the complex, which is only partially resolved in the electron density map.

The heterodimer of yTAF1 and yTAF7 displays unusual saddle-like organization with dimensions of ~78 Å × 98 Å × 102 Å (Fig. 1 *C* and *D*). There are two notable flanking arms (arm 1 and arm 2) spaced by ~78 Å, with a cradle between the arms and a heterodimerization surface at the base of the complex (Fig. 1 *E* and *F*). The surface is mostly hydrophobic but with two prominent electronegative pockets exposed to solvent at the base of the cradle (Fig. 1*G*). Projection of amino acid conservation scores onto the complex surface reveals that the arms, cradle, and pockets display high cross-species identity, suggestive of conserved function (Fig. 1*H* and Figs. S3 and S4).

The yTAF1/7 heterodimerization surface is remarkably large and is comprised of several components (Fig. 2 and Fig. S5). The TAF1/7 interface is largely hydrophobic with polar residues at the peripheral interaction region (Fig. 2 *A* and *B*). In total, heterodimerization buries ~6,800 $Å^2$ of solvent-accessible surface and contains a total of 43 salt bridges and numerous hydrogen bonds between the two TAFs, as documented by PISA analysis (17). The main hydrophobic portion of the heterodimerization surface is composed of multiple β strands, which form an interlocking triple barrel (Fig. 2*C* and Fig. S5). Two barrels are, respectively, derived mostly from yTAF1 or yTAF7 residues and related by pseudo-twofold symmetry along a central barrel (Fig. 2*C*). N-terminal loops of yTAF1 (aa 462–501)

Table 1. Data collection and refinement statistics

Data collection	
Space group	$P2_12_12_1$
Cell dimensions	
a, *b*, *c*, Å	122.22, 123.66, 229.61
α, β, γ , °	90.0, 90.0, 90.0
Resolution, Å	2.9 (3.05–2.89)*
R_{merge}	11.8 (77.7)
$I/\sigma I$	11.1 (2.0)
Completeness, %	99.9 (99.7)
Redundancy	7.6 (5.0)
Refinement	
Resolution, Å	83.6–2.9
No. reflections	77,943
R_{work}/R_{free}	20.8/24.3
No. atoms	
Protein	16,456
Ligand/ion	2
Water	13
B factors	
Protein	80.30
Ligand/ion	106.7
Water	56
rmsds	
Bond lengths, Å	0.004
Bond angles, °	0.806
Ramachandran	
Preferred/allowed/disallowed, %	95.61/4.25/0.14

*Highest resolution shell is shown in parentheses.

