Architecture of TAF11/TAF13/TBP complex suggests novel regulatory state in General Transcription Factor TFIID function

Kapil Gupta[1,2], Aleksandra A. Watson[3], Tiago Baptista[4-7], Elisabeth Scheer[4-7], Anna L. Chambers[1], Christine Koehler[8], Juan Zou[9], Ima Obong-Ebong[10], Eaazhisai Kandiah[2,11], Arturo Templador[2], Adam Round[2], Eric Forest[11], Petr Man[12], Christoph Bieniossek[2], Ernest D. Laue[3], Edward A. Lemke[8], Juri Rappsilber[9], Carol V. Robinson[10], Didier Devys[4-7], Laszlo Tora[4-7*] and Imre Berger[1*]

[1] The School of Biochemistry, Biomedical Sciences and BrisSynBio Centre, University of Bristol, Tankard's Close, Bristol BS8 1TD, United Kingdom

[2] European Molecular Biology Laboratory, 71 Avenue des Martyrs, 38000 Grenoble, France

[3] Department of Biochemistry, University of Cambridge, 80 Tennis Court Road, Cambridge CB2 1GA, United Kingdom

[4] Institut de Génétique et de Biologie Moléculaire et Cellulaire IGBMC, Illkirch, France

[5] Centre National de la Recherche Scientifique, UMR7104, Illkirch, France

[6] Institut National de la Santé et de la Recherche Médicale, U964, Illkirch, France

[7] Université de Strasbourg, Illkirch, France

[8] European Molecular Biology Laboratory, Meyerhofstrasse 1, 69117 Heidelberg, Germany

[9] Wellcome Trust Centre for Cell Biology, University of Edinburgh, Max Born Crescent, Edinburgh, EH9 3BF, United Kingdom and Chair of Bioanalytics, Institute of Biotechnology, Technische Universität Berlin, 13355 Berlin, Germany

[10] Physical and Theoretical Chemistry Laboratory, South Parks Road, Oxford OX1 3QZ, United Kingdom

[11] Institut de Biologie Structurale IBS, 71 Avenue des Martyrs, 38042 Grenoble, France

[12] BioCeV - Institute of Microbiology, The Czech Academy of Sciences, Prumyslova 595, 252 50 Vestec and Faculty of Science, Charles University, Hlavova 8, 128 43 Prague, Czech Republic

[#] Corresponding authors: Laszlo Tora (laszlo.tora@igbmc.fr) & Imre Berger (imre.berger@bristol.ac.uk)

ABSTRACT

General transcription factor TFIID is a key component of RNA polymerase II transcription initiation. Human TFIID is a megadalton-sized complex comprising TATA-binding protein (TBP) and 13 TBP-associated factors (TAFs). TBP binds to core promoter DNA, recognizing the TATA-box. We identified a ternary complex formed by TBP and the histone fold (HF) domain-containing TFIID subunits TAF11 and TAF13. We demonstrate that TAF11/TAF13 competes for TBP binding with TATA-box DNA, and also with the N-terminal domain of TAF1 previously implicated in TATA-box mimicry. In an integrative approach combining crystal coordinates, biochemical analyses and data from cross-linking mass-spectrometry (CLMS), we determine the architecture of the TAF11/TAF13/TBP complex, revealing TAF11/TAF13 interaction with the DNA binding surface of TBP. We identify a highly conserved C-terminal TBP-binding domain (CTID) in TAF13 which is essential for supporting cell growth. Our results thus have implications for cellular TFIID assembly and suggest a novel regulatory state for TFIID function.

INTRODUCTION

Eukaryotic gene expression is a highly regulated process which is controlled by a plethora of proteins, arranged in multiprotein complexes including the general transcription factors (GTFs), Mediator and RNA polymerase II (Pol II) (Gupta et al. 2016; Thomas and Chiang 2006). Regulated class II gene transcription is initiated by sequential nucleation of GTFs and Mediator on core promoter DNA (Rhee and Pugh 2012). The GTF TFIID is a cornerstone in this process and links cellular signaling events with regulatory DNA elements and the components of the transcription machinery (Albright and Tjian 2000). A basal transcription system which supports initiation can be reconstituted with TBP and the GTFs TFIIA, TFIIB, TFIIE, TFIIF and TFIIH *in vitro*, however, TFIID is required to respond to activators (Hampsey and Reinberg 1999). In mammalian cells, the promoters of virtually all protein-encoding genes are occupied by TFIID, and loss of TFIID components causes embryonic lethality (Gegonne et al. 2012; Kim et al. 2005; Mohan et al. 2003). TFIID subunits are thought to mediate cross-talk with epigenetic modifications on nucleosomes and regulatory DNA elements in promoter regions (Vermeulen et al. 2007; Verrijzer et al. 1995). X-ray crystallography revealed many details of TFIID components at near atomic resolution (Gupta et al. 2016). Cryo-electron microscopy (cryo-EM) provided essential insight into TFIID architecture and promoter DNA interaction (Bieniossek et al. 2013; Cianfrocco et al. 2013; Louder et al. 2016). The recent identification of a discrete TAF-containing complex in the cytoplasm of cells provided first insight into holo-TFIID assembly from preformed sub-modules, regulated by nuclear import mechanisms (Trowitzsch et al. 2015).

Canonical human TFIID comprises TATA-binding protein (TBP) and 13 TBP-associated factors (TAFs) (Matangkasombut et al. 2004, Muller and Tora 2014). Furthermore, non-canonical TFIID and TAF-containing complexes have been identified regulating spermatogenesis and stem cell development (Goodrich and Tjian 2010; Maston et

al. 2012; Muller et al. 2010). A nuclear core-TFIID complex was identified made up of two copies each of TAF4, 5, 6, 9 and 12 (Bieniossek et al. 2013; Wright et al. 2006). Biochemical and structural studies established the histone-fold domain (HFD) as a key TAF-TAF interaction motif within TFIID (Gangloff et al. 2001a). TAF3, 4, 6, 8, 9, 10, 11, 12 and 13 contain HFDs and assemble specifically into heterodimers (TAF3-10, TAF4-12, TAF6-9, TAF8-10 and TAF11-13) (Birck et al. 1998; Gangloff et al. 2001b; Werten et al. 2002; Xie et al. 1996).

TFIID recognizes core promoter DNA via its TATA-box binding protein subunit, TBP. TBP is central to transcription regulation in eukaryotes and is the only subunit present in the transcription initiation complexes of each of the three RNA polymerases (Koster et al. 2015; Thomas and Chiang 2006; Tora and Timmers 2010). TBP consists of a highly variable N-terminal domain with less well understood function and a conserved DNA-binding C-terminal core domain comprising two symmetrical pseudo-repeats (Thomas and Chiang 2006; Tora and Timmers 2010). Crystal structures of the conserved TBP core revealed a saddle-like shape with a concave DNA binding surface recognizing the minor groove of TATA-box containing DNA (Kim et al. 1993; Nikolov et al. 1996; Nikolov et al. 1992).

The DNA-binding activity of TBP/TFIID is tightly regulated by gene-specific co-factors that can activate or inhibit transcription (Koster et al. 2015; Tora and Timmers 2010). The mechanism of a number of these regulatory factors has been described in molecular detail. The TFIID component TAF1 was found to associate with the concave DNA-binding surface of TBP via its N-terminal domain (TAF1-TAND), exhibiting TATA-box mimicry (Anandapadamanaban et al. 2013). TAF1-TAND, unstructured in isolation, was found to adopt a three-dimensional structure closely resembling the TATA-element is shape and charge distribution when bound to TBP (Burley and Roeder 1998; Liu et al. 1998). This interaction is conserved in yeast, *Drosophila* and human (Anandapadamanaban et al. 2013;

Burley and Roeder 1998; Liu et al. 1998; Mal et al. 2004). The recent high-resolution structure of TBP bound to yeast TAF1-TAND revealed anchoring patterns in transcriptional regulation shared by TBP interactors, providing insight into the competitive multiprotein TBP interplay critical to transcriptional regulation (Anandapadamanaban et al. 2013). Mot1 is an ATP dependent inhibitor of TBP/TATA-DNA complex formation (Auble and Hahn 1993). Mot1 regulates the genomic distribution of TBP and was shown to influence transcription levels both positively and negatively (Pereira et al. 2003). Recent structural analysis revealed the molecular mechanism of Mot1 wrapping around TBP resembling a bottle opener, with a 'latch helix' blocking the concave DNA-binding surface of TBP and acting as a chaperone to prevent DNA re-binding to ensure promoter clearance (Wollmann et al. 2011). Mot1 and negative co-factor 2 (NC2) are thought to cooperate in gene-specific repression of TBP activity (Hsu et al. 2008). The GTF TFIIA, on the other hand, competes with NC2 for TBP (Kamada et al. 2001; Xie et al. 2000) and promotes TBP/DNA interactions in a ternary TFIIA/TBP/DNA complex, facilitating formation of and stabilizing the preinitiation complex (PIC). Interaction of TFIIA with TBP results in the exclusion of negative factors that would interfere with PIC formation, and TFIIA acts as a coactivator assisting transcriptional activators in increasing transcription levels (Bleichenbacher et al. 2003).

Genetic and biochemical experiments suggested that the TFIIA/TBP/DNA complex is further stabilized by the histone-fold containing TFIID subunits TAF11 and TAF13 conveying the formation of a TAF11/TAF13/TFIIA/TBP/DNA assembly (Kraemer et al. 2001; Lavigne et al. 1999; Robinson et al. 2005). We therefore set out to investigate this putative pentameric complex in detail. Unexpectedly, we did not observe a stabilization of TFIIA/TBP/DNA complex by TAF11/TAF13 but found a marked destabilization of the TFIIA/TBP/DNA complex by TAF11/TAF13, resulting in the release of free DNA and the

121 formation of a stable ternary complex formed by TAF11/TAF13 and TBP. We analysed the
122 TAF11/TAF13/TBP complex biochemically and structurally utilizing a comprehensive,
123 integrative approach. We report a novel interaction of the TAF11/TAF13 dimer binding to
124 the concave DNA-binding groove of TBP, thus excluding TATA-box containing DNA.
125 Using pull-down experiments with immobilized TAF1-TAND, we demonstrate competition
126 between TAF11/TAF13 and TAF1-TAND for TBP binding. We identify a novel C-terminal
127 TBP-binding domain (CTID) within TAF13 which is highly conserved from yeast to man.
128 We probe key residues within this TAF13 CTID by mutagenesis *in vitro* and *in vivo* in cell
129 growth experiments, revealing a key role of this domain for viability. We contrast the
130 TAF11/TAF13 interaction with other TBP DNA-binding groove interactors including Mot1
131 and discuss the implications of our findings in the context of TFIID assembly. Based on our
132 results, we propose a novel, functional state of TFIID in transcription regulation.
133

RESULTS

Identification of a novel TAF11/TAF13/TBP ternary complex

We set out to analyze the structure of a putative pentameric TAF11/TAF13/TFIIA/TBP/DNA complex (Kraemer et al. 2001; Lavigne et al. 1999; Robinson et al. 2005), with the objective to better understand the possible roles of TAF11/TAF13 in TFIID function. First we purified human TAF11/TAF13 complex and TBP to homogeneity (Figure 1A). TFIIA in human cells is made from two precursor polypeptides, TFIIAαβ and TFIIAγ, with TFIIAαβ processed *in vivo* into two separate polypeptides, α and β, by proteolytic cleavage mediated by the protease Taspase1 (Hoiby et al. 2007). Recombinant human TFIIA is typically produced in *E. coli* by refolding from three separate polypeptides expressed in inclusion bodies, which correspond to the native α, β and γ chains (Bleichenbacher et al. 2003). To facilitate recombinant human TFIIA production, we designed a single-chain construct (TFIIA$^{s\text{-}c}$) by connecting α, β and γ by flexible linkers, based on atomic coordinates taken from the crystal structure of human TFIIA/TBP/DNA complex (Bleichenbacher et al. 2003). TFIIA$^{s\text{-}c}$ could be produced in high amounts in soluble form in *E. coli* and purified to homogeneity without any need for refolding steps (see Methods section). We stored highly purified TFIIA$^{s\text{-}c}$ at 4°C, and observed the formation of needle-shaped crystals in the storage buffer after several weeks. We improved the crystals manually and determined the 2.4Å crystal structure of TFIIA$^{s\text{-}c}$ (Figure 1A, Figure 1–Figure Supplement 1, Table 1). The crystal structure revealed a virtually identical conformation of unliganded TFIIA$^{s\text{-}c}$ as compared to TFIIA in the TFIIA/TBP/DNA complex. Moreover, the crystal structure highlighted the importance of the connecting loops we had introduced in stabilizing the three-dimensional crystal lattice (Figure 1–Figure Supplement 1). TFIIA$^{s\text{-}c}$ was active in a band-shift assay with TBP and adenovirus major late promoter (AdMLP) TATA-DNA (Figure 1A), similar to purified TFIIA using the classical refolding protocol (Bleichenbacher et al. 2003).

Next, we attempted to reconstitute the TAF11/TAF13/TFIIA/TBP/DNA complex following published procedures (Kraemer et al. 2001; Robinson et al. 2005). Titrating TAF11/TAF13 dimer to a preformed TFIIA/TBP/DNA complex had been reported to stabilize TFIIA/TBP/DNA in band-shift assays using AdMLP TATA-DNA (Robinson et al. 2005). Surprisingly, in our titration experiments, TAF11/TAF13 did not stabilize the preformed TFIIA/TBP/DNA complex but, in marked contrast, resulted in TAF11/TAF13-dependent release of free promoter-containing DNA in band-shift assays (Figure 1B). We analysed the possible underlying intermolecular interactions that may be formed between the individual components TAF11/TAF13, TFIIA$^{s\text{-}c}$, TBP and AdMLP TATA-DNA. We observed that TFIIA$^{s\text{-}c}$ and TAF11/TAF13 did not interact in our experiments (Figure 1–Figure Supplement 2). Combining the TAF11/TAF13 dimer with TBP, in contrast, revealed a novel TAF11/TAF13/TBP complex that was stable in size exclusion experiments (Figure 1C). We concluded that human TAF11/TAF13 did not further stabilize the preformed TFIIA/TBP/DNA complex, but rather sequestered TBP from this complex giving rise to a novel assembly comprising TAF11, TAF13 and TBP.

TATA-DNA and TAF1-TAND compete with TAF11/TAF13 for TBP binding

We tested competition between TAF11/TAF13/TBP formation and TBP binding to AdMLP DNA and showed that the ternary TAF11/TAF13/TBP complex remained stable in the presence of excess AdMLP DNA (Figure 1D). Thus, our results indicate that TAF11/TAF13 and TATA-DNA compete for at least parts of the same binding interface within TBP, and that once TAF11/TAF13 is bound to TBP, TATA-DNA binding is precluded.

TAF1 had been shown previously to bind to the DNA-binding surface of TBP via its TAND domain (Anandapadamanaban et al. 2013; Burley and Roeder 1998; Liu et al. 1998). We produced human TAF1-TAND fused to maltose-binding protein (MBP) and immobilized highly purified fusion protein to an amylose column (Methods). We added preformed,

purified TAF11/TAF13/TBP complex in a pull-down assay using immobilized TAF1-TAND. We found that TAF1-TAND effectively sequestered TBP from the TAF11/TAF13/TBP complex, evidenced by TAF11/TAF13 eluting in the flow-through fraction. Elution by maltose, in contrast, revealed a TAF1-TAND/TBP complex (Figure 1E). Together, these findings substantiate our view that TAF11/TAF13, TAF1-TAND and AdMLP TATA-DNA all interact with the concave DNA-binding surface of TBP, and that the interactions are mutually exclusive.

TAF11, TAF13, TBP form a 1:1:1 complex

Next we set out to determine the subunit stoichiometry within the TAF11/TAF13/TBP complex, by using two complementary methods, analytical ultra-centrifugation (AUC) and native mass-spectrometry (native MS). Both methods were consistent in revealing a 1:1:1 complex (Figure 2A, B; Figure 2–Figure Supplement 1). Collision-induced dissociation (CID) in native MS, evidenced a TBP monomer and a TAF11/TAF13 dimer (Figure 2B). Of note, our control experiment using highly purified TBP evidenced a dimer in the native MS (Figure 2–Figure Supplement 1). To obtain first three-dimensional information of the TAF11/TAF13/TBP complex, we performed small-angle X-ray scattering (SAXS) experiments (Figure 2C, Figure 2–Figure Supplement 2, Table 2). Superimposition of the SAXS envelopes of the dimeric TAF11/TAF13 complex on one hand, and the ternary TAF11/TAF13/TBP complex on the other, provided further evidence that TAF11/TAF13 binds to the concave surface of TBP and underscored the presence of one copy each of the constituent proteins in the based on steric considerations. TAF11 and TAF13 contain unstructured regions, notably in the N-terminal part of TAF11, giving rise to a pan-handle shaped extension in the TAF11/TAF13 SAXS envelope, which is not present in SAXS envelopes of a truncated TAF11/TAF13 dimer comprising the more globular histone fold domains only (data not shown). This extension, which was likewise observed in the

TAF11/TAF13/TBP SAXS data, facilitated superimposition of the respective envelopes considerably. In aggregate, our results are consistent with a ternary assembly which accommodates one copy each of TAF11, TAF13 and TBP in the complex.

TAF11/TAF13 interacts with the concave DNA-binding surface of TBP

With the objective to elucidate the nature of the TAF11/TAF13 interaction with TBP, and to provide direct evidence that TAF11/TAF13 indeed engages the DNA-binding concave surface of TBP, we performed hydrogen-deuterium exchange/mass spectrometry (HDX-MS) (Rajabi et al. 2015) experiments (Figure 2D, Table 3). We analysed unbound TAF11/TAF13 and TAF11/TAF13/TBP, respectively, and compared changes in the deuteration levels in proteolytic peptides by MS indicating the level of protection of the respective regions in TAF11, TAF13 or TBP, respectively, upon ternary complex formation. The HDX-MS results underscored that the DNA binding surface of TBP was recognized by TAF11/TAF13, evidenced by decrease in deuteration levels which corresponds to increased protection from the solvent of peptides located in the concave surface of TBP upon TAF11/TAF13 binding. The extent of protection within TBP further indicates that the binding of TAF11/TAF13 engages both symmetric pseudo-repeats in TBP, thus spanning the entire concave interface (Figure 2D). Interestingly, we identified one peptide (AA residues 157-167) in TBP which evidenced an increased level of deuteration upon TAF11/TAF13/TBP complex formation in the HDX-MS experiments (Figure 2D). This peptide is located at the dyad relating the two pseudo-symmetric repeats in TBP. We interpret this result as an indication that this particular region within TBP is more protected in a presumed TBP dimer which dissociates when TAF11/TAF13 is binding and the 1:1:1 complex is formed. Our HDX-MS experiments provide direct evidence that TAF11/TAF13 engage to the concave DNA-binding surface of TBP, in excellent agreement with our above described biochemical experiments involving TAF11/TAF13, TBP, AdMLP DNA and TAF1-TAND.

Architecture of the TAF11/TAF13/TBP complex

We proceeded to determine the architecture of the TAF11/TAF13/TBP complex by using a comprehensive, integrative multi-parameter approach. We utilized the available crystal structure of TBP (Nikolov et al. 1992) as well as the crystal structure of the globular histone-fold containing domains of the TAF11/TAF13 dimer (Birck et al. 1998) in our approach, and combined these atomic coordinates with our native MS, SAXS, AUC and HDX-MS results. We acquired distance constraints to define our structural model by carrying out cross-linking/mass-spectrometry (CLMS) experiments using two different approaches. We first carried out a series of CLMS experiments using the cross-linker bis(sulfosuccinimidyl)suberate (BS3) (Figure 3–Figure Supplement 1; Table 4). BS3 cross-links primary amines on the side chain of lysine residues and the N-terminus of polypeptide chains. Inclusion of the BS3 CLMS derived distance constraints into our calculation already evidenced that the TAF11/TAF13 engaged the concave DNA-binding surface of TBP. In addition, we carried out site-specific UV-induced CLMS experiments utilizing MultiBacTAG (Koehler et al. 2016), a method we recently developed to unlock protein complex chemical space (Figure 3–Figure Supplement 2). MultiBacTAG relies on genetic code expansion (GCE) and employs a modified MultiBac baculoviral genome into which we engineered expression cassettes encoding for the orthogonal pyrolysine tRNA ($tRNA^{Pyl}$) / tRNA synthetase (PylRS) pair from *Methanosarcina mazei*. Infection of insect cell cultures with a MultiBacTAG virus containing TAF11/TAF13 mutants harboring an AMBER stop codon resulted in efficient incorporation of the UV-activatable amino acid diazirin lysine (DiAzK) provided in the culture medium, leading to site-specific cross-links upon UV irradiation of the reconstituted complex (Koehler et al. 2016) (Figure 3–Figure Supplement 2).

Our final three-dimensional TAF11/TAF13/TBP ternary complex model accommodated more than 90% of all experimental constraints and evidenced a tight

association of the TAF11/TAF13 histone-fold regions with the concave surface of TBP, giving rise to a compact structure (Figure 3, Figure 3–Figure Supplement 3). To validate our approach, we carried out calculations using alternative starting models. For instance, we rotated TBP by 180° around its axes to artificially expose the convex surface to TAF11/TAF13, or, alternatively, to reverse the location of the N- and C-terminal stir-ups of TBP (data now shown). These alternative calculations were far inferior in accommodating experimental spatial and distance restraints, in addition to being inconsistent with our biochemical data, thus substantiating our TAF11/TAF13/TBP structural model.

A highly conserved domain within TAF13 is required for TAF11/TAF13 interaction with TBP and for supporting cell growth.

In our TAF11/TAF13/TBP complex, the HF domains of TAF11 and TAF13 invade the concave DNA binding surface of TBP. Moreover, the structural model conveys that the C-terminal extension abutting the histone fold domain of TAF13 may play a prominent role in stabilizing the interaction with TBP. We analyzed TAF13 primary sequences from yeast to human (Figure 4). Sequence alignments revealed a very high degree of sequence conservation in this C-terminal TAF13 domain, with key residues virtually identical in all TAF13 proteins analyzed (Figure 4A). Based on our observation, we mutated these signature residues in the human TAF13 C-terminal region and analyzed the effect of the mutations on TAF11/TAF13/TBP complex formation. In particular, we analyzed two TAF13 mutants, A and B (Figure 4). In Mutant A, we substituted with alanine conserved residues located in the center of the TAF11/TAF13/TBP interface, while in Mutant B we changed conserved amino acid residues located more to the periphery (Figure 4 – Figure Supplement 1A). Both TAF13 mutants readily formed dimers with TAF11. SEC experiments evidenced impairment of TBP interaction with both mutants, but with notable differences. In case of Mutant A, TBP interaction was completely abolished and the TAF11/TAF13/TBP complex was not detected.

In case of Mutant B, on the other hand, TBP interaction was likewise diminished, however, residual TAF11/TAF13/TBP complex formation was clearly observed (Figure 4B). These resultsprovide evidence for a C-terminal TBP interaction domain (CTID) in TAF13, which is highly conserved throughout evolution. As the human TAF13 CTID is very well conserved, we generated two mutants in *Saccharomyces cerevisiae* (sc) Taf13 CTID that were mutating the same amino acid residues that were deleterious in the human TAF11/TAF13/TBP complex interactions. With plasmids expressing these scTaf13 mutants (scMutant A and scMutant B) we carried out *in vivo* rescue experiments in yeast, using previously described temperature sensitive *Taf13* mutant strains (Shen et al. 2003; Lemaire et al. 2000). While plasmids expressing either wild type (WT) or Mutant B Taf13 proteins supported growth at the non-permissive temperature (both on solid media and in suspension cell culture at close to wild-type levels), Mutant A did not rescue the lethal phenotype at 37°C, evidencing that the conserved mutations in this mutant effectively arrested cell growth (Figure 4C).

We corroborated our results by means of a novel Taf13 degron yeast strain (Warfield et al. 2017). In this strain, endogenous Taf13 is fused to an auxin-inducible degron (AID) tag resulting in Taf13-AID depletion upon addition of a chemical, indole-3acetic acid (IAA). In excellent agreement with the ts strain experiments, the plasmids expressing wild-type or Mutant B Taf13 supported growth after IAA addition, while expression of Mutant A lead to cell growth arrest (Figure 4D). Thus, our observations consistently suggest that the amino acids in TAF13 CTID, which when mutated destroy TAF11/TAF13 interactions with TBP, are required for functional TFIID formation.

Co-immunoprecipitation experiments reveal cytoplasmic TAF11/TAF13 and TBP dynamics in nuclear holo-TFIID

We recently demonstrated that human TFIID assembly involves preformed cytosolic and nuclear submodules (Trowitzsch et al. 2015), and we now asked whether the human

309 TAF11/TAF13/TBP complex would likewise represent such a sub-assembly. To this end we
310 performed co-immunoprecipitations (co-IPs) from HeLa cell cytosolic and nuclear extracts
311 using an anti-TAF11 antibody (Figure 5). We found dimeric TAF11/TAF13 complex in the
312 cytosol representing the complete HF pair. We could not detect TBP in cytosolic co-IPs,
313 however, our experiments evidenced TAF7 association with cytoplasmic TAF11/TAF13. The
314 anti-TAF11 co-IP from nuclear extract, in contrast, contained all TFIID components.
315 Surprisingly, normalized spectral abundance factor (NSAF) (Zybailov et al. 2007) analyses of
316 several distinct anti-TFIID IPs (i.e. anti-TAF1 and anti-TAF7) from NE indicate that less
317 than half of the nuclear TFIID specimens appear to contain stably bound TBP, implying
318 considerable dynamics in TBP association with TFIID in the nucleus, possibly regulated by
319 TBP containing TFIID submodules such as TAF11/TAF13/TBP.
320

DISCUSSION

Given the central role of TBP in eukaryotic transcription, it is not surprising that the activity of TBP is highly regulated, both positively and negatively, by GTFs, cofactors and gene-specific activators. In addition to TFIIA, a number of protein factors were identified to regulate the DNA binding capability of TBP. Several of these factors interact directly with the concave DNA binding surface and are capable of displacing TATA-box containing promoter DNA. The structures of the TAF1 N-terminal domains from *Drosophila* and yeast were determined, exhibiting TATA-box mimicry when bound to TBP (Figure 6, Figure 6–Figure Supplement 1). TBP and BTAF1 in mammals, or TBP and Mot1p in yeast, form a heterodimeric complex called B-TFIID (Auble and Hahn 1993). Mot1p was shown to utilize a flexible loop to target the DNA-binding surface of TBP thus precluding TATA-DNA binding (Wollmann et al. 2011). In our present study, we analyzed the interactions amongst TAF11/TAF13, TBP, TFIIA and TATA-box containing promoter DNA. TAF11 and TAF13 form a tight dimeric complex held together by pairing of the histone fold domains contained within these TAFs (Birck et al. 1998). Existing data conveyed a putative TAF11/TAF13/TFIIA/TBP/DNA complex which may represent a molecular building block in early stage preinitiation complex formation. In marked contrast to previous reports, our careful and well-calibrated titration of TAF11/TAF13 to preformed TFIIA/TBP/DNA did not result in a stabilization of the TFIIA/TBP/DNA complex. Rather, we observed that titration of TAF11/TAF13 resulted in the binding of TAF11/TAF13 to TBP and the release of free promoter DNA. Dissection of the underlying molecular interactions revealed a stable ternary complex comprising the TAF11/TAF13 HF pair and TBP. Our comprehensive multi-parameter approach revealed a compact 3-D structure in which TAF11/TAF13 bound tightly to the concave DNA-binding surface of TBP, fully consistent with our observations that TAF11/TAF13 could displace TATA-box containing DNA from a TBP/DNA complex. In

addition, we identified in our experiments a novel C-terminal domain within TAF13 that is essential for binding to TBP, and moreover markedly conserved throughout evolution. We generated TAF13 mutants and could demonstrate that mutations of key residues within this highly conserved domain, while not perturbing TAF13/TAF11 interactions in our experiments, had a profound effect on TBP binding *in vitro* and cell growth *in vivo*, effectively resulting in cell growth arrest. Taken together, our results indicate that the interaction of TAF11/TAF13 with TBP supports viability. Furthermore, we also demonstrated that TAF1-TAND and TAF11/TAF13 can compete for TBP binding. In summary, we provide compelling evidence that TAF11/TAF13, TATA-box DNA and TAF1-TAND share the same interaction interface in TBP. Careful inspection of the molecular modes by which different interactors engage the concave DNA-binding surface of TBP reveals that TAF11/TAF13 is unique in spanning the entire concave groove including the stir-ups of the saddle-shape adopted by TBP, with a solvent-excluded surface comparable to TATA-box DNA binding to TBP (Figure 6, Table 5).

Holo-TFIID is thought to exist in distinct structural states, based on cryo-EM analyses (Cianfrocco et al. 2013; Louder et al. 2016). In the canonical state, TAF1 is proposed to associate via its TAND domain to the DNA-binding surface of TBP thereby inhibiting TBP/TFIID binding to TATA-box containing core promoter DNA. In the activated state, TFIID was proposed to undergo major conformational rearrangements, likely involving interactions with transcriptional activators, thus unmasking TBP to promote DNA binding stabilized by TFIIA initiating transcription. Our results suggest that, in TFIID, several distinct TBP/TAF interactions exist, which are formed to forestall unwanted TFIID/DNA interactions which could otherwise lead for instance to cryptic transcription initiation on genomic regions that do not contain promoter elements. In addition to the TAF1-dependent inhibited canonical state, we propose an alternative inhibitory TFIID state in which the TAF11/TAF13 HF pair

blocks TBP from binding TATA-box containing promoter DNA (Figure 7). This alternative inhibited state may serve as a further point of transcriptional control, possibly depending on promoter context or additional gene regulatory factors bound. Our results imply that, in a given TFIID complex, the TAF1-dependent and the TAF11/TAF13 HF pair-dependent TBP blocking activities are mutually exclusive, but they may compete with each other to ascertain full blocking activity. Interestingly however, it appears that this TAF11/TAF13 HF pair-dependent TBP binding/blocking activity is essential/required for normal TFIID function, because when we interfered in the TBP binding through mutating the CTID, yeast growth was compromised at the non-permissive conditions. Thus, it is not clear at the moment whether or not the TAF1-dependent and the TAF11/TAF13 HF pair-dependent TBP blocking activities are really competing with each other, or would be simply part of a step-wise TFIID conformational change, or "activation", process that would allow TFIID to bind to DNA only when open promoter structures would become available. Further experiments will be needed to answer these exciting questions. Transcription activators and chromatin remodeling factors may direct inhibited TFIID to specific promoters, which could be poised to be transcribed by histone H3K4 trimethylation, and alleviate the TBP-blocking through TAF-interactions or by TAF-chromatin mark interactions. Alternatively, it is conceivable that once TFIID is brought to a promoter by interactions with transcription activators and positive chromatin marks (i.e. histone H3K4me3), DNA and TFIIA together may synergize to liberate the TATA-box binding surface of TBP from the inhibitory TAF-interactions.

The general roles of individual TAFs and the holo-TFIID complex are increasingly better understood, the mechanisms by which the cell assembles this essential multiprotein complex however remains largely enigmatic. The existence of discrete TFIID subassemblies containing a subset of TAFs, such as nuclear core-TFIID and the TAF2/TAF8/TAF10 complex present in the cytoplasm, provides evidence that holo-TFIID may be assembled in a

396 regulated manner in the nucleus from preformed submodules (Bieniossek et al. 2013; Gupta
397 et al. 2016; Trowitzsch et al. 2015). We analyzed TAF11-containing complexes by co-IP
398 experiments from the cytoplasm and the nucleus of HeLa cells. In cytosolic co-IPs, we
399 identified the TAF11/TAF13 histone fold pair, suggesting that this TFIID submodule may
400 also be preformed in the cytoplasm (Figure 5). Note however that this cytoplasmic
401 TAF11/TAF13 building block did not contain detectable amounts of TBP in our experiments,
402 suggesting that trimeric TAF11/TAF13/TBP, within TFIID or as a discrete TFIID
403 submodule, would be formed in the nucleus. Interestingly, we also identified TAF7
404 associated with TAF11/TAF13 in the cytoplasm, hinting at putative novel interactions
405 between TAF11/TAF13 and TAF7. TAF1 forms stable complexes with both TAF7 and TBP
406 (Gupta et al. 2016). We speculate that cytoplasmic TAF7/TAF11/TAF13 may represent an
407 assembly intermediate towards a TAF1/TAF7/TAF11/TAF13/TBP module, which may
408 integrate into a core-TFIID and TAF2/TAF8/TAF10 containing "8TAF" assembly
409 (Trowitzsch et al. 2015) in the formation pathway to the complete nuclear holo-TFIID
410 complex. In this TAF1/TAF7/TAF11/TAF13/TBP module, TBP would be tightly bound to
411 either TAF1 or TAF11/TAF13, which could serve to ascertain that this putative TFIID
412 submodule is efficiently blocked from any potentially detrimental interactions with DNA
413 until holo-TFIID formation is completed.