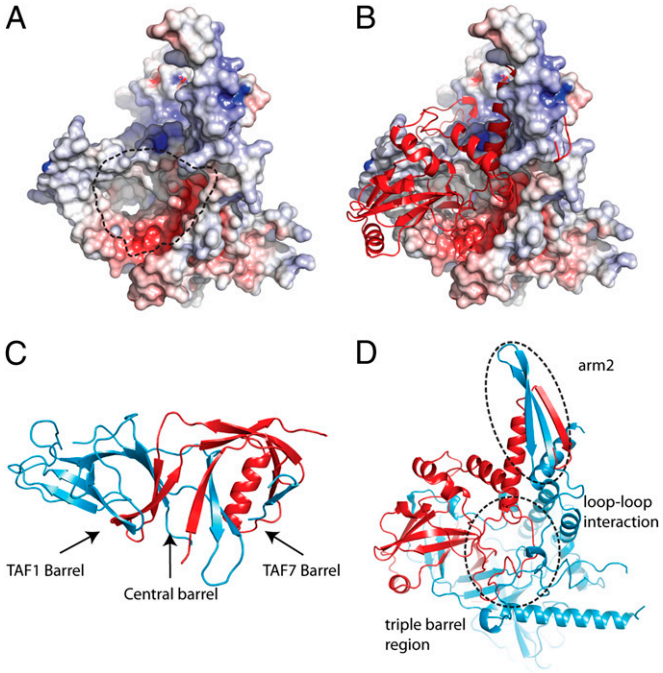

Fig. 2. Organization of the yTAF1/7 heterodimerization surface. (*A*) Electrostatic potential distribution of the TAF1 region involved in the interaction with TAF7 (circled). (*B*) Ribbon representation of TAF7 (red) within the complex interaction surface of TAF1 showing the loop–loop interaction region of the yTAF1 and yTAF7 subunit is dominated by electrostatic interaction, whereas triple barrel regions are dominated by hydrophobic interaction. (*C*) Ribbon representation of the triple barrel region formed by yTAF1(blue) and yTAF7 (red). (*D*) The loop–loop interaction of two TAF subunits adjacent to the triple barrel region and the interaction of β strands bracketed by α-helices between two subunits at the arm 2 region (circled).

and yTAF7 (aa 91–113) wrap the yTAF1 barrel, extending the mutual fold in this region of the complex. Following the triple barrel motif, yTAF1 diverges from yTAF7 (Fig. S5). The yTAF1 β strands are followed by a long loop that connects the triple barrel to arm 1, which resembles a winged helix–turn–helix (wHTH) motif, the cradle region, and the entire back surface of the complex. All of these features are derived entirely from yTAF1 residues. yTAF1 then engages in additional heterodimer contacts with yTAF7 within arm 2 and in a loop–loop interaction region at the base of arm 2, close to the TAF7 barrel (Fig. 2*D*). This extensive mutual folding pattern may explain why we found that the yTAF1 fragment could not be expressed in isolation in *Escherichia coli* and could only be coexpressed with yTAF7.

To look for structural homologs of the TAF1/7 complex, we performed a DALI search (18). The analysis revealed partial similarities to motifs in RNA Pol I and III-associated proteins (19, 20) and RNA Pol II-associated factor TFIIF (Rap30/Rap74) (21), which all contain triple barrels. In the latter case, TAF7 β strands show close overlap with β strands of the TFIIF Rap30 subunit (21) (Z score of 5.6 and rmsd in 84 Cα atom positions of 2.9 Å), but the disposition of TAF1 β strands bears little or no similarity to that of the Rap74 strands of the TFIIF dimerization module (Fig. S6). The yTAF1 arm 1 also shows close similarity to the Rap30 DNA binding wHTH domain (22) (Z score of 4.3 and rmsd in 65 Cα atom positions of 2.8 Å) (Fig. S6). Most importantly, the DALI search revealed no obvious structural similarities between yTAF1 and any known HATs. It has previously been proposed that TAF1 harbors putative Gly–X–Gly motifs, which mediate acetyl Co-A substrate recognition that is required for HAT activity (9, 23). Visual inspection reveals, however, that these motifs are actually located at different positions in yeast (in the loop connecting strands β7 and β8 in the triple barrel region) and human (near the end of helix α9 in the cradle region) TAF1 sequences (Fig. S3). Although the putative yeast Gly–X–Gly motif is completely surface exposed, repeated attempts to demonstrate acetyl Co-A binding to the yTAF1/7 complex failed (Fig. S7). Together, these facts led us to suggest that yTAF1 may not be a HAT.

We next analyzed the two highly conserved electronegative pockets on the yTAF1/7 surface (Fig. 3*A*). Both pockets are large (pocket 1, ~3,600 $Å^3$ and pocket 2, ~3,100 $Å^3$) and separated only by the yTAF1 N-terminal loop (aa 486–502). Whereas pocket 1 is tunnel shaped, is only partially accessible to the surface, and runs deep within the interior of the complex (Fig. S8), pocket 2 is completely solvent exposed and sits on the top of the TAF1 barrel (Fig. 3*B*). The pocket 2 base is electronegative (Fig. 3*C*) and composed of residues from the yTAF7 N-terminal loop (aa 111–114) and the β6 strand. The wall of pocket 2 is predominantly hydrophobic (Fig. 3*D*) and derived mostly from the residues of the yTAF1 N terminus loop region (aa 478–502) that spans the TAF1 barrel and loops connecting the yTAF1 β strands (β1–β2, β2–β3, and β7–β8).