414 In the nucleus, IPs utilizing antibodies against several different TFIID specific TAFs
415 co-precipitated all known TFIID subunits, although with variable stoichiometry. Strikingly,
416 stoichiometry analyses carried out by NSAF calculations of our nuclear anti-TAF IPs
417 indicated that TBP was only present in less than half of the TFIID specimens, when
418 compared to TAF1 or TAF7 for example, suggesting that TFIID-type complexes may exist
419 which do not contain TBP. TBP is thought to be highly mobile structurally in the context of
420 holo-TFIID, with recent cryo-EM studies predicting large-scale migration of TBP within

421 distinct TFIID conformational states (Cianfrocco et al. 2013; Louder et al. 2016). Our
422 observations point to an additional level of compositional TBP and/or TAF dynamics in
423 TFIID formation, raising the interesting possibility that the accretion of TBP in holo-TFIID
424 may be regulated by partially assembled nuclear TFIID building blocks including for instance
425 the TAF11/TAF13/TBP complex we identified.
426

METHODS

DNA constructs. TAF11, TAF13 and TAF1-TAND were cloned in MultiBac baculovirus/insect cell transfer plasmids (Berger et al. 2004; Fitzgerald et al. 2006) and TBP^{fl}, TBP^{c} and $TFIIA^{s\text{-}c}$ expression constructs were cloned in *E. coli* expression plasmids. Constructs for genetic code expansion using the MultiBacTAG system were generated by PCR as described (Koehler et al. 2016). All constructs were verified by DNA sequencing.

Coding sequences of full-length TAF11 (Uniprot accession number Q15544) and TAF13 (Uniprot accession code Q15543) were synthesized at GenScript (New Jersey, USA). TAF11 contained an N-terminal hexahistidine tag spaced by a restriction site for Tobacco Etch Virus (TEV) NIA protease. TAF13 was cloned into MCS1 of the pFL acceptor plasmid from the MultiBac(Berger et al. 2004) suite via restriction sites NsiI and XhoI. TAF11 was inserted into MCS2 of the pFL-TAF13 plasmid via restriction sites RsrII and EcoRI.

The coding sequences for TFIIA α (AA2-59; Uniprot accession number P52655), β (AA302-376; Uniprot accession number P52655) and γ (AA2-110; Uniprot accession number P52657) were arranged into a single open reading frame by adding DNA sequences encoding for linkers L1(-DGKNTANSANTNTV-) and L2(-SRAVDGELFDT-). A C-terminal hexahistidine-tag was added to facilitate purification. The complete coding sequence was inserted into the bacterial expression plasmid pET28a via restriction sites NcoI and XhoI.

TBP full-length (UniProt accession number P20226) was synthesized (Genscript, New Jersey) and cloned via restriction enzymes NdeI and KpnI into a pET28a plasmid containing a hexahistidine-tag with a TEV cleavage site. TBP core (AA 155-335, UniProt accession number P20226) was generated from this plasmid by polymerase chain reaction (PCR).

The TAF1-TAND coding sequence (AA 26-168; UniProt accession number P21675) was cloned into a modified pUCDM vector coding for an engineered N-terminal TEV-

452 cleavable maltose-binding protein (MBP) tag using sequence and ligation independent
453 cloning (SLIC) (Li and Elledge 2007).

454 **Preparation of TAF11/TAF13 complex.** The human TAF11/TAF13 complex was co-
455 expressed in Sf21 insect cells using the MultiBac system (Berger et al. 2004). DNA encoding
456 for an N-terminal hexa-histidine tag and a protease cleavage site for tobacco etch virus (TEV)
457 NIa protease was added to the 5' end of the TAF11 open reading frame and cloned into pFL
458 plasmid (Berger et al. 2004). Cells were resuspended in Talon Buffer A (25mM Tris pH 8.0,
459 150mM NaCl, 5mM imidazole with complete protease inhibitor (Roche Molecular
460 Biochemicals). Cells were lysed by freeze-thawing (3 times), followed by centrifugation at
461 40,000g in Ti70 rotor for 60 min to clear the lysate. TAF11/TAF13 complex was first bound
462 to talon resin, pre-equilibrated with Talon Buffer A, followed by washes with Talon Buffer
463 A, then Talon Buffer HS (25mM Tris pH 8.0, 1M NaCl, 5mM imidazole and complete
464 protease inhibitor) and then again Talon Buffer A. TAF11/TAF13 complex was eluted using
465 Talon Buffer B (25mM Tris pH 8.0, 150mM NaCl, 200mM imidazole and complete protease
466 inhibitor). Fractions containing the TAF11/TAF13 complex were dialyzed overnight against
467 HiTrapQ Buffer A (50mM Tris pH 8.0, 150mM NaCl, 5mM β-ME and complete protease
468 inhibitor). Complex was further purified using ion exchange chromatography (IEX) with a
469 HiTrapQ column pre-equilibrated with HiTrapQ Buffer A. After binding, column was
470 washed with HiTrapQ Buffer A and TAF11/TAF13 eluted using a continuous gradient of
471 HiTrapQ Buffer B (50mM Tris pH 8.0, 1M NaCl, 5mM β-ME and complete protease
472 inhibitor) from 0-50%, followed by a step gradient to 100%. The complex was further
473 purified by size exclusion chromatography (SEC) with a SuperdexS75 10/300 column in SEC
474 buffer (25mM Tris pH 7.5, 300 mM NaCl, 1mM EDTA, 1mM DTT and complete protease
475 inhibitor). Mutants of TAF13 were generated by self-SLIC reaction (Haffke et al. 2013) and
476 complexes expressed and purified as wild-type.

Preparation of TBP. Full-length human TBP with an N-terminal oligo-histidine tag was expressed in *E. coli* BL21 (DE3) STAR cells at 30°C. Cells were lysed in Talon Buffer A by using a French press. Lysate was cleared by centrifugation at 40,000 g for 60 min. TBP^{fl} was eluted from TALON resin with Talon Buffer B using a continuous gradient. The tag was removed by TEV protease cleavage during dialysis overnight into Dialysis Buffer (25mM Tris pH 8.0, 300 mM NaCl, 5mM β-ME) and a reverse IMAC step was used to remove uncleaved protein. TBP^{fl} was polished using a SuperdexS75 16/60 equilibrated in SEC Buffer. The conserved TBP core (TBP^{c}) was expressed in *E. coli* Rosetta (DE3) cells at 18°C and purified as described for TBP^{fl}.

Preparation of double-stranded TATA-box containing promoter DNA substrate. AdMLP) TATA-DNA was prepared from synthetic olgonucleotides d(ctgctataaaaggctg) and d(cagccttttatagcag) (BioSpring GmbH) by mixing the complementary strands in equimolar amounts in Annealing Buffer (10 mM Tris pH 8.0, 50mM KCl, 5mM $MgCl_2$), heating to 96°C for 2 min and slow-cooling to room temperature.

Design and production of $TFIIA^{s-c}$. A construct encoding human TFIIA was prepared by structure-based design starting from the TFIIA/TBP/DNA structure (PDB ID 1NVP) by introducing a linker (L1) with sequence DGKNTANSANTNTVP between the TFIIA γ chain and the α chain. Similarly, a second linker (L2) with sequence SRAVDGELFDT was introduced connecting the α chain with the β chain giving rise to a single-chain $TFIIA^{s-c}$ construct γ-L1-α-L2-β encompassing 240 amino acid residues in total. The gene encoding for $TFIIA^{s-c}$ was cloned in a pET28a plasmid resulting in frame with a C-terminal hexa-histidine tag. $TFIIA^{s-c}$ was expressed in *E. coli* BL21 (DE3) cells at 18°C. Cells were lysed using a French press in Binding Buffer (20 mM Tris–HCl pH 7.4, 150 mM NaCl and complete protease inhibitor). The lysate was cleared by centrifugation at 40,000*g* for 45 min, and

loaded on a Talon affinity column. After 10 column volumes of washing with Binding Buffer, TFIIA$^{s\text{-}c}$ was eluted using Elution Buffer (20 mM Tris–HCl pH 7.4, 150 mM NaCl and 250 mM imidazole). Eluted protein was dialyzed overnight in 20mM Tris-HCl pH 7.4, 150mM NaCl, 0.5mM EDTA and 1mM DTT and loaded onto a Heparin column. TFIIA$^{s\text{-}c}$ was eluted with HS Buffer (20mM Tris-HCl pH 7.4, 1M NaCl, 0.5mM EDTA and 1mM DTT) applying a gradient, and polished by SEC with a Superdex75 column equilibrated in SEC Buffer 2 (20mM Tris-HCl pH 7.4, 150mM NaCl, 0.5mM EDTA and 1mM DTT). Purified TFIIA$^{s\text{-}c}$ protein was aliquoted and stored in Storage Buffer (20mM Tris pH 8.0, 1mM DTT, 0.5mM EDTA and 150mM NaCl) at −80°C.

X-ray crystallography. Large crystals of TFIIA$^{s\text{-}c}$ were obtained by vapor diffusion at room temperature from a protein solution concentrated to 15mg/ml in Storage Buffer and equilibrated against 20 mM Tris pH 8.0 with 25 mM NaCl in the reservoir. Best crystals were obtained by streak-seeding with the TFIIA$^{s\text{-}c}$ crystals spontaneously formed in the Eppendorf tube used for storing the protein. Crystals were harvested and mounted using perfluoropolyether (PFO-X175/08) as cryo-protectant. X-ray diffraction data were collected using a Pilatus 6M detector at beamline ID29 at the European Synchrotron Radiation Facility (ESRF). TFIIA$^{s\text{-}c}$ crystals diffracted to 2.4Å resolution. Data were processed using XDS (Kabsch 2010). The structure was determined by molecular replacement (MR) utilizing software PHASER (McCoy 2007), with the TFIIA crystal coordinates from the human TBP/TFIIA/TATA-DNA structure (Bleichenbacher et al. 2003) used as a search model. Iterative cycles of refinement and model building were performed using REFMAC5 and COOT from the CCP4 suite (Winn et al. 2011). Residues 2-210 could be modeled unambiguously in the electron density maps. Refinement converged at R and R_{free} values of 0.18 and 0.24, respectively. Refinement statistics are provided in Table 1. TFIIA$^{s\text{-}c}$ coordinates and structure factors were submitted to the PDB (5M4S).

Band shift assay. Samples for electrophoretic mobility shift assay were prepared by mixing annealed dsDNA (2 μM) with TBP^{c} (4 μM) or TBP^{fl} (4 μM), respectively, and $TFIIA^{s\text{-}c}$ (6μM) in EMSA Reaction Buffer (10mM Tris pH 8.0, 60mM KCl, 10mM MgCl2, 10% glycerol, 2.5mM DTT). Purified TAF11/TAF13 was added to aliquots of this mix at increasing concentrations (2 μM to 64 μM) followed by 1.5 hour incubation on ice. Samples were analysed by non-denaturing 5% polyacryl-amide gel electrophoresis (PAGE) using EMSA Running Buffer (25 mM Tris, 190 mM Glycine, 5 mM Mg Acetate, pH 8.0). Gels were stained with ethidium bromide (EtBr).

Preparation of TAF11/TAF13/TBP complex. TAF11/TAF13 (wild-type and mutants) was mixed with TBP in a 1:1.1 molar ratio in Complex Reaction Buffer (25mM Tris pH 8.0, 300 mM NaCl, 1mM EDTA, 1mM DTT, Leupeptin, Pepstatin and complete protease inhibitor) and incubated on ice for 90 min, followed by SEC with a SuperdexS200 10/300 column pre-equilibrated in Reaction Buffer.

Analytical ultracentrifugation. Purified TAF11/TAF13/TBP complex was analysed by analytical ultracentrifugation (AUC) in an An-60Ti rotor in a Beckman XL-I analytical ultracentrifuge (Beckman Coulter) at 42,000 rpm at 10°C for 16 h. Data was analysed with software Sedfit (Schuck 2000).

Native mass spectrometry. Proteins and complexes were buffer exchanged into 150 mM ammonium acetate pH7.5 before MS analysis using Vivaspin 10 kDa MWCO (Sartorius, Goettingen, Germany). 2 ul of the protein solution was then nano-electrosprayed from a gold coated borosilicate glass capillaries made in the Robinson lab (Hernandez and Robinson 2007). All mass spectrometry measurements were performed on a QToF instrument optimized for high mass measurements in positive ion mode. MS spectra were recorded at capillary and cone voltages of 1.7 kV and 80 V respectively. Other instrument parameters

were ToF pressure 1.76 x 10^6 mbar and analyser pressure of 2 x 10^4 mbar. For the collision induced dissociation the collision energy was increased up to 100 V to induce dissociation. All mass spectra were calibrated using an aqueous solution of caesium iodide and data were processed using MassLynx software V4.1.

TAF1-TAND production and competition assay. Human TAF1-TAND (Liu et al. 1998; Mal et al. 2004) containing an N-terminal maltose-binding protein (MBP) tag was expressed in Sf21 insect cells using MultiBac (Berger et al. 2004). Cells were lysed in MBP Binding Buffer (20mM Tris pH 8.0, 300mM NaCl, 1mM EDTA, 1mM DTT, complete protease inhibitor) by freeze-thawing, followed by centrifugation at 20,000 rpm in a JA25.5 rotor for 45 min. MBPTAF1-TAND was bound to amylose resin pre-equilibrated in MBP Binding Buffer, followed by extensive washing (20 column volumes). MBPTAF1-TAND loaded resin was incubated with an excess of preformed TAF11/TAF13/TBP complex for 60 min at 4°C. The column was washed and flow-through as well as wash fractions were collected. Bound protein(s) were eluted using MBP Elution buffer (20mM Tris pH 8.0, 300mM NaCl, 1mM EDTA, 1mM DTT, 10mM Maltose, complete protease inhibitor). Samples were analysed by SDS-PAGE followed by staining with Coomassie Brilliant Blue (Sigma Aldrich).

Small-angle X-ray scattering experiments. Small-angle X-ray scattering (SAXS) experiments were carried out at the ESRF BioSAXS beamline BM29 (Pernot et al. 2013). 30µl of each of purified TAF11/TAF13/TBP, TAF11/TAF13 and TBP at a range of concentrations (Table 2) and SAXS Sample Buffer (25mM Tris pH 8.0, 300 mM NaCl, 1mM EDTA, 1mM DTT and complete protease inhibitor) were exposed to X-rays and scattering data collected using the robotic sample handling available at the beamline. Ten individual frames were collected for every exposure, each 2 seconds in duration, using the Pilatus 1M detector (Dectris AG). Data were processed with the ATSAS software package (Petoukhov et

al. 2012). Individual frames were processed automatically and independently within the EDNA framework (Incardona et al. 2009), yielding individual radially averaged curves of normalized intensity versus scattering angle $S=4\pi Sin\theta/\lambda$. Additional data reduction within EDNA utilized the automatic data processing tools of ATSAS software package, to combine timeframes, excluding any data points affected by aggregation induced by radiation damage, yielding the average scattering curve for each exposure series. Matched buffer measurements taken before and after every sample were averaged and used for background subtraction. Merging of separate concentrations and further analysis steps were performed manually using PRIMUS (Petoukhov et al. 2012). Forward scattering (I(0)) and radius of gyration (Rg) were calculated from the Guinier approximation, to compute the hydrated particle volume using the Porod invariant and to determine the maximum particle size (Dmax) from the pair distribution function computed by GNOM (Petoukhov et al. 2012). 40 *ab initio* models were calculated for each sample, using DAMMIF (Petoukhov et al. 2012), and then aligned, compared and averaged (evidencing minimal variation) using DAMAVER (Petoukhov et al. 2012). The most representative model for TAF11/TAF13/TBP and TAF11/TAF13 selected by DAMAVER were compared to each other as well as the known structure of TBP, with overlays of the resulting models generated in PyMOL. The fits to the experimental data of the models and the theoretical scattering of the calculated structures were generated with CRYSOL (Petoukhov et al. 2012).

Cross-linking/mass spectrometry (CLMS) experiments.

BS3 cross-linking*:* Purified TAF11/TAF13/TBP and TAF11/TAF13 complexes were cross-linked separately by BS3 at complex/BS3 ratio of 1:25 [w/w] in Cross-linking Buffer (25mM HEPES, pH 8.0, 300 mM NaCl, 1mM DTT, 1 mM EDTA and complete protease inhibitor) for 2 h on ice. The reaction was quenched by adding saturated ammonium bicarbonate solution followed by incubation on ice (45 min). Cross-linked samples were then further

599 purified by injecting on a SuperdexS200 10/300 column. Peak fractions containing purified
600 cross-linked samples were concentrated using pin concentrators (Amicon) and separated by
601 SDS-PAGE using a NuPAGE 4-12% bis-Tris gel (Thermo Fisher Scientific).