Interestingly, yTAF1Gly561, a conserved residue that is subject to a temperature-sensitive (ts) mutation (Gly561Asp) that interferes with cell cycle progression in cultured eukaryotic cells (23, 24), is located next to aromatic residues that form the pocket 2 wall (Fig. S9*A*). Inspection of potential influences of Gly to Asp substitution reveals steric clashes with neighboring residues of both subunits that could destabilize the complex (Fig. S9*B*). Indeed, attempts to purify the stable yTAF1/7 complex with the ts mutation were unsuccessful. Consideration of the location of the ts mutation in the yTAF1/7 complex raises the possibility that pocket 2 may have key roles in cell cycle and gene regulation.

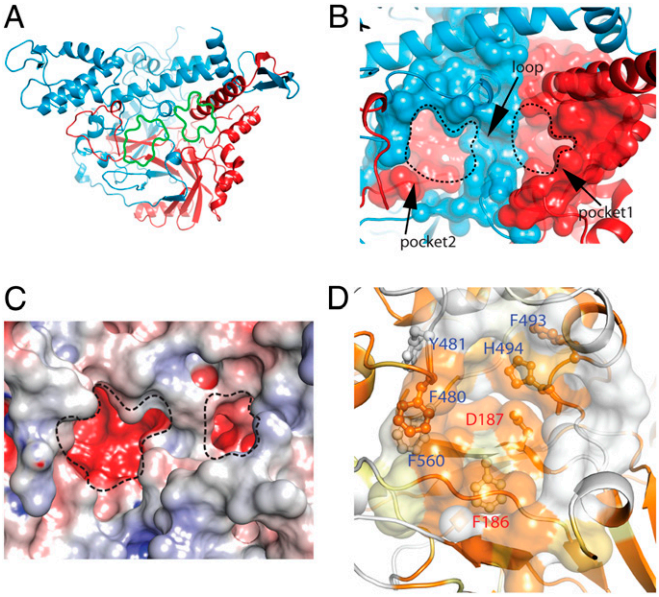

Fig. 3. Surface pockets of the yTAF1/7 complex. (*A*) Location of the two pockets (green lines) at the base of the cradle region shown on the ribbon representation of the complex. (*B*) Surface representation of the pockets, showing that the two pockets are separated by the N-terminal loop region of yTAF1. Note that the base of pocket 2 is formed by TAF7 (red), whereas the N-terminal loop region of yTAF1 (blue) mainly forms the wall of the pocket. (*C*) Blow-up view of two pockets showing the highly negative charge distribution at the base of the pockets from red (−10 k_BT/e) to blue (+10 k_BT/e). (*D*) Surface conservation and composition of pocket 2. Aromatic residues derived from yTAF1 (455–960) and yTAF7 (91–320) subunits are labeled in blue and red text, respectively.

To better understand functions of yTAF1/7 surface pockets, we performed fitting of the yTAF1/7 complex structure within available cryo-electron microscopic (EM) maps of the TFIID complex (25). The yTAF1/7 complex displayed excellent fit within lobe A of the EM map with a correlation coefficient of 0.93. The orientation of yTAF1/7 subunits within the EM map shows that the arm 1 wHTH motif is located at the extreme edge of TFIID lobe A, whereas surface pockets face outwards (Fig. 4 *A* and *B*). This raises the possibility that TAF1/7 surface pockets lie in close proximity to chromatin.

To test the idea that TAF1/7 might display specific chromatin interactions, we screened yTAF1/7 interactions with N-terminal tails of histones H2A, H2B, H3, and H4 containing various post-translational modifications in a histone peptide array. Based on the preliminary observations, we performed biolayer interferometry assays with biotinylated yTAF1/7 immobilized on a biosensor tip and H3 peptides as analytes. The study revealed that yTAF1/7 binds selectively to the H3-K27me3 (19–35) peptide with an apparent K_d value of 3.3 μM (Fig. 4*C* and Table S1). The complex fails to bind unmodified H3 tails or equivalent peptides bearing either a H3K27me3S28ph double mark or a single H3S28ph mark. This suggests that H3K27me3 and H3S28ph marks could, together, constitute a methyl-phosphorylation (methyl/phos) switch, which regulates yTAF1/7 association with chromatin.

Preliminary assessments suggested that the yTAF1/7 pocket 2 is a likely candidate for recognition of the H3K27me3 repressive mark; the pocket 2 wall aromatic residues could create a hydrophobic aromatic cage, favorable for accommodation of the hydrophobic trimethylated lysine side chain (26), whereas the negatively charged pocket floor could neutralize positive charges on the H3K27me3 peptide. We tested this possibility using mutational analysis combined with peptide affinity studies. Substitution of evolutionary conserved yTAF1 Phe493, His494, Phe480 residues, and yTAF7 Phe186 with Ala (Fig. 3*D*) permitted stable yTAF1/7 formation but completely abolished binding to H3K27me3 (19–35) peptide (Table S2). Other targeted mutations disrupted stable yTAF1/7 complex formation (Table S2). Thus, pocket 2 aromatic residues are absolutely essential for recognition of the H3K27me3 mark.

Discussion

TAF7 interactions with TAF1 are highly important for regulation of TFIID function; TAF7 displays key transcriptional checkpoint control activities that help to ensure correct PIC assembly before transcriptional initiation (15). It is currently proposed that TAF7 could achieve this effect by acting as a dissociable inhibitor of an intrinsic TAF1 HAT activity. Our new structure of the TAF1/7 complex, reported here, suggests that this model could require re-evaluation. We find that the TAF1/7 complex is characterized by a large, complex, and intertwined hydrophobic heterodimer interface. This finding strongly suggests that TAF1 is only likely to be stable in the presence of a suitable binding partner, such as TAF7, and that TAF7 dissociation would disrupt TAF1 folding and functional viability. Moreover, analysis of the structural features of the complex do not favor the concept that yTAF1, itself, is a HAT.

Interestingly, our study does reveal, for the first time to our knowledge, that the yTAF1/7 complex can bind specifically to a transcriptionally repressive H3K27me3 mark, which is present in both mammalian and yeast cells (27). This raises the surprising possibility that TAF1/7 interactions with H3K27me3 could play important and hitherto unsuspected roles in TFIID function and regulation of PIC assembly. Although assessment of the potential roles of this interaction in cultured cells and experimental organisms is beyond the scope of this crystallographic study, two lines of evidence suggest that pocket 2 interactions with H3K27me3 are likely functionally important. First, residues that comprise pocket 2, which mediates H3K27me3 recognition, are well conserved in other species. Indeed, we expressed and purified equivalent domains of human TAF1 and TAF7 as a stable complex and found that human TAF1/7 also binds selectively to H3K27me3 with a K_d of 10 μM (Fig. S10) and this interaction is also dependent upon the putative H3 K27methyl/S28phos switch. Second, the evolutionarily conserved TAF1 residue that is subject to ts mutation (Gly561Asp in yeast and G716Asp in human) and is associated with defects in cell cycle progression and gene expression is located near the rear edge of pocket 2. This further signifies the likely functional importance of this region of TAF1 (24). Trimethylation on H3K27 is written and recognized by polycomb repressive complexes (PRCs) in higher eukaryotes, raising the question of why this mark is present in yeast (27), which lacks PRC. A hitherto unknown role for the TFIID complex in K27me3 recognition could help to explain this apparent paradox.

What are the possible roles of the TAF1/7 interaction with K27me3? We propose two ideas. Based on our findings and studies published to date regarding the role of TAF7 and TAF1 in transient transcriptional repression required for effective PIC assembly (28), we suggest that the TAF1/7 complex could read the repressive H3K27me3 mark to help induce this temporary transcriptional arrest and that subsequent H3S28 phosphorylation could reverse this effect to induce transcription initiation. It will also be interesting to see if TAF1/7 plays a role in recognition of bivalent chromatin in embryonic stem cells, which contains both the repressive H3K27me3 mark and an active H3K4me3 mark, and in cross-talk with the TAF3–PHD domain that recognizes the H3K4me3 mark on genes that are marked with bivalent chromatin (6, 29–31). Further experiments would be required to test these, and other, possibilities, but our findings will provide a strong structural basis for future studies of the possible regulatory interplay between the TAF1/7 complex and H3K27me3/H3S28ph switch in transcription initiation.