602 The gel bands corresponding to cross-linked complexes were reduced, alkylated and
603 trypsin digested following standard procedures (Maiolica et al. 2007) and digested peptides
604 were desalted using C18 StageTips (Rappsilber et al. 2007). Peptides were analysed on an
605 LTQ Orbitrap Velos mass spectrometer (Thermo Fisher Scientific) that was coupled with a
606 Dionex Ultimate 3000 RSLC nano HPLC system. The column with a spray emitter (75-µm
607 inner diameter, 8-µm opening, 250-mm length; New Objectives) was packed with C18
608 material (ReproSil-Pur C18-AQ 3 µm; Dr Maisch GmbH, Ammerbuch-Entringen, Germany)
609 using an air pressure pump (Proxeon Biosystems) (Ishihama et al. 2002). Mobile phase A
610 consisted of water and 0.1% formic acid. Mobile phase B consisted of 80% acetonitrile and
611 0.1% formic acid. Peptides were loaded onto the column with 2% B at 500 nl/min flow rate
612 and eluted at 200 nl/min flow rate with two gradients: linear increase from 2% B to 40% B in
613 90 minutes or 120 mins; then increase from 40% to 95% B in 11 minutes. The eluted peptides
614 were directly sprayed into the mass spectrometer.

615 Peptides were analysed using a high/high strategy (Chen et al. 2010): both MS spectra
616 and MS2 spectra were acquired in the Orbitrap. FTMS full scan spectra were recorded at
617 100,000 resolution. The eight highest intensity peaks with a charge state of three or higher
618 were selected in each cycle for fragmentation. The fragments were produced using CID with
619 35% normalized collision energy and detected by the Orbitrap at 7500 resolution. Dynamic
620 exclusion was set to 90s and repeat count was 1.

621 ***DiAzKs cross-linking****:* An unnatural amino acid, DiAzKs, was introduced at K34 position of
622 TAF13 using genetic code expansion (GCE) we implemented recently in our
623 baculovirus/insect cell system (MultiBacTAG) (Koehler et al. 2016). TAF11/TAF13-

K34DiAzKs was purified similarly as wild type. Briefly, TAF11/TAF13-K34DiAzKs (as well as wild type) and TBP were mixed in 1:1.25 molar ratio in Incubation Buffer (25mM Tris, pH 8.0, 300 mM NaCl, 1mM DTT, 1 mM EDTA and complete protease inhibitor) and incubated on ice for 2 h. Reactions were then split into 2 aliquots. One aliquot was UV irradiated for 15 minutes on ice using a 345 nm filter with an approximately 40 cm distance to the 1000 W lamp. Cross-linked samples were then separated on SDS-PAGE using a NuPAGE 4-12% bis-Tris gel (Thermo Fisher Scientific). Gel bands were processed as above for BS3 CLMS.

Peptides were analysed on an Orbitrap Fusion Lumos Tribrid mass Spectrometer (Thermo Fisher Scientific) coupled to a Dionex UltiMate 3000 RSLC nano HPLC system using a 75 µmx50cm PepMap EASY-Spray column (Thermo Fisher Scientific). Eluted peptides were directly sprayed into the mass spectrometer through EASY-Spray source (Thermo Fisher Scientific) and analysed using a high/high strategy (Chen et al. 2010): both MS spectra and MS2 spectra were acquired in the Orbitrap. MS1 spectra were recorded at 120,000 resolution and peptides with charge state of 3 to 8 were selected for fragmentation at top speed setting. The fragments were produced using HCD with 30% normalized collision energy and detected by the Orbitrap at 15000 resolution. Dynamic exclusion was set to 60s and repeat count was 1. Peak lists were generated by MaxQuant (version 1.5.3.30) (Cox and Mann 2008) at default parameters except for “top MS/MS peaks per 100 Da” being set to 100. Cross-linked peptides were matched to spectra using Xi software (ERI, Edinburgh).

Hydrogen-deuterium exchange/mass spectrometry. Hydrogen-deuterium exchange/mass spectrometry (HDX-MS) experiments were fully automated using a PAL autosampler (CTC Analytics). This controlled the start of the exchange and quench reactions, the proteolysis temperature (4 °C), the injection of the deuterated peptides, as well as management of the injection and washing valves; it also triggered the acquisition of the mass spectrometer and

649 HPLC and UPLC pumps. A Peltier-cooled box (4 °C) contained two Rheodyne automated
650 valves, a desalting cartridge (Trap Acquity UPLC Protein BEH C18 2.1x5mm,Waters) and a
651 UPLC column (Acquity UPLC BEH C18 1.7µm 1x100mm, Waters). HDX-MS reactions
652 were carried out using either TAF11/TAF13 or TBP alone or in complex at a concentration of
653 20 µM. Deuteration was initiated by a 5-fold dilution of the protein samples (10 µl) with the
654 same buffer in D_2O (40 µl). The proteins were deuterated for 15 sec or 2 min at 4 °C. The
655 exchange was quenched using 50 µl of 200 mM glycine-HCl, pH 2.5 at 4 °C. The proteins or
656 complexes were digested online with immobilized porcine pepsin (Sigma) and recombinant
657 nepenthesin-1. The peptides were desalted for 6 min using a HPLC pump (Agilent
658 Technologies) with 0.1% formic acid in water, at a flow rate of 100 µl/min. Desalted peptides
659 were separated using a UPLC pump (Agilent Technologies) at 50 µl/min for 10 min with 15–
660 70% gradient B (Buffer A: 0.1% formic acid in water; Buffer B: 0.1% formic acid in 95%
661 acetonitrile), followed by 1 min at 100% B. The peptide masses were measured using an
662 electrospray-TOF mass spectrometer (Agilent 6210) in the 300–1300 m/z range. Each
663 deuteration experiment was conducted in triplicate. The Mass Hunter (Agilent Technologies)
664 software was used for data acquisition. The HD Examiner software (Sierra Analytics) was
665 used for HDX-MS data processing. Identification of peptides generated by the digestion was
666 done as described previously (Giladi et al. 2016). Different proteases (pepsin, nepenthesin-1,
667 nepenthesin-2, rhizopuspepsin) or their combinations were tested for protein digestion with
668 pepsin-nepenthesin-1 pair providing the best digestion parameters and sequence coverage.

669 **Integrative multiparameter-based model building and refinement.** Initial models of the
670 two component structures (TAF11/TAF13, TBP) were taken from the PDB (1BH8 and
671 1CDW) (Birck et al. 1998; Nikolov et al. 1996). 1BH8 was extended to include a helix
672 structure missing from the complete histone-fold domain as described before (Birck et al.
673 1998). The structure of the complex was constructed in a two-stage workflow. Initially, a

model of the structured core of the complex was constructed by rigid body docking using the HADDOCK webserver (de Vries et al. 2010). The resulting complex structures and their scores were visually analysed against the SAXS data to select the highest scoring structure that fit within the SAXS envelopes.

The selected complex with the highest scores was then refined integrating the cross-linking data. The HADDOCK complex was used as an input to MODELLER 9.14 (Webb and Sali 2014) with the complete sequences (including loop structures). Observed cross-links were included as restraints in the refinement with a mean distance of 11.4 Å. Refinement was performed iteratively until more than 90% of all distance constraints could be accommodated while maintaining the fit to the SAXS envelope.

Cell growth experiments

Yeast Taf13 wild-type (WT), as well as Mutant A and Mutant B, were cloned along with native promoters into the LEU2 (auxotrophic marker) containing plasmid pRS415 (Genscript Corp., Piscataway, NJ, USA) by using the BamHI and NotI restriction enzyme sites. Constructs thus generated were transformed into yeast strain BY4741 (comprising endogenous wild-type Taf13) as well as the temperature sensitive (ts) yeast strains TSA797 (ts *taf13*) and TSA636 (ts *taf13*) (EuroSCARF, SRD GmbH, Germany). Transformed yeast containing the plasmids were restreaked onto selective media and grown at permissive (30°C) or non-permissive (37°C) temperatures, and plates imaged. To determine growth rates, ts strains transformed with empty vector or Taf13 expression plasmids were grown in liquid media at 37°C. In a separate experiment, the above constructs were transformed into a Taf13-AID auxin-inducible degron strain obtained from Steven Hahn (Warfield et al. 2017), and grown at 30°C in liquid media supplemented with 500μM indole-3-acetic acid (IAA) or an equivalent volume of DMSO (used to prepare IAA stocks). Empty degron strain was used as a negative control. Absorbance at 600nm was measured every hour for all cell growth

699 experiments. Three (ts strains) or two (degron strain) independent experiments were
700 performed and data was normalized against the first time point taken. Average absorbance
701 was plotted against time, standard errors of mean (SEM) were calculated over each data
702 point.

703 For spot assays, overnight cultures of empty degron strain as well as degron strain
704 transformed with Taf13 wild-type and mutant expression plasmids were washed and
705 resuspended in milli-Q water to obtain identical densities. Serial 10-fold dilutions were
706 spotted on solid media (YPD, or synthetic drop-out media -LEU) supplemented with 500µM
707 IAA or DMSO, and incubated at 30°C for 48-72 hours.

708 **TFIID immunoprecipitation experiments.** HeLa cells were grown in suspension culture.
709 10^{11} cells were harvested by centrifugation and a nuclear extract was prepared according to a
710 modified protocol (Dignam et al. 1983). Briefly, nuclei were prepared by resuspending the
711 pellets in 4 packed cell volume (PCV) of 50 mM Tris-HCl, pH 7.9; 1 mM EDTA; 1 mM DTT
712 and proteinase inhibitors and opening the cells with a Dounce-homogenizer. Nuclei were
713 collected by centrifugation and lysed in 4 PCV of 50 mM Tris-HCl, pH 7.9; 25% glycerol;
714 500 mM NaCl; 0.5 mM EDTA; 1 mM DTT and protease inhibitors by powerful strokes. The
715 lysate was centrifuged at 50,000 g for 20 minutes. The supernatant was filtered and proteins
716 precipitating in 30% (w/v) $(NH_4)_2SO_4$ were pelleted. They were resuspended in 50 mM Tris-
717 HCl, pH 7.9; 20% glycerol; 100 mM KCl; 5 mM $MgCl_2$; 1 mM DTT and dialysed against the
718 same buffer.

719 For immunopreciptation. 200 µl protein G-Sepharose (Pharmacia) was incubated with
720 approximately 50 µg of the different antibodies (as indicated). Washed antibody-bound beads
721 were then mixed with 4 mg of pre-cleared HeLa cell nuclear extract and incubated overnight
722 at 4°C. Antibody-protein G Sepharose bound protein complexes were washed three times
723 with IP buffer (25 mM Tris-HCl pH 7.9, 10% (v/v) glycerol, 0.1% NP40, 0.5 mM DTT, 5

mM $MgCl_2$) containing 0.5 M KCl and twice with IP buffer containing 100 mM KCl. Immunoprecipitated proteins were eluted from the protein G columns with 0.1 M glycine (pH 2.5) and quickly neutralized with 2 M Tris-HCl (pH 8.8).

For analysis by mass spectrometry, samples were reduced, alkylated and digested with LysC and trypsin at 37°C overnight. They were then analysed using an Ultimate 3000 nano-RSLC (Thermo Fischer Scientific) coupled in line with an Orbitrap ELITE (Thermo Fisher Scientific). Briefly, peptides were separated on a C18 nano-column with a linear gradient of acetonitrile and analysed in a Top 20 CID (Collision Induced Dissociation) data-dependent mass spectrometry. Data were processed by database searching using SequestHT (Thermo Fisher Scientific) with Proteome Discoverer 1.4 software (Thermo Fisher Scientific) against the Human Swissprot database (Release 2013_04, 20225 entries). Precursor and fragment mass tolerance were set at 7 ppm and 0.5 Da respectively. Trypsin was set as enzyme, and up to 2 missed cleavages were allowed. Oxidation (M) was set as variable modification, and Carbamidomethylation (C) as fixed modification. Peptides were filtered with a 5 % FDR (false discovery rate) and rank 1. Proteins were identified with 1 peptide.

FIGURE LEGENDS

Figure 1: TAF11/TAF13 and TBP form ternary complex. (**A**) Human TAF11/TAF13 complex, TBP full-length (TBP^{fl}) and core (TBP^{c}), and a single-chain version of TFIIA ($TFIIA^{s-c}$) were purified to homogeneity as shown by SDS-PAGE (left). TFIIA subunits α(AA2–59), β(AA302–376) and γ(AA2–110) (Bleichenbacher et al. 2003) were connected with linkers L1(-DGKNTANSANTNTV-) and L2(-SRAVDGELFDT-) as indicated (middle, top). $TFIIA^{s-c}$ crystallized during purification (Figure 1–Figure Supplement 1). The 2.4Å X-ray structure is shown in a cartoon representation with a section of electron density at 1σ (middle, bottom). $TFIIA^{s-c}$ was assayed by band-shift for activity (right). DNA, Adenovirus major late promoter (AdMLP) DNA. (**B**) Probing formation of putative pentameric TAF11/TAF13/TBPTFIIA/DNA complex by band-shift assay (Kraemer et al. 2001; Robinson et al. 2005). TAF11/TAF13 titration to the TFIIA/TBP/DNA complex results in DNA release. (**C**) SEC analysis reveals a stable TAF11/TAF13/TBP ternary complex. Elution fractions (1-7) were analysed by SDS-PAGE (inset). IN, equimolar mixture of TAF11/TAF13 and TBP. No interactions were found between TAF11/TAF13/TBP complex and TFIIA (Figure 1–Figure Supplement 2). (**D**) TAF11/TAF13 competes with DNA for TBP binding, evidenced by SEC. Elution fractions (1-6, 11-16) were analysed by SDS-PAGE and ethidium-bromide stained agarose gel (inset). IN, preformed TAF11/TAF13/TBP complex; DNA, AdMLP DNA; M, protein molecular weight marker and DNA ladder, respectively. (**E**) Immobilized human TAF1 N-terminal domain (TAF1-TAND) (Anandapadamanaban et al. 2013) efficiently depletes TBP from preformed TAF11/TAF13/TBP complex. RS^{IN}, amylose resin with MBP-tagged TAF1-TAND bound; IN, preformed TAF11/TAF13/TBP; M: protein marker; FT, flow-through fraction; W, wash fraction; E, maltose elution fraction.

Figure 2: TAF11/TAF13/TBP interactions. (**A**) Sedimentation velocity analytical ultracentrifugation (AUC) experiments reveal a major peak consistent with a 1:1:1

765 TAF11/TAF13/TBP ternary complex. The second, smaller peak corresponds to excess
766 TAF11/TAF13. (**B**) Native mass spectrometry of the TAF11/TAF13/TBP complex confirms
767 1:1:1 stoichiometry. Collision induced dissociation (CID) results in TBP monomer and
768 TAF11/TAF13. Experimental masses are provided (inset). Calculated masses: 20659 Da
769 (TBP); 40691 Da (TAF11/TAF13); 61351 Da (TAF11/TAF13/TBP). Mass spectra of the
770 TAF11/TAF13/TBP complex and TBP dimer are shown in Figure 2–Figure Supplement 1.
771 (**C**) Small-angle X-ray scattering (SAXS) data of TAF11/TAF13/TBP and TAF11/TAF13.
772 Fit of experimental scattering (dots) and scattering calculated from *ab initio* models (solid
773 lines) is provided in a log plot (left). SAXS envelopes of TAF11/TAF13 (yellow),
774 TAF11/TAF13/TBP (grey) and an overlay are shown (right). TBP (PDB ID 1CDW) and
775 TAF11/TAF13 (PDB ID 1BH8) were placed into the overlay. A distinct extension
776 (panhandle), likely arising from unstructured tail regions in TAF11/TAF13, enabled
777 alignment of the SAXS envelopes. Kratky plot, P(r) and Guinier analysis for corresponding
778 complexes are provided in Figure 2–Figure Supplement 2. (**D**) TAF11/TAF13/TBP and
779 TAF11/TAF13 were analysed by hydrogen-deuterium exchange/mass spectrometry (HDX-
780 MS) (Rajabi et al. 2015). Changes in the deuteration level of selected peptides in TAF11,
781 TAF13 or TBP are depicted in diagrams (top row). Peptides protected upon ternary complex
782 formation are coloured in red in cartoon representations of TAF11/TAF13 and TBP (bottom
783 row). One peptide in TBP (grey, far right) becomes more accessible, hinting at disassembly
784 of a TBP dimer when TAF11/TAF13 binds (Figure 2–Figure Supplement 1). All peptides
785 implicated in TAF11/TAF13 binding map to the concave DNA-binding surface of TBP. *(N)*
786 and *(C)* indicate N- and C-terminal TBP stirrups, respectively.

787 **Figure 3: Architecture of TAF11/TAF13/TBP complex.** TAF11/TAF13/TBP complex
788 architecture was determined by using an integrative multi-parameter approach, We utilized
789 the crystal structure of TBP (Nikolov et al. 1996) as well as the crystal structure of the

790 TAF11/TAF13 dimer (Birck et al. 1998) combined with our native MS, SAXS, AUC and
791 HDX-MS results as well as distance constraints from CLMS experiments (Figure 3–Figure
792 Supplement 1, Figure 3–Figure Supplement 2). The structure of the TAF11/TAF13/TBP
793 ternary complex is shown in a cartoon representation in stereo (top) and as a space filling
794 model (devoid of unstructured regions) in three views (bottom). Three axes (x, y, z, drawn as
795 arrows) illustrate the special relation between the views. TAF11 is colored in blue, TAF13 in
796 magenta and TBP in green. This model satisfies >90 % of the experimental constrains (Figure
797 3–Figure Supplement 3, Tables 3 and 4).