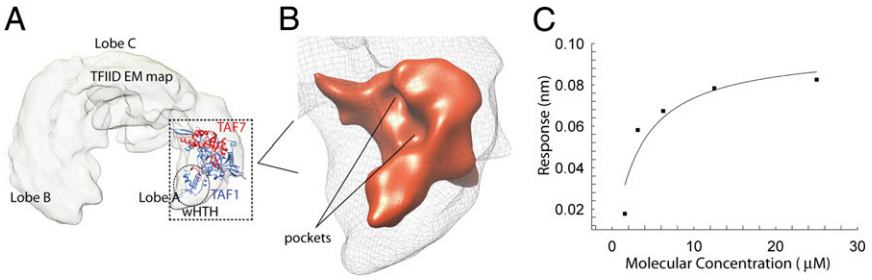

Fig. 4. Location of the yTAF1/7 complex structure suggests a role in chromatin recognition. (*A*) Fitted yTAF1/7 complex within lobe A of the EM map (emdb-5026) (25) of TFIID with the wHTH domain pointing outward and triple barrel positioned upward. (*B*) Blow-up view of the fitted yTAF1/7 complex within lobe A of the EM map. The resolution of the complex was set to 10 Å to show the outward-facing location of TAF1/7 pockets 1 and 2. (*C*) H3K27me3 (19–34) binding to the yTAF1/7 complex. The plot shows the response versus peptide concentration derived from the binding data. The H3K27me3 peptide binds to the complex with 3.3 μM affinity, with Rmax = 0.0975 and R^2 = 0.8792.

Materials and Methods

Limited Proteolysis and Mass Spectrometry. TAF1/7 interacting regions were identified by limited proteolysis studies with various enzymes (trypsin, chymotrypsin, papain) on endogenous yeast TFIID purified via tandem affinity tag (protA/calmodulin binding peptide) at the C terminus of the TAF1 protein and immobilized on the calmodulin beads. The partial digestion time course was analyzed by SDS/PAGE and mass spectrometry on an Agilent LC-MSD trap SL mass spectrometer to reveal protected fragments corresponding to the entire C terminus of TAF1 from amino acids ~430–1070 and N-terminal domain of TAF7 from amino acids ~80–320. We synthesized the respective TAF subunit gene fragments optimized for the *E. coli* expression host and cloned the synthetic genes into the pET-duet vector (Novagen) for coexpression of these two proteins with His affinity tag.

Protein Expression and Purification. *E. coli* BL21 cells were amplified, and TAF1 and TAF7 expression was initiated by isopropyl β-D-1-thiogalactopyranoside (IPTG) induction at 30 °C and purified over Talon beads by virtue of the associated TAF7 His tag. The protein complex was further purified over a Superdex 200 Gel filtration column (GE Healthcare Life Sciences). The resulting complex shows only one peak in the S200 column with an equimolar ratio of TAF1 and TAF7 in the SDS/PAGE gel (Fig. S2*B*). The purified protein was concentrated to ~12 mg/mL for crystallization. Subsequent crystallization trials were unsuccessful. Further limited proteolysis with trypsin, elastase, Gluc, and chymotrypsin combined with LC-MS analysis defined fragments of TAF1 (aa 455–960) and TAF7 (aa 85–320) that were proteolytically resistant, and coexpression of this complex revealed that it was well expressed without truncation products. We found that the complex produced showers of fiber-like crystals on a crystallization trial with the addition of thermolysin and subtilisin (in dilution of 1:1,000) to the protein complex in the crystallization drop. Another round of LC-MS study revealed a truncation at the low-complexity sequence region of TAF1 at aa ~830–850. We therefore generated different-sized loop deletion fragments of TAF1 in complex with TAF7 (aa 85–320). Finally deletion of amino acids from 836 to 855 of the TAF1 fragment in complex with TAF7 (aa 85–320) produced 3D crystals for X-ray diffraction study. Se-Met substituted protein complex was expressed in the BL21(834) cell with 100 μg/mL SeMet per liter of culture. SeMet incorporation was at ~90% efficiency and was confirmed by mass spectrometry.

Crystallization and Data Collection. Crystals of the yTAF1/7 and seMet substituted complex were grown at room temperature by vapor diffusion method by using an equal volume of protein-to-well solution over a reservoir solution containing 150 mM ammonium citrate and 20% (wt/vol) PEG 3350. The crystals were cryoprotected by brief soaking in mother liquor containing 12% (vol/vol) ethylene glycol and flash cooled in liquid nitrogen. Single-wavelength anomalous dispersion (SAD) data were collected with Se-Met crystals at a resolution of 2.9 Å at The Northeastern Collaborative Access Team (NECAT) beamline 24 ID C at advanced photon source (APS). The diffraction data were processed and scaled with Package HKL2000 (32). The yTAF1/7 crystallizes in space group $P2_12_12_1$ with typical unit cell parameters of $a = 122.2$ Å, $b = 123.6$ Å, $c = 229.6$ Å, $\alpha = 90.0°$, $\beta = 90.0°$, and $\gamma = 90.0°$. Each asymmetric unit contained three heterodimers of the yTAF1/7 complex.

Structure Determination and Refinement. The SAD dataset was collected to 2.9 Å at a wavelength of 0.97918 Å. The selenium sites in the asymmetric unit were initially located and used to generate preliminary phases calculated to 4.0 Å by using the program PHENIX (33). The starting model was then built manually with the help of Coot (34) and refined by PHENIX. Initially all of the three heterodimers were individually traced through density, with Coot and noncrystallographic symmetry (NCS) averaging applied for the initial stages of refinement. Subsequent stages of refinement did not include any NCS, and we were able to extend the phase to the limit of the data (2.9 Å). Data collection and refinement statistics are summarized in Table 1. Figures were generated by using PYMOL software (www.pymol.org). The final model, which includes three copies of yTAF1/7, was analyzed using PHENIX and PROCHECK (35). The individual copies superimpose with an average rmsd of 1.5 Å for common Cα positions.

Fitting of the yTAF1/7 Complex Within the Cryo EM Map of TFIID. The fitting of the yTAF1/7 complex within the cryo EM map of TFIID (emdb5026) (25) was performed with the Chimera (36) UCSF software. The EM map was segmented using Segger segmentation method (37) into four different regions. Lobe A was chosen to fit the yTAF1/7 molecule using the Chimera software. During the fitting process, a density map for the complex structure was generated. A cross-correlation score for the fit was calculated based on the generated density map. The structure was adjusted within the segmented region of lobe A for a better cross-correlation score. The density map for the structure was generated at the same resolution of the EM map of TFIID at ~20 Å. The final cross-correlation score for the fit is 0.93.

Peptide Array. We used a modified peptide array from Active Motif (cat. no. 13001) to screen for TAF1/7 interactions. The array was blocked with the buffer provided with the kit, washed three times, and the membrane was incubated with purified 5 μM His-tagged protein complex overnight at 4 °C. The membrane was washed three times with the wash buffer provided and incubated with anti-his antibody at a 1:3,000 dilution for 1 h at room temperature. The unbound antibody was washed three times, and the membrane was again incubated with horseradish peroxidase conjugated antibody (at 1:2,500) in blocking buffer for 1 h at room temperature. Finally, the membrane was submerged in the ECL developing solution provided with the kit, and the image was captured on X-ray film. The exposure time was 5–10 s.

Biolayer Interferometry Peptide Binding Assay. Peptide binding to the native or mutated yTAF1/7 protein complexes was measured by biolayer interferometry on an Octet-RED 384 instrument (Pall ForteBio). All of the binding studies were performed in binding buffer 20 mM Tris pH 7.8, 175 mM NaCl, 0.1 mg/mL BSA, 0.05% Tween20 using streptavidin biosensors. Samples or buffer were dispensed into 96-well microtiter plates at a volume of 200 μL per well. Operating temperature was maintained at 27 °C with a 1,000 rpm rotary agitation. SA Streptavidin biosensor tips (Pall FortéBio) were prewet for 10 min with buffer to establish a baseline before protein immobilization. Biotinylated yTAF1/7 proteins (molecular weight, ~90 KDa) were immobilized onto the biosensors for 15 min at a concentration of 20 μg/mL. The immobilization level attained was 3–4 nm. Binding association of the peptides with biosensor tips was monitored for 3 min, and subsequent disassociation in buffer was monitored for 2 min. Histone peptide tails (mw, ~1.5 KDa) were tested at concentrations of 1.56, 3.12, 6.25, 12.5, 25, 50, and 100 μM. The apparent affinities of peptides and yTAF1/7 were calculated from equilibrium measurements and, when appropriate, global fits of the kon and koff values, yielding similar values.