798 **Figure 4: Highly conserved C-terminal TBP-interaction domain (CTID) in TAF13**
799 **required for survival** (**A**) Sequence alignments reveal a highly conserved C-terminal TBP
800 interaction domain (CTID) in TAF13 comprising virtually identical signature residues in
801 TAF13 from yeast to man. Residues that were mutated in the CTID of TAF13 are indicated
802 by arrows, giving rise to two mutant TAF13 proteins (Mutant A, B). The locations of the
803 mutated residues in TAF11/TAF13/TBP are illustrated in Figure 4 - Figure Supplement 1A.
804 (**B**) SEC analysis demonstrates complete abolition of the TBP binding by TAF11/TAF13 in
805 case of Mutant A. In case of Mutant B, residual interaction with TBP is observed (marked by
806 asterisk). Elution fractions (1-8) were analysed by SDS-PAGE (inset). IN, equimolar mixture
807 of TAF11/TAF13 and TBP. (**C**) Cell growth experiments in yeast containing temperature
808 sensitive (ts) Taf13 on solid media plates at permissive (30°C) and non-permissive (37°C)
809 temperatures are shown on the left. EV, empty vector; WT, wild-type Taf13; MutA, MutB,
810 Taf13 mutants A and B; TSA797, TSA636, yeast strains harboring distinct temperature
811 sensitive Taf13 mutants (Shen et al. 2003; Lemaire et al. 2000). Corresponding absorbance
812 plots displaying growth curves of temperature sensitive strains in liquid media at the non-
813 permissive temperature (37°C) are provided on the right. Polynomial fits are shown as dotted
814 lines. Standard errors of mean (SEM) are shown as bars. The corresponding growth curves

for strain BY4741 used as a control, are shown in Figure 4–Figure Supplement 1B. (**D**) Cell growth experiments in yeast containing Taf13 fused to an auxin-inducible degron tag (AID) is shown in spot assays on solid media plates (YPD, -LEU) on the left, in presence or absence of indole-3-acetic acid (IAA) which activates Taf13-AID depletion. 13-AID, Taf13 degron-tag fusion (Warfield et al. 2017); YPD, yeast total media; -LEU, synthetic drop-out media. Corresponding absorbance plots displaying growth curves in presence or absence of IAA are shown on the right.

Figure 5: TAF/TFIID co-immunoprecipitations. Orbitrap mass spectroscopic analyses of proteins co-immunoprecipitated from nuclear (NE) or cytoplasmic HeLa cell extracts using mouse monoclonal antibodies against the indicated TAFs. The stoichiometry of the TAFs and TBP in the purified TFIID complexes was calculated by determining normalized spectral abundance factors (NSAFs) (Sardiu et al. 2008; Zybailov et al. 2007). Each column is the average of three independent MS runs. Blue arrows indicate the bait in each immunoprecipitation.

Figure 6: Distinct modes of TBP binding involving the concave DNA-binding surface. The interaction interfaces of TBP binders are shown in a cartoon representation (top). Interactors shown are TATA-box DNA and protein interactors including the TAF11/TAF13 dimer. The binding modes are further illustrated using space-filling models depicting the corresponding electrostatic surface potentials (bottom). The interacting region representing the concave DNA-binding surface of TBP is delimited by dashed lines. Structures shown are TBP on one hand, as well as TATA-DNA (PDB ID 1CDW), a second copy of TBP from the crystal structure of unliganded TBP (PDB ID 1TBP), TAF1-TAND (PDB ID 1TBA), Mot1 (PDB ID 3OC3) and TAF11/TAF13, respectively, on the other. (N) and (C) indicate N- and C-terminal TBP. The “TATA-box mimicry” by TAF1-TAND in shape and charge

839 distribution is evident. TAF11/TAF13 engage the entire concave DNA binding surface of
840 TBP including the stir-ups.

841 **Figure 7: Novel TFIID regulatory state comprising TAF11/TAF13/TBP.** TFIID is
842 thought to exist in an inhibited 'canonical' state with TAF1-TAND bound to TBP's DNA
843 binding surface (bottom left). Activated states of TFIID (right) bind promoter DNA stabilized
844 by TFIIA (Cianfrocco et al. 2013; Louder et al. 2016; Papai et al. 2010). Our results suggest a
845 novel, alternative TFIID inhibited state comprising TAF11/TAF13/TBP (top left). TFIID is
846 shown in a cartoon representation based on previous EM studies (Cianfrocco et al. 2013;
847 Louder et al. 2016; Papai et al. 2010). TFIID lobes A, B and C are indicated. TAF11/TAF13
848 are placed in lobe A as suggested by immune-labeling analysis (Leurent et al. 2002).
849 Promoter DNA is colored in grey.

850 **Figure 1–Figure Supplement 1: Crystal structure of TFIIA^{s-c} at 2.4 Å resolution.**
851 (**A**) Superimposition of human TFIIA from a TFIIA/TBP/DNA ternary complex (green, PDB
852 ID: 1NVP) and uncomplexed TFIIA^{s-c} in our crystal (blue) is shown in a cartoon
853 representation in a stereo view. The structures are virtually identical (rms deviation 0.753 Å).
854 C, N are C- and N-terminal end of single-chain; L1, L2 are Linker regions. (**B**) Crystal
855 packing of TFIIA^{s-c}, highlighting the critical role of engineered linkers (yellow spheres) in
856 mediating the majority of crystal contacts. The unit cell is indicated (grey lines). (**C**) A close
857 up highlights the interactions mediated by the linker region between symmetry related
858 molecules in the crystal lattice. Symmetry related molecules are colored in grey or green,
859 respectively. Hydrogen bonds (distance cut-off 3.7Å) are shown as black dashed lines.
860 Amino acid residue numbers are indicated. A section of the 2Fo-Fc electron density map
861 contoured at 1σ is shown (grey mesh). Crystal statistics are provided in Table 1.

Figure 1–Figure Supplement 2: Size exclusion chromatography (SEC) analysis of TAF11/TAF13/TBP and TFIIA$^{s\text{-}c}$.

Interaction analysis of TAF11/TAF13/TBP complex with TFIIA$^{s\text{-}c}$ by SEC. Preformed TAF11/TAF13/TBP complex and TFIIA$^{s\text{-}c}$ were combined and injected on a Superdex S200 10/300 column (left). Peak elution fractions were further analyzed by SDS-PAGE (right). TAF11/TAF13/TBP complex and TFIIA$^{s\text{-}c}$ eluted in separate peaks indicating a lack of interaction at the conditions studied. IN, Input sample; M, BioRad prestained marker (MW of marker bands are indicated). Elution fractions are numbered 1 to 11.

Figure 2–Figure Supplement 1: Native mass spectrometry (MS) analysis of TAF11/TAF13/TBP complex.

Mass spectrum of the TAF11/TAF13/TBP complex recorded from a buffer solution containing 150 mM ammonium acetate (pH7.5). Two charge state series are observed corresponding to complex with (circles filled in purple) and without (circles filled in orange) an oligo-histidine affinity purification tag, respectively. Calculated masses: 20659 Da (TBP); 22765 Da (TBP with His-tag); 45531 Da (TBP dimer); 40691 Da (TAF11/TAF13). The inset shows the native MS spectrum recorded for separately purified TBP alone, evidencing a TBP dimer.

Figure 2–Figure Supplement 2: Small Angle X-ray Scattering (SAXS) of TAF11/TAF13 and TAF11/TAF13/TBP.

Kratky plot, P(r) and Guinier analysis of the SAXS data of TAF11/TAF13/TBP and TAF11/TAF13, respectively are shown. (**A**) Kratky plots for TAF11/TAF13/TBP (light gray) and TAF11/TAF13 (light orange) indicate globular and flexible domains. (**B**) Guinier plots for TAF11/TAF13/TBP (left) and TAF11/TAF13 (right) are shown; Rg values are indicated.

Residuals of the linear fit are represented as green lines. (**C**) Distance distribution analysis for TAF11/TAF13/TBP (left) and TAF11/TAF13 (right). SAXS statistics are provided in Table 2.

Figure 3–Figure Supplement 1: Cross-linking/mass spectrometry (CLMS).

TAF11/TAF13/TBP complex and TAF11/TAF13 complex were cross-linked separately using bissulfosuccinimidyl suberate (BS3) and then further purified by SEC using a Superdex 200 10/300 column. Cross-linked and purified TAF11/TAF13/TBP complex (lane 1) and TAF11/TAF13/TBP complex (lane 2) were analyzed by SDS-PAGE (left). The top band corresponds to cross-linked TAF11/TAF13/TBP, the middle band to TAF11/TAF13; the band on the bottom to uncross-linked TAF11. Cross-linked TAF11/TAF13 is present also in the cross-linked TAF11/TAF13/TBP sample. Bands were excised, reduced, alkylated, trypsin digested and desalted followed by mass spectrometric analysis. Cross-links were identified between TAF11 and TAF13, TAF11 and TBP as well as TAF13 and TBP. Cross-links between TAF11-TAF13, TAF11-TBP and TAF13-TBP are shown using 5% FDR cutoff data (right). The N-terminal region of TAF11 is lacking amino acids R and K, and consequently, no cross-links were observed with BS3. Observed cross-links are listed in Table 4.

Figure 3–Figure Supplement 2: Site-specific cross-linking of TAF11 and TAF13 by Genetic Code Expansion (GCE)

The UV-activatable amino acid DiAzKs (Diazirine-Lysine) was introduced at position K34 in TAF13 by genetic code expansion (MultiBacTAG) (Koehler et al. 2016). Thus labelled TAF11/TAF13 complex and TBP were reconstituted into TAF11/TAF13/TBP complex, which was then exposed to UV light. A cartoon representation of TAF/TAF13/TBP is shown (left) with a zoom-in on the site of ncAA introduction (right). K34 of TAF13 was chosen as it appears to be located at/near the interfaces between TAF13, TAF11 and TBP in our model of

the ternary complex. Cross-linked sample then was separated by SDS-PAGE, excised from gel, reduced, alkylated, trypsin digested and desalted followed by mass spectrometry analysis. Specific cross-linking patterns were obtained upon TAF11/TAF13 complex formation with TBP. Example of mass spectrum for a cross-linked peptide is shown below.

Figure 3–Figure Supplement 3: Mapping CLMS and HDX-MS data on TAF11/TAF13/TBP complex.

(**A**) Cross-links identified by BS3 CLMS mapped on the TAF11/TAF13/TBP ternary complex. The current model satisfies >90 % of all observed cross-links. Cross-linking sites are depicted as red balls. (**B**) Regions on TBP exhibiting changes in deuteration levels upon complex formation with TAF11/TAF13 are highlighted in red.

Figure 4 – Figure Supplement 1: **TAF11 CTID mutant studies (A)** Location of amino acid residues mutated in TAF11 Mutant A and Mutant B, respectively, mapped on the TAF11/TAF13/TBP architecture. Mutated amino acids are drawn in van-der-Waals representation. TAF11 and TAF13 are shown as ribbons colored in blue and magenta, respectively. TBP is shown in surface representation colored in green. Amino acids mutated in Mutant A are at the center of the TAF11/TAF13 interaction surface with TBP, while amino acids mutated in Mutant B are more in the periphery of the interaction. **(B)** Growth curves for the wild-type strain, BY4741, are shown. BY4741 was transformed with empty vector (EV), vector containing the gene for wild-type TAF13 (WT) and the two mutants (A, B), respectively, and grown at 37°C. Polynomial fits for each growth curve are shown as dotted lines. Standard errors of mean (SEM) calculated from three independent experiments are shown as bars.

Figure 6–Figure Supplement 1: TAF1-TAND TATA-box mimicry in *Drosophila* and Yeast

933 Cartoon representation (top row) and electrostatic surface potential (bottom row) of TBP
934 (left, PDB ID: 1TBP), TATA-DNA (PDB ID: 1CDW), and TAF1-TAND from *Drosophila*
935 (PDB ID: 1TBA) and Yeast (PDB ID: 4B0A). The DNA binding surface of TBP is shown
936 (left), as well as the corresponding complementary surfaces of the interactors (right). Red
937 indicates regions with negative charge, blue indicates regions with positive charge. The shape
938 and surface charge similarity between *Drosophila* TAF1-TAND and TATA-DNA is evident
939 ('TATA-box mimicry'), but is less pronounced in TAF1-TAND from yeast. Yeast TAF1-
940 TAND binding is more extensive involving also the convex surface of TBP (Table 5).
941

REFERENCES

Albright SR, Tjian R (2000) TAFs revisited: more data reveal new twists and confirm old ideas. Gene 242: 1-13.

Anandapadamanaban M, Andresen C, Helander S, Ohyama Y, Siponen MI, Lundstrom P, Kokubo T, Ikura M, Moche M, Sunnerhagen M (2013) High-resolution structure of TBP with TAF1 reveals anchoring patterns in transcriptional regulation. Nat Struct Mol Biol 20: 1008-14. doi: 10.1038/nsmb.2611

Auble DT, Hahn S (1993) An ATP-dependent inhibitor of TBP binding to DNA. Genes Dev 7: 844-56.

Berger I, Fitzgerald DJ, Richmond TJ (2004) Baculovirus expression system for heterologous multiprotein complexes. Nat Biotechnol 22: 1583-7. doi: 10.1038/nbt1036

Bieniossek C, Papai G, Schaffitzel C, Garzoni F, Chaillet M, Scheer E, Papadopoulos P, Tora L, Schultz P, Berger I (2013) The architecture of human general transcription factor TFIID core complex. Nature 493: 699-702. doi: 10.1038/nature11791

Birck C, Poch O, Romier C, Ruff M, Mengus G, Lavigne AC, Davidson I, Moras D (1998) Human TAF(II)28 and TAF(II)18 interact through a histone fold encoded by atypical evolutionary conserved motifs also found in the SPT3 family. Cell 94: 239-249. doi: Doi 10.1016/S0092-8674(00)81423-3

Bleichenbacher M, Tan S, Richmond TJ (2003) Novel interactions between the components of human and yeast TFIIA/TBP/DNA complexes. Journal of Molecular Biology 332: 783-793. doi: Doi 10.1016/S0022-2836(03)00887-8

Burley SK, Roeder RG (1998) TATA box mimicry by TFIID: autoinhibition of pol II transcription. Cell 94: 551-3.

Chen ZA, Jawhari A, Fischer L, Buchen C, Tahir S, Kamenski T, Rasmussen M, Lariviere L, Bukowski-Wills JC, Nilges M, Cramer P, Rappsilber J (2010) Architecture of the RNA polymerase II-TFIIF complex revealed by cross-linking and mass spectrometry. EMBO J 29: 717-26. doi: 10.1038/emboj.2009.401

Cianfrocco MA, Kassavetis GA, Grob P, Fang J, Juven-Gershon T, Kadonaga JT, Nogales E (2013) Human TFIID binds to core promoter DNA in a reorganized structural state. Cell 152: 120-31. doi: 10.1016/j.cell.2012.12.005

Cox J, Mann M (2008) MaxQuant enables high peptide identification rates, individualized p.p.b.-range mass accuracies and proteome-wide protein quantification. Nat Biotechnol 26: 1367-72. doi: 10.1038/nbt.1511

de Vries SJ, van Dijk M, Bonvin AM (2010) The HADDOCK web server for data-driven biomolecular docking. Nat Protoc 5: 883-97. doi: 10.1038/nprot.2010.32

Dignam JD, Lebovitz RM, Roeder RG (1983) Accurate transcription initiation by RNA polymerase II in a soluble extract from isolated mammalian nuclei. Nucleic Acids Res 11: 1475-89.

980 Fitzgerald DJ, Berger P, Schaffitzel C, Yamada K, Richmond TJ, Berger I (2006) Protein
981 complex expression by using multigene baculoviral vectors. Nat Methods 3: 1021-1032.
982 doi: Doi 10.1038/Nmeth983

983 Gangloff YG, Romier C, Thuault S, Werten S, Davidson I (2001a) The histone fold is a key
984 structural motif of transcription factor TFIID. Trends Biochem Sci 26: 250-7.