Isothermal Titration Calorimetry. Calorimetric titrations were carried out at room temperature with an isothermal titration calorimetry (ITC) instrument from Microcal. Purified yTAF1/7 complex protein was maintained in 25 mM Tris·HCl (pH 7.8), 150 mM NaCl, 1 mM EDTA, and 1 mM *tris*(2-carboxyethyl)phosphine buffer for the ITC binding experiment. Acetyl Co-A was dissolved in the same buffer. Data analysis was done with Microcal Origin 5.0 software.

ACKNOWLEDGMENTS. We thank Senapathy Rajagopalan and Stephen Ayers for comments. This study was supported by National Institutes of Health Grants DK41482 (to P.W.) UO1 GM094614 (to R.J.F.; with a subcomponent to P.W.), and Cancer Prevention Research Institute of Texas Grant RP011444-P3 (to P.W.). J.-Å.G. is grateful for a grant from the Robert A. Welch Foundation (E-0004). This work is based on research conducted at the Advanced Photon Source on the Northeastern Collaborative Access Team beamlines, which are supported by National Institute of General Medical Sciences Grant P41 GM103403 from the National Institutes of Health.

1. Bieniossek C, et al. (2013) The architecture of human general transcription factor TFIID core complex. *Nature* 493(7434):699–702.
2. Grant PA, et al. (1998) A subset of TAF(II)s are integral components of the SAGA complex required for nucleosome acetylation and transcriptional stimulation. *Cell* 94(1):45–53.
3. Leurent C, et al. (2002) Mapping histone fold TAFs within yeast TFIID. *EMBO J* 21(13):3424–3433.
4. Cianfrocco MA, et al. (2013) Human TFIID binds to core promoter DNA in a reorganized structural state. *Cell* 152(1-2):120–131.
5. Jacobson RH, Ladurner AG, King DS, Tjian R (2000) Structure and function of a human TAFII250 double bromodomain module. *Science* 288(5470):1422–1425.
6. van Ingen H, et al. (2008) Structural insight into the recognition of the H3K4me3 mark by the TFIID subunit TAF3. *Structure* 16(8):1245–1256.
7. Tavassoli P, et al. (2010) TAF1 differentially enhances androgen receptor transcriptional activity via its N-terminal kinase and ubiquitin-activating and -conjugating domains. *Mol Endocrinol* 24(4):696–708.
8. Ribeiro JR, Lovasco LA, Vanderhyden BC, Freiman RN (2014) Targeting TBP-associated factors in ovarian cancer. *Front Oncol* 4. Available at www.ncbi.nlm.nih.gov/pmc/articles/PMC3949196/. Accessed April 10, 2014.
9. Mizzen CA, et al. (1996) The TAF(II)250 subunit of TFIID has histone acetyltransferase activity. *Cell* 87(7):1261–1270.