985 Gangloff YG, Sanders SL, Romier C, Kirschner D, Weil PA, Tora L, Davidson I (2001b)
986 Histone folds mediate selective heterodimerization of yeast TAF(II)25 with TFIID
987 components yTAF(II)47 and yTAF(II)65 and with SAGA component ySPT7. Mol Cell
988 Biol 21: 1841-53. doi: 10.1128/MCB.21.5.1841-1853.2001

989 Gegonne A, Tai X, Zhang J, Wu G, Zhu J, Yoshimoto A, Hanson J, Cultraro C, Chen QR,
990 Guinter T, Yang Z, Hathcock K, Singer A, Rodriguez-Canales J, Tessarollo L, Mackem S,
991 Meerzaman D, Buetow K, Singer DS (2012) The general transcription factor TAF7 is
992 essential for embryonic development but not essential for the survival or differentiation of
993 mature T cells. Mol Cell Biol 32: 1984-97. doi: 10.1128/MCB.06305-11

994 Giladi M, Almagor L, van Dijk L, Hiller R, Man P, Forest E, Khananshvili D (2016)
995 Asymmetric Preorganization of Inverted Pair Residues in the Sodium-Calcium Exchanger.
996 Sci Rep 6: 20753. doi: 10.1038/srep20753

997 Goodrich JA, Tjian R (2010) Unexpected roles for core promoter recognition factors in cell-
998 type-specific transcription and gene regulation. Nat Rev Genet 11: 549-58. doi:
999 10.1038/nrg2847

1000 Gupta K, Sari-Ak D, Haffke M, Trowitzsch S, Berger I (2016) Zooming in on Transcription
1001 Preinitiation. J Mol Biol 428: 2581-91. doi: 10.1016/j.jmb.2016.04.003

1002 Haffke M, Viola C, Nie Y, Berger I (2013) Tandem recombineering by SLIC cloning and
1003 Cre-LoxP fusion to generate multigene expression constructs for protein complex
1004 research. Methods Mol Biol 1073:131-40. doi: 10.1007/978-1-62703-625-2_11.

1005 Hampsey M, Reinberg D (1999) RNA polymerase II as a control panel for multiple
1006 coactivator complexes. Curr Opin Genet Dev 9: 132-9. doi: 10.1016/S0959-
1007 437X(99)80020-3

1008 Hernandez H, Robinson CV (2007) Determining the stoichiometry and interactions of
1009 macromolecular assemblies from mass spectrometry. Nat Protoc 2: 715-26. doi:
1010 10.1038/nprot.2007.73

1011 Hoiby T, Zhou H, Mitsiou DJ, Stunnenberg HG (2007) A facelift for the general transcription
1012 factor TFIIA. Biochim Biophys Acta 1769: 429-36. doi: 10.1016/j.bbaexp.2007.04.008

1013 Hsu JY, Juven-Gershon T, Marr MT, 2nd, Wright KJ, Tjian R, Kadonaga JT (2008) TBP,
1014 Mot1, and NC2 establish a regulatory circuit that controls DPE-dependent versus TATA-
1015 dependent transcription. Genes Dev 22: 2353-8. doi: 10.1101/gad.1681808

1016 Incardona MF, Bourenkov GP, Levik K, Pieritz RA, Popov AN, Svensson O (2009) EDNA:
1017 a framework for plugin-based applications applied to X-ray experiment online data
1018 analysis. J Synchrotron Radiat 16: 872-9. doi: 10.1107/S0909049509036681

1019 Ishihama Y, Rappsilber J, Andersen JS, Mann M (2002) Microcolumns with self-assembled
1020 particle frits for proteomics. J Chromatogr A 979: 233-9.

1021 Kabsch W (2010) Xds. Acta Crystallogr D Biol Crystallogr 66: 125-32. doi:
1022 10.1107/S0907444909047337

1023 Kamada K, Shu F, Chen H, Malik S, Stelzer G, Roeder RG, Meisterernst M, Burley SK
1024 (2001) Crystal structure of negative cofactor 2 recognizing the TBP-DNA transcription
1025 complex. Cell 106: 71-81.

1026 Kim JL, Nikolov DB, Burley SK (1993) Co-crystal structure of TBP recognizing the minor
1027 groove of a TATA element. Nature 365: 520-7. doi: 10.1038/365520a0

1028 Kim TH, Barrera LO, Zheng M, Qu C, Singer MA, Richmond TA, Wu Y, Green RD, Ren B
1029 (2005) A high-resolution map of active promoters in the human genome. Nature 436: 876-
1030 80. doi: 10.1038/nature03877

1031 Koehler C, Sauter PF, Wawryszyn M, Girona GE, Gupta K, Landry JJ, Fritz MH, Radic K,
1032 Hoffmann JE, Chen ZA, Zou J, Tan PS, Galik B, Junttila S, Stolt-Bergner P, Pruneri G,
1033 Gyenesei A, Schultz C, Biskup MB, Besir H, Benes V, Rappsilber J, Jechlinger M, Korbel
1034 JO, Berger I, Braese S, Lemke EA (2016) Genetic code expansion for multiprotein
1035 complex engineering. Nat Methods. doi: 10.1038/nmeth.4032

1036 Koster MJ, Snel B, Timmers HT (2015) Genesis of chromatin and transcription dynamics in
1037 the origin of species. Cell 161: 724-36. doi: 10.1016/j.cell.2015.04.033

1038 Kraemer SM, Ranallo RT, Ogg RC, Stargell LA (2001) TFIIA interacts with TFIID via
1039 association with TATA-binding protein and TAF40. Mol Cell Biol 21: 1737-46. doi:
1040 10.1128/MCB.21.5.1737-1746.2001

1041 Lavigne AC, Gangloff YG, Carre L, Mengus G, Birck C, Poch O, Romier C, Moras D,
1042 Davidson I (1999) Synergistic transcriptional activation by TATA-binding protein and
1043 hTAFII28 requires specific amino acids of the hTAFII28 histone fold. Mol Cell Biol 19:
1044 5050-60.

1045 Lemaire M, Collart MA (2000) The TATA-binding protein-associated factor yTafII19p
1046 functionally interacts with components of the global transcriptional regulator Ccr4-Not
1047 complex and physically interacts with the Not5 subunit. J Biol Chem 275(35):26925-34.
1048 doi: 10.1074/jbc.M002701200

1049 Leurent C, Sanders S, Ruhlmann C, Mallouh V, Weil PA, Kirschner DB, Tora L, Schultz P
1050 (2002) Mapping histone fold TAFs within yeast TFIID. EMBO J 21: 3424-33. doi:
1051 10.1093/emboj/cdf342

1052 Li MZ, Elledge SJ (2007) Harnessing homologous recombination in vitro to generate
1053 recombinant DNA via SLIC. Nat Methods 4: 251-6. doi: 10.1038/nmeth1010

1054 Liu D, Ishima R, Tong KI, Bagby S, Kokubo T, Muhandiram DR, Kay LE, Nakatani Y, Ikura
1055 M (1998) Solution structure of a TBP-TAF(II)230 complex: protein mimicry of the minor
1056 groove surface of the TATA box unwound by TBP. Cell 94: 573-83.

1057 Louder RK, He Y, Lopez-Blanco JR, Fang J, Chacon P, Nogales E (2016) Structure of
1058 promoter-bound TFIID and model of human pre-initiation complex assembly. Nature 531:
1059 604-9. doi: 10.1038/nature17394

1060 Maiolica A, Cittaro D, Borsotti D, Sennels L, Ciferri C, Tarricone C, Musacchio A,
1061 Rappsilber J (2007) Structural analysis of multiprotein complexes by cross-linking, mass
1062 spectrometry, and database searching. Mol Cell Proteomics 6: 2200-11. doi:
1063 10.1074/mcp.M700274-MCP200

1064 Mal TK, Masutomi Y, Zheng L, Nakata Y, Ohta H, Nakatani Y, Kokubo T, Ikura M (2004)
1065 Structural and functional characterization on the interaction of yeast TFIID subunit TAF1
1066 with TATA-binding protein. Journal of Molecular Biology 339: 681-93. doi:
1067 10.1016/j.jmb.2004.04.020

1068 Maston GA, Zhu LJ, Chamberlain L, Lin L, Fang M, Green MR (2012) Non-canonical TAF
1069 complexes regulate active promoters in human embryonic stem cells. Elife 1: e00068. doi:
1070 10.7554/eLife.00068

1071 Matangkasombut O, Auty R, Buratowski S (2004) Structure and function of the TFIID
1072 complex. Adv Protein Chem 67: 67-92. doi: 10.1016/S0065-3233(04)67003-3

1073 McCoy AJ (2007) Solving structures of protein complexes by molecular replacement with
1074 Phaser. Acta Crystallogr D Biol Crystallogr 63: 32-41. doi: 10.1107/S0907444906045975

1075 Mohan WS, Jr., Scheer E, Wendling O, Metzger D, Tora L (2003) TAF10 (TAF(II)30) is
1076 necessary for TFIID stability and early embryogenesis in mice. Mol Cell Biol 23: 4307-18.

1077 Muller F, Tora L (2014) Chromatin and DNA sequences in defining promoters for
1078 transcription initiation. Biochim Biophys Acta 1839: 118-28. doi:
1079 10.1016/j.bbagrm.2013.11.003

1080 Muller F, Zaucker A, Tora L (2010) Developmental regulation of transcription initiation:
1081 more than just changing the actors. Curr Opin Genet Dev 20: 533-40. doi:
1082 10.1016/j.gde.2010.06.004

1083 Nikolov DB, Chen H, Halay ED, Hoffman A, Roeder RG, Burley SK (1996) Crystal
1084 structure of a human TATA box-binding protein/TATA element complex. Proc Natl Acad
1085 Sci U S A 93: 4862-7.

1086 Nikolov DB, Hu SH, Lin J, Gasch A, Hoffmann A, Horikoshi M, Chua NH, Roeder RG,
1087 Burley SK (1992) Crystal structure of TFIID TATA-box binding protein. Nature 360: 40-
1088 6. doi: 10.1038/360040a0

1089 Papai G, Tripathi MK, Ruhlmann C, Layer JH, Weil PA, Schultz P (2010) TFIIA and the
1090 transactivator Rap1 cooperate to commit TFIID for transcription initiation. Nature 465:
1091 956-60. doi: 10.1038/nature09080

1092 Pereira LA, Klejman MP, Timmers HT (2003) Roles for BTAF1 and Mot1p in dynamics of
1093 TATA-binding protein and regulation of RNA polymerase II transcription. Gene 315: 1-
1094 13.

Pernot P, Round A, Barrett R, De Maria Antolinos A, Gobbo A, Gordon E, Huet J, Kieffer J, Lentini M, Mattenet M, Morawe C, Mueller-Dieckmann C, Ohlsson S, Schmid W, Surr J, Theveneau P, Zerrad L, McSweeney S (2013) Upgraded ESRF BM29 beamline for SAXS on macromolecules in solution. J Synchrotron Radiat 20: 660-4. doi: 10.1107/S0909049513010431

Petoukhov MV, Franke D, Shkumatov AV, Tria G, Kikhney AG, Gajda M, Gorba C, Mertens HD, Konarev PV, Svergun DI (2012) New developments in the program package for small-angle scattering data analysis. J Appl Crystallogr 45: 342-350. doi: 10.1107/S0021889812007662

Rajabi K, Ashcroft AE, Radford SE (2015) Mass spectrometric methods to analyze the structural organization of macromolecular complexes. Methods 89: 13-21. doi: 10.1016/j.ymeth.2015.03.004

Rappsilber J, Mann M, Ishihama Y (2007) Protocol for micro-purification, enrichment, pre-fractionation and storage of peptides for proteomics using StageTips. Nat Protoc 2: 1896-906. doi: 10.1038/nprot.2007.261

Rhee HS, Pugh BF (2012) Genome-wide structure and organization of eukaryotic pre-initiation complexes. Nature 483: 295-301. doi: 10.1038/nature10799

Robinson MM, Yatherajam G, Ranallo RT, Bric A, Paule MR, Stargell LA (2005) Mapping and functional characterization of the TAF11 interaction with TFIIA. Mol Cell Biol 25: 945-57. doi: 10.1128/MCB.25.3.945-957.2005

Sardiu ME, Cai Y, Jin J, Swanson SK, Conaway RC, Conaway JW, Florens L, Washburn MP (2008) Probabilistic assembly of human protein interaction networks from label-free quantitative proteomics. Proc Natl Acad Sci U S A 105: 1454-9. doi: 10.1073/pnas.0706983105

Schuck P (2000) Size-distribution analysis of macromolecules by sedimentation velocity ultracentrifugation and lamm equation modeling. Biophys J 78: 1606-19. doi: 10.1016/S0006-3495(00)76713-0

Shen WC, Bhaumik SR, Causton HC, Simon I, Zhu X, Jennings EG, Wang TH, Young RA, Green MR (2003) Systematic analysis of essential yeast TAFs in genome-wide transcription and preinitiation complex assembly. EMBO J 22(13):3395-402. doi: 10.1093/emboj/cdg336

Thomas MC, Chiang CM (2006) The general transcription machinery and general cofactors. Crit Rev Biochem Mol Biol 41: 105-78. doi: 10.1080/10409230600648736

Tora L, Timmers HT (2010) The TATA box regulates TATA-binding protein (TBP) dynamics in vivo. Trends Biochem Sci 35: 309-14. doi: 10.1016/j.tibs.2010.01.007

Trowitzsch S, Viola C, Scheer E, Conic S, Chavant V, Fournier M, Papai G, Ebong IO, Schaffitzel C, Zou J, Haffke M, Rappsilber J, Robinson CV, Schultz P, Tora L, Berger I (2015) Cytoplasmic TAF2-TAF8-TAF10 complex provides evidence for nuclear holo-TFIID assembly from preformed submodules. Nat Commun 6: 6011. doi: 10.1038/ncomms7011

1135 Vermeulen M, Mulder KW, Denissov S, Pijnappel WW, van Schaik FM, Varier RA,
1136 Baltissen MP, Stunnenberg HG, Mann M, Timmers HT (2007) Selective anchoring of
1137 TFIID to nucleosomes by trimethylation of histone H3 lysine 4. Cell 131: 58-69. doi:
1138 10.1016/j.cell.2007.08.016

1139 Verrijzer CP, Chen JL, Yokomori K, Tjian R (1995) Binding of TAFs to core elements
1140 directs promoter selectivity by RNA polymerase II. Cell 81: 1115-25.

1141 Webb B, Sali A (2014) Comparative Protein Structure Modeling Using MODELLER.
1142 Current Protocols in Bioinformatics. John Wiley & Sons, Inc.

1143 Werten S, Mitschler A, Romier C, Gangloff YG, Thuault S, Davidson I, Moras D (2002)
1144 Crystal structure of a subcomplex of human transcription factor TFIID formed by TATA
1145 binding protein-associated factors hTAF4 (hTAF(II)135) and hTAF12 (hTAF(II)20). J
1146 Biol Chem 277: 45502-9. doi: 10.1074/jbc.M206587200

1147 Winn MD, Ballard CC, Cowtan KD, Dodson EJ, Emsley P, Evans PR, Keegan RM, Krissinel
1148 EB, Leslie AG, McCoy A, McNicholas SJ, Murshudov GN, Pannu NS, Potterton EA,
1149 Powell HR, Read RJ, Vagin A, Wilson KS (2011) Overview of the CCP4 suite and current
1150 developments. Acta Crystallogr D Biol Crystallogr 67: 235-42. doi:
1151 10.1107/S0907444910045749

1152 Warfield L, Ramachandran S, Baptista T, Tora L, Devys D, Hahn S Transcription of nearly
1153 all yeast RNA Polymerase II-transcribed genes is dependent on transcription factor TFIID
1154 (submitted 2017).

1155 Wollmann P, Cui S, Viswanathan R, Berninghausen O, Wells MN, Moldt M, Witte G,
1156 Butryn A, Wendler P, Beckmann R, Auble DT, Hopfner KP (2011) Structure and
1157 mechanism of the Swi2/Snf2 remodeller Mot1 in complex with its substrate TBP. Nature
1158 475: 403-7. doi: 10.1038/nature10215

1159 Wright KJ, Marr MT, 2nd, Tjian R (2006) TAF4 nucleates a core subcomplex of TFIID and
1160 mediates activated transcription from a TATA-less promoter. Proc Natl Acad Sci U S A
1161 103: 12347-52. doi: 10.1073/pnas.0605499103

1162 Xie J, Collart M, Lemaire M, Stelzer G, Meisterernst M (2000) A single point mutation in
1163 TFIIA suppresses NC2 requirement in vivo. EMBO J 19: 672-82. doi:
1164 10.1093/emboj/19.4.672

1165 Xie X, Kokubo T, Cohen SL, Mirza UA, Hoffmann A, Chait BT, Roeder RG, Nakatani Y,
1166 Burley SK (1996) Structural similarity between TAFs and the heterotetrameric core of the
1167 histone octamer. Nature 380: 316-22. doi: 10.1038/380316a0

1168 Zybailov BL, Florens L, Washburn MP (2007) Quantitative shotgun proteomics using a
1169 protease with broad specificity and normalized spectral abundance factors. Mol Biosyst 3:
1170 354-60. doi: 10.1039/b701483j

1171

Accession codes. Atomic coordinates and structure factors have been deposited in the Protein Data Bank (PDB ID 5M4S). Proteomics data has been submitted to PRIDE (Accession number PXD005676)

ACKNOWLEDGEMENT

We thank all members of our laboratories for helpful discussions. We are grateful to Max Nanao, Aurelien Deniaud, Christian Becke, Moreno Wichert and Timothy J. Richmond for valuable contributions. Frederic Garzoni is acknowledged for MultiBac expressions. We thank Steven Hahn for kindly providing the Taf13-AID strain and Tony Weil for providing yeast TAF constructs and antibody reagents. This work was supported by the Agence Nationale de Recherche (ANR, France) DiscoverIID (to LT and IB). CK and EAL are funded by the Baden-Wuerttemberg Stiftung (Germany). This work was supported by the Wellcome Trust (Senior Research Fellowship to JR: 103139, Centre core grant: 092076, instrument grant: 108504). LT is recipient of a European Research Council (ERC) Advanced grant (Birtoaction). IB is recipient of a Senior Investigator Award from the Wellcome Trust. This research received support from BrisSynBio, a BBSRC/EPSRC Research Centre for synthetic biology at the University of Bristol (BB/L01386X/1).