10. Dikstein R, Ruppert S, Tjian R (1996) TAFII250 is a bipartite protein kinase that phosphorylates the base transcription factor RAP74. *Cell* 84(5):781–790.
11. Leurent C, et al. (2004) Mapping key functional sites within yeast TFIID. *EMBO J* 23(4): 719–727.
12. Irvin JD, Pugh BF (2006) Genome-wide transcriptional dependence on TAF1 functional domains. *J Biol Chem* 281(10):6404–6412.
13. Lee TI, et al. (2000) Redundant roles for the TFIID and SAGA complexes in global transcription. *Nature* 405(6787):701–704.
14. Gegonne A, Weissman JD, Singer DS (2001) TAFII55 binding to TAFII250 inhibits its acetyltransferase activity. *Proc Natl Acad Sci USA* 98(22):12432–12437.
15. Gegonne A, Weissman JD, Zhou M, Brady JN, Singer DS (2006) TAF7: A possible transcription initiation check-point regulator. *Proc Natl Acad Sci USA* 103(3):602–607.
16. Liu D, et al. (1998) Solution structure of a TBP-TAF(II)230 complex: Protein mimicry of the minor groove surface of the TATA box unwound by TBP. *Cell* 94(5):573–583.
17. Krissinel E, Henrick K (2007) Inference of macromolecular assemblies from crystalline state. *J Mol Biol* 372(3):774–797.
18. Holm L, Sander C (1993) Protein structure comparison by alignment of distance matrices. *J Mol Biol* 233(1):123–138.
19. Geiger SR, et al. (2010) RNA polymerase I contains a TFIIF-related DNA-binding subcomplex. *Mol Cell* 39(4):583–594.
20. Taylor NMI, Baudin F, von Scheven G, Müller CW (2013) RNA polymerase III-specific general transcription factor IIIC contains a heterodimer resembling TFIIF Rap30/Rap74. *Nucleic Acids Res* 41(19):9183–9196.
21. Gaiser F, Tan S, Richmond TJ (2000) Novel dimerization fold of RAP30/RAP74 in human TFIIF at 1.7 A resolution. *J Mol Biol* 302(5):1119–1127.
22. Groft CM, Uljon SN, Wang R, Werner MH (1998) Structural homology between the Rap30 DNA-binding domain and linker histone H5: Implications for preinitiation complex assembly. *Proc Natl Acad Sci USA* 95(16):9117–9122.
23. Dunphy EL, Johnson T, Auerbach SS, Wang EH (2000) Requirement for TAF(II)250 acetyltransferase activity in cell cycle progression. *Mol Cell Biol* 20(4):1134–1139.
24. Hayashida T, et al. (1994) The CCG1/TAFII250 gene is mutated in thermosensitive G1 mutants of the BHK21 cell line derived from golden hamster. *Gene* 141(2): 267–270.
25. Papai G, et al. (2009) Mapping the initiator binding Taf2 subunit in the structure of hydrated yeast TFIID. *Structure* 17(3):363–373.
26. Yun M, Wu J, Workman JL, Li B (2011) Readers of histone modifications. *Cell Res* 21(4): 564–578.
27. Garcia BA, et al. (2007) Organismal differences in post-translational modifications in histones H3 and H4. *J Biol Chem* 282(10):7641–7655.
28. Pennington KL, Marr SK, Chirn G-W, Marr MT, 2nd (2013) Holo-TFIID controls the magnitude of a transcription burst and fine-tuning of transcription. *Proc Natl Acad Sci USA* 110(19):7678–7683.
29. Bernstein BE, et al. (2006) A bivalent chromatin structure marks key developmental genes in embryonic stem cells. *Cell* 125(2):315–326.
30. Mikkelsen TS, et al. (2007) Genome-wide maps of chromatin state in pluripotent and lineage-committed cells. *Nature* 448(7153):553–560.
31. Liu Z, Scannell DR, Eisen MB, Tjian R (2011) Control of embryonic stem cell lineage commitment by core promoter factor, TAF3. *Cell* 146(5):720–731.
32. Otwinowski Z, Minor W (1997) Processing of x-ray diffraction data collected by oscillation methods. *Methods Enzymol* 276:307–326.
33. Adams PD, et al. (2010) *PHENIX*: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D Biol Crystallogr* 66(Pt 2):213–221.
34. Emsley P, Cowtan K (2004) Coot: Model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60(Pt 12 Pt 1):2126–2132.
35. Laskowski RA, MacArthur MW, Moss DS, Thornton JM (1993) PROCHECK: A program to check the stereochemical quality of protein structures. *J Appl Cryst* 26(Part 2): 283–291.
36. Pettersen EF, et al. (2004) UCSF Chimera—A visualization system for exploratory research and analysis. *J Comput Chem* 25(13):1605–1612.
37. Pintilie GD, Zhang J, Goddard TD, Chiu W, Gossard DC (2010) Quantitative analysis of cryo-EM density map segmentation by watershed and scale-space filtering, and fitting of structures by alignment to regions. *J Struct Biol* 170(3):427–438.