AUTHOR CONTRIBUTIONS

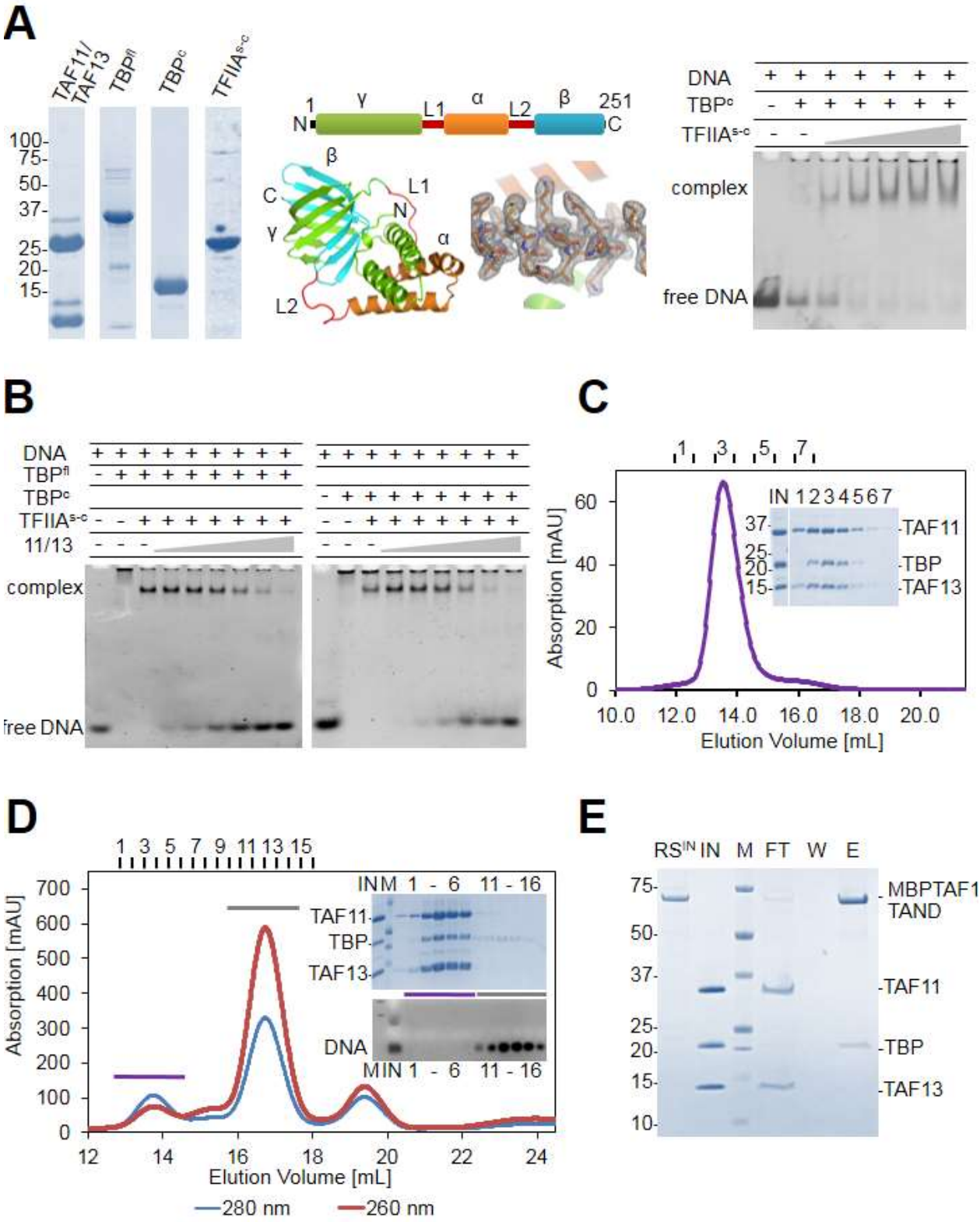
IB and LT conceived this research. KG prepared samples and carried out experiments. All other authors provided critical analytical tools, expertise or reagents. KG, LT and IB co-wrote the manuscript with input from all authors.

COMPETING FINANCIAL INTEREST STATEMENT

1197 The authors declare no competing financial interest.

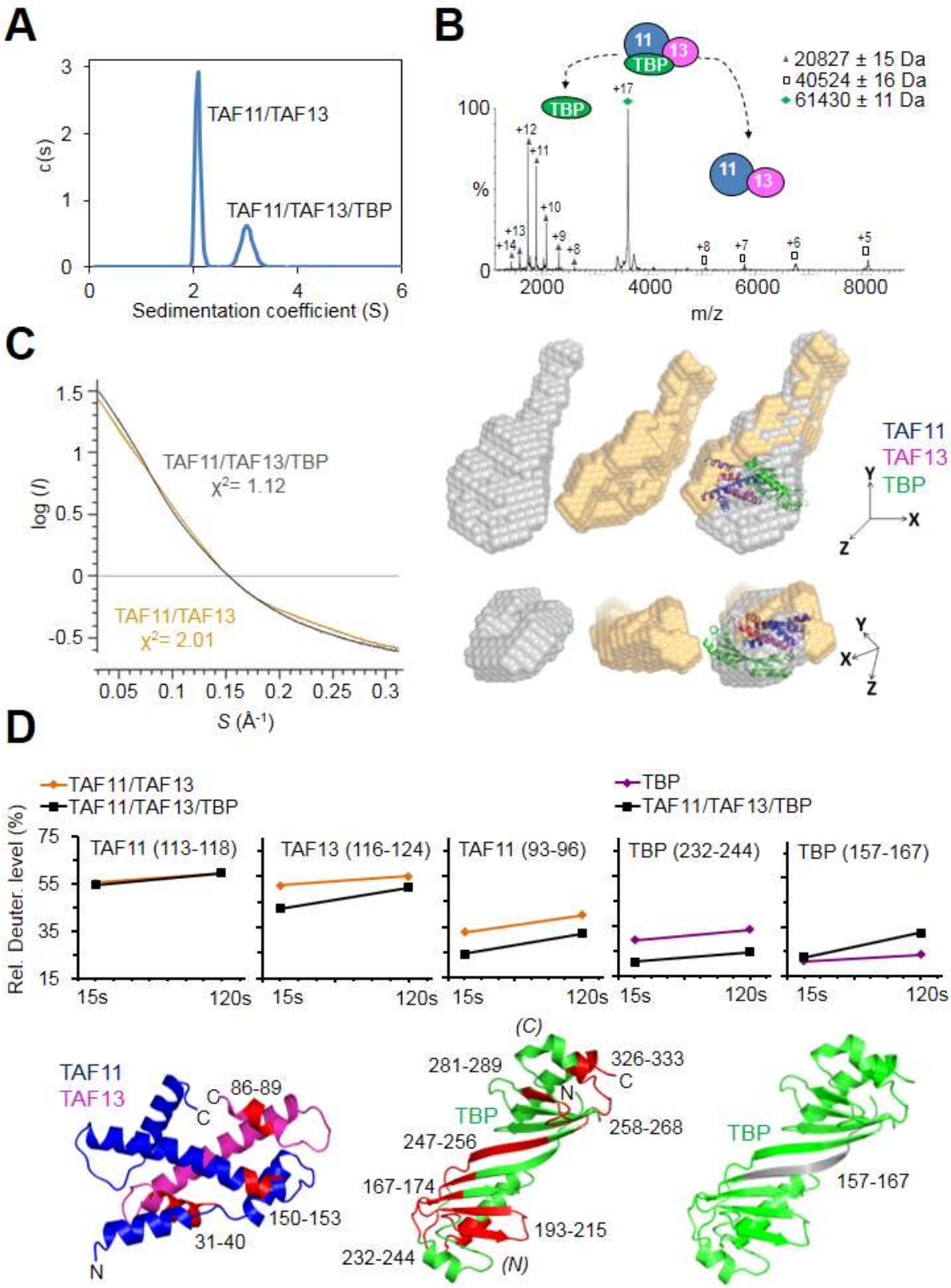
1198 **FIGURES**

1199 **Figure 1**



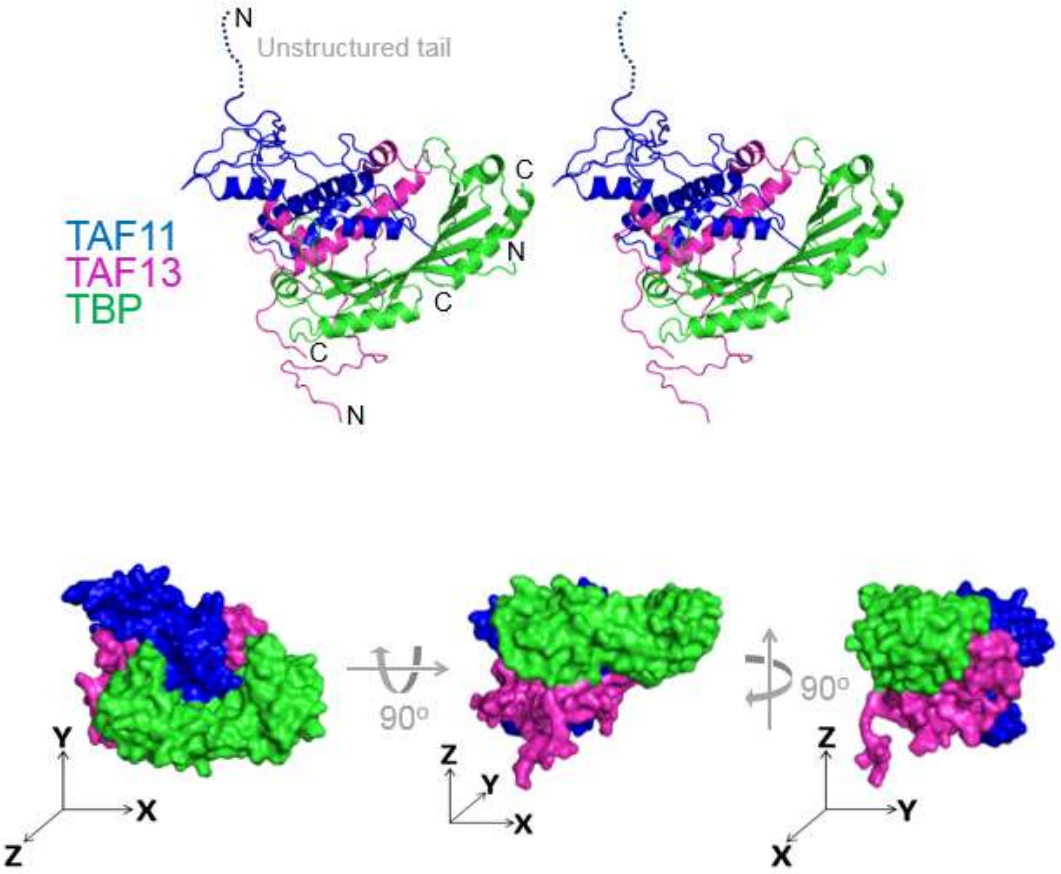
1200
1201

1202 **Figure 2**



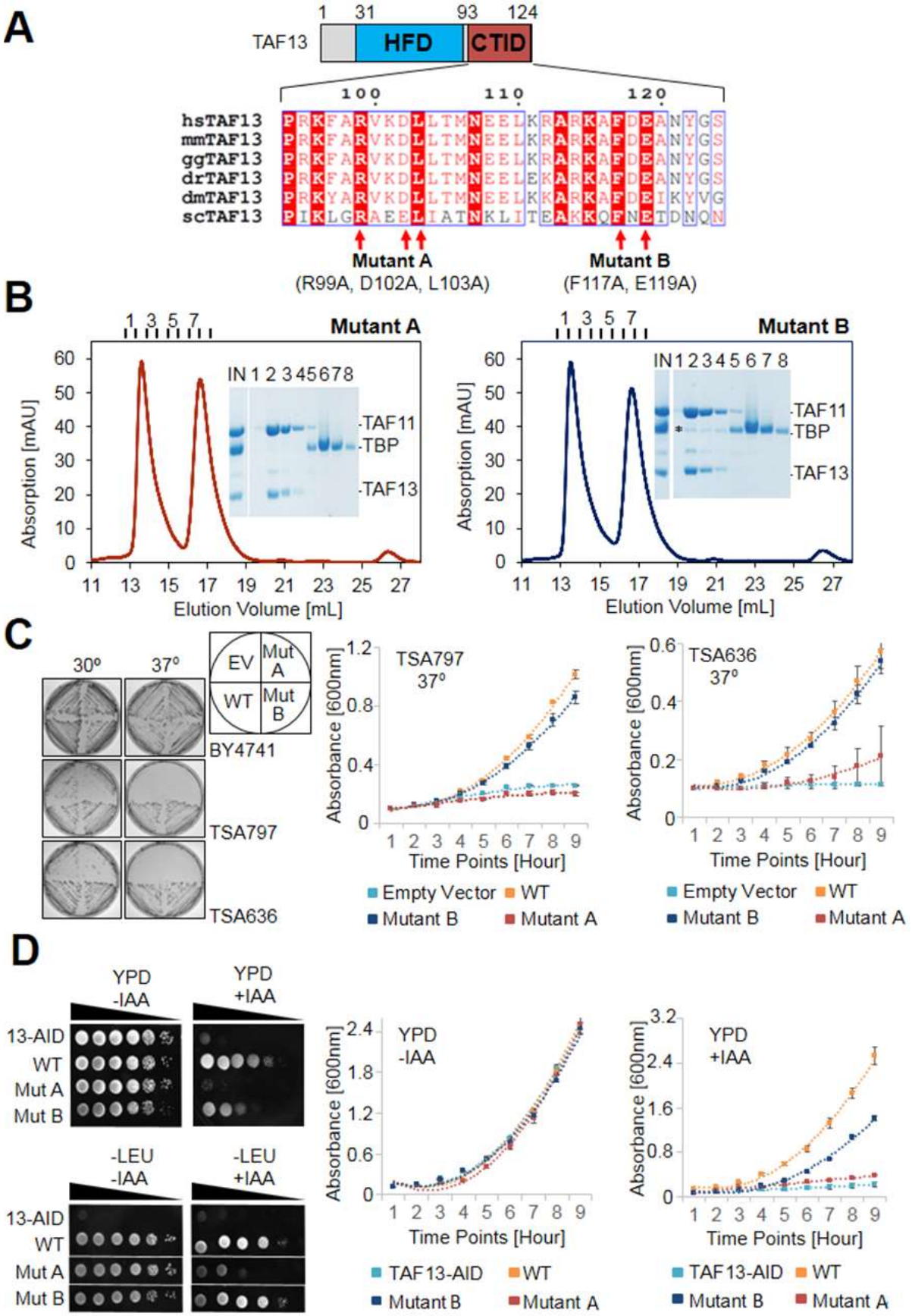
1203
1204

1205 **Figure 3**



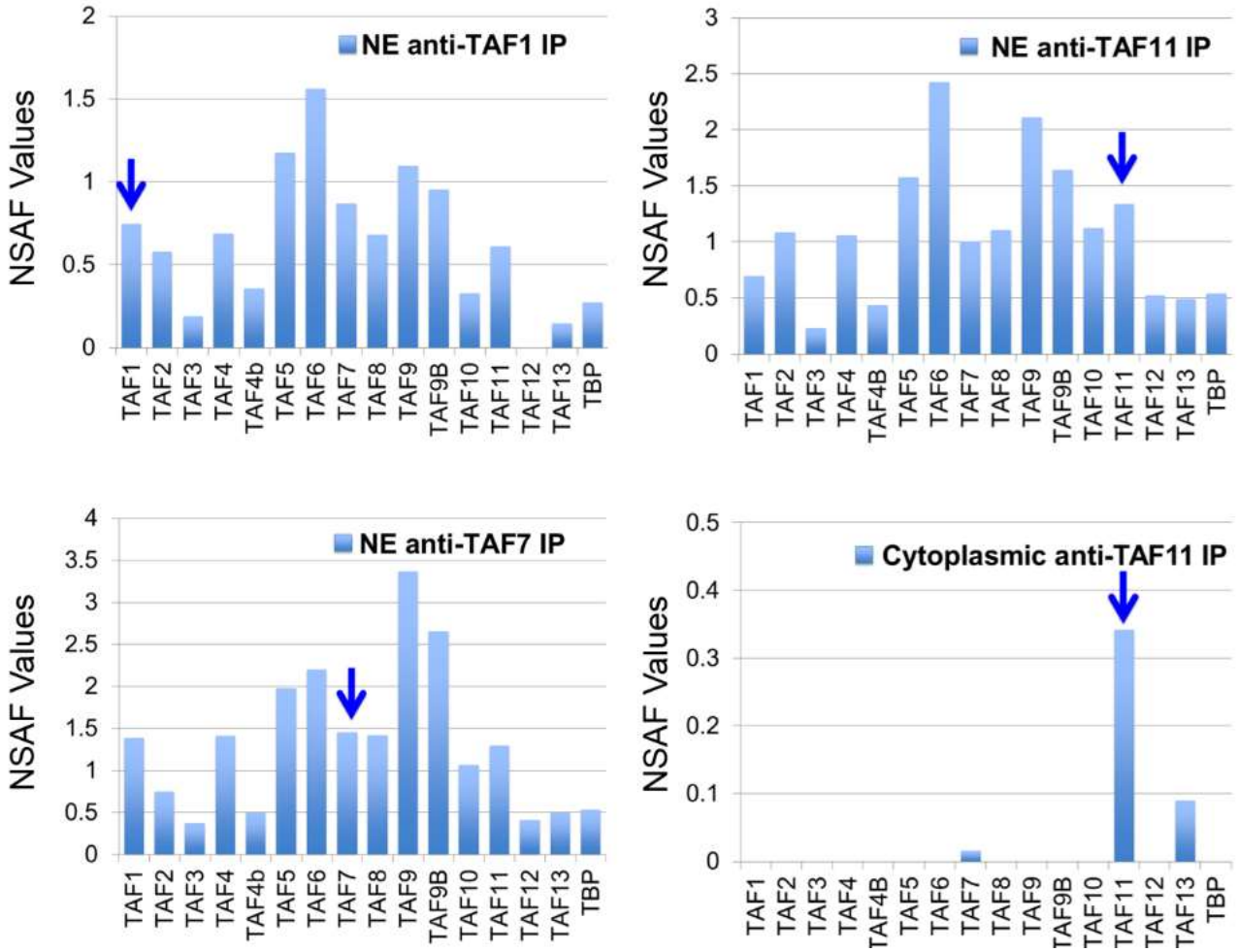
1206

1207 **Figure 4**

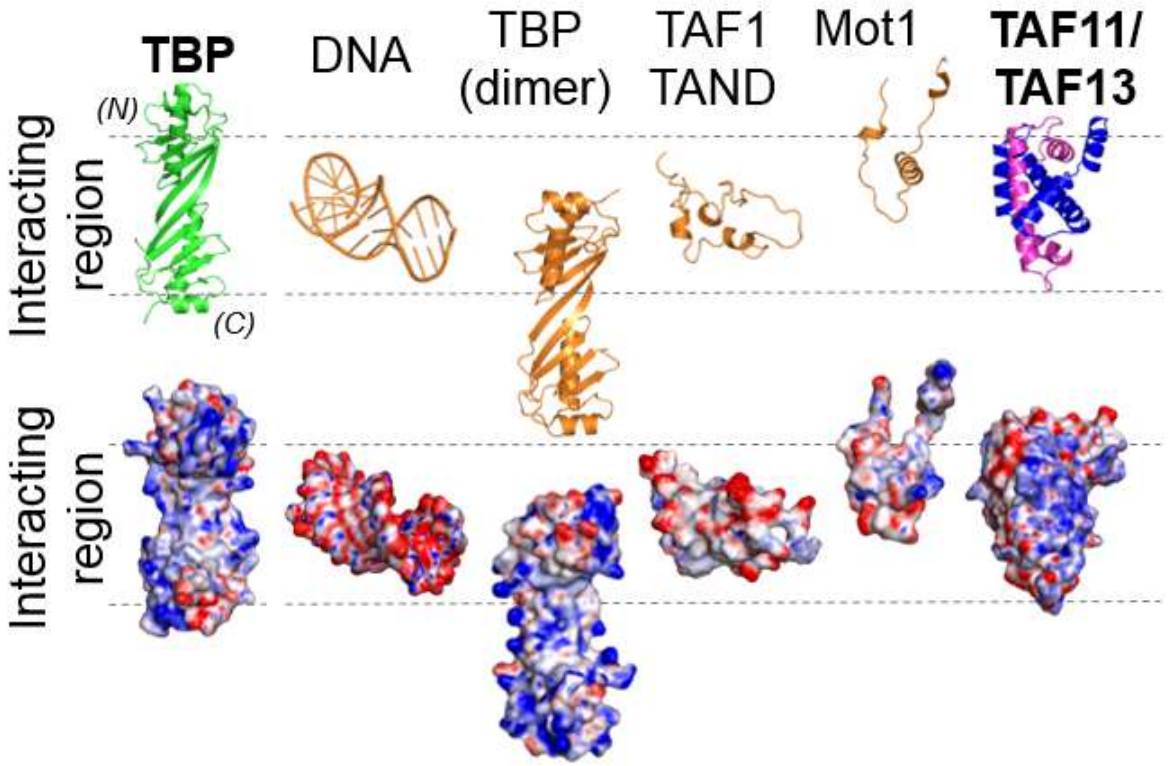


1208

Figure 5

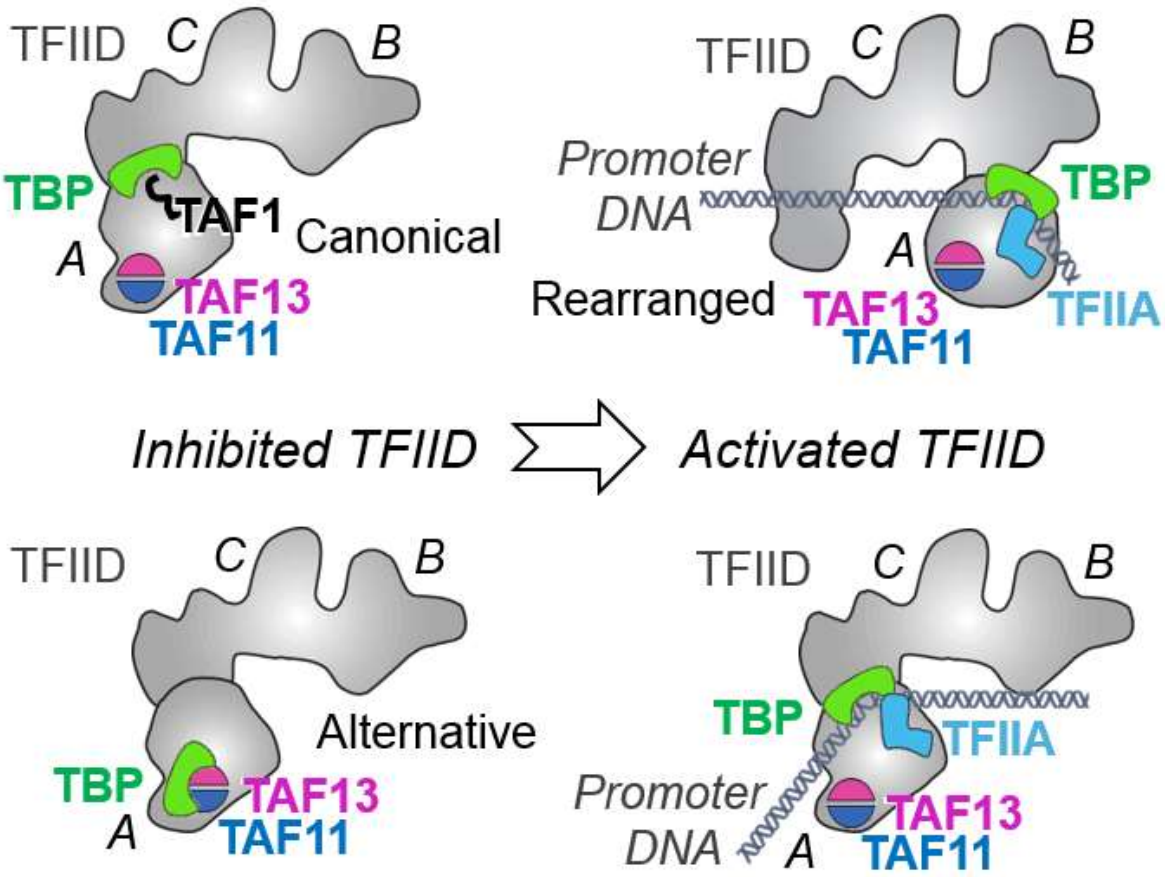


1211 **Figure 6**

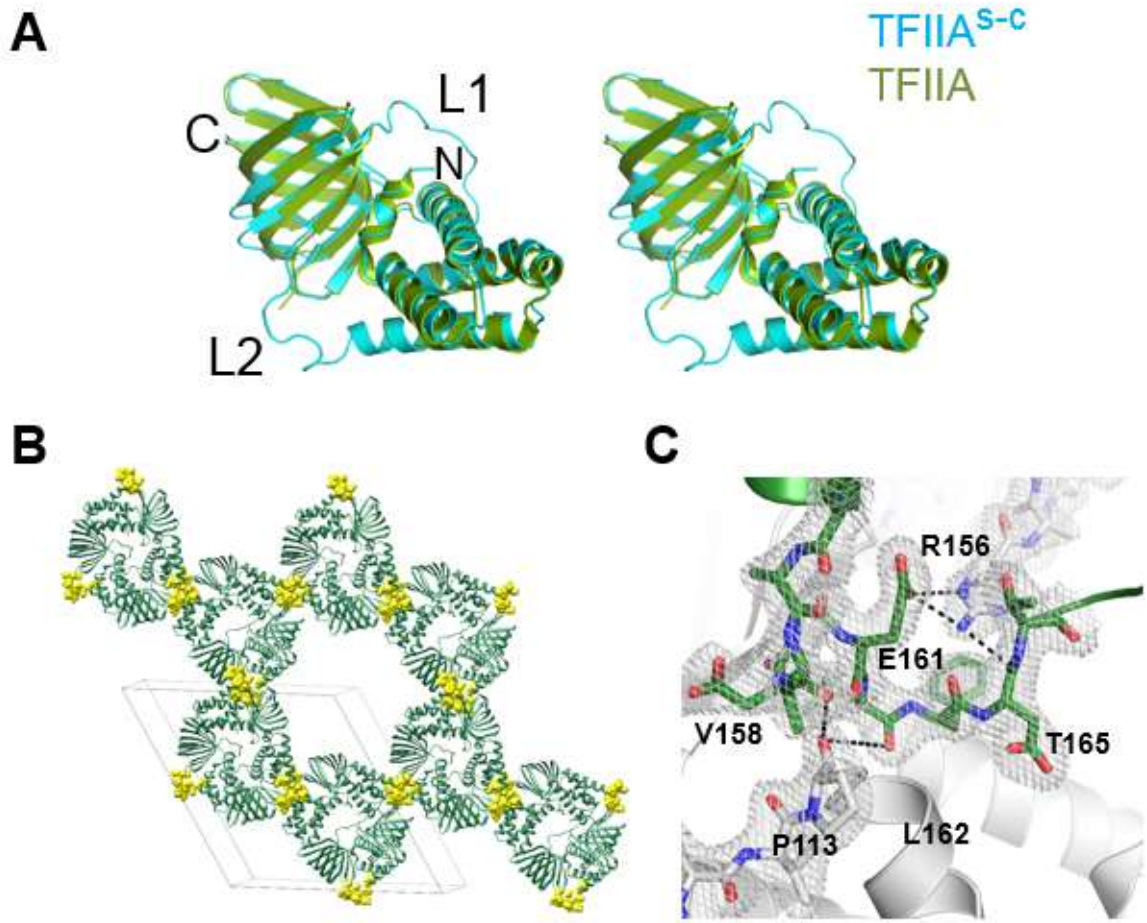


1212

Figure 7



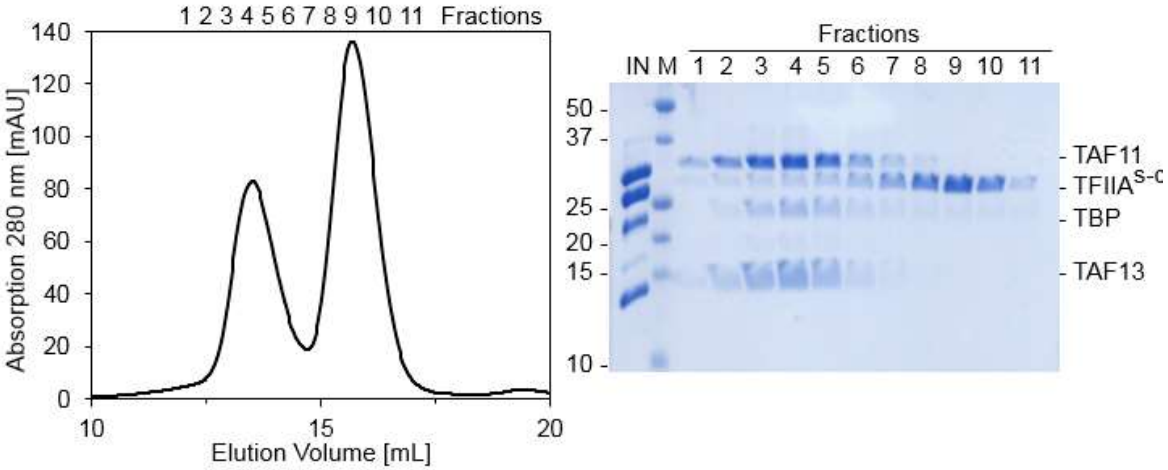
1216 **Figure 1–Figure Supplement 1**



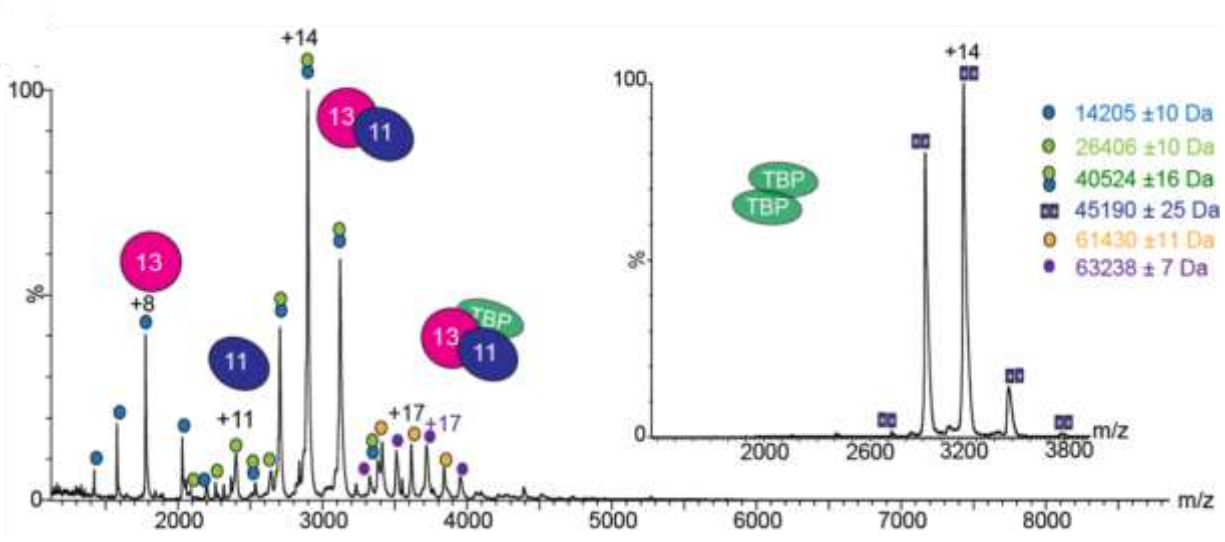
1217

1218

Figure 1–Figure Supplement 2



1222 **Figure 2–Figure Supplement 1**



1223

1224

Figure 2–Figure Supplement 2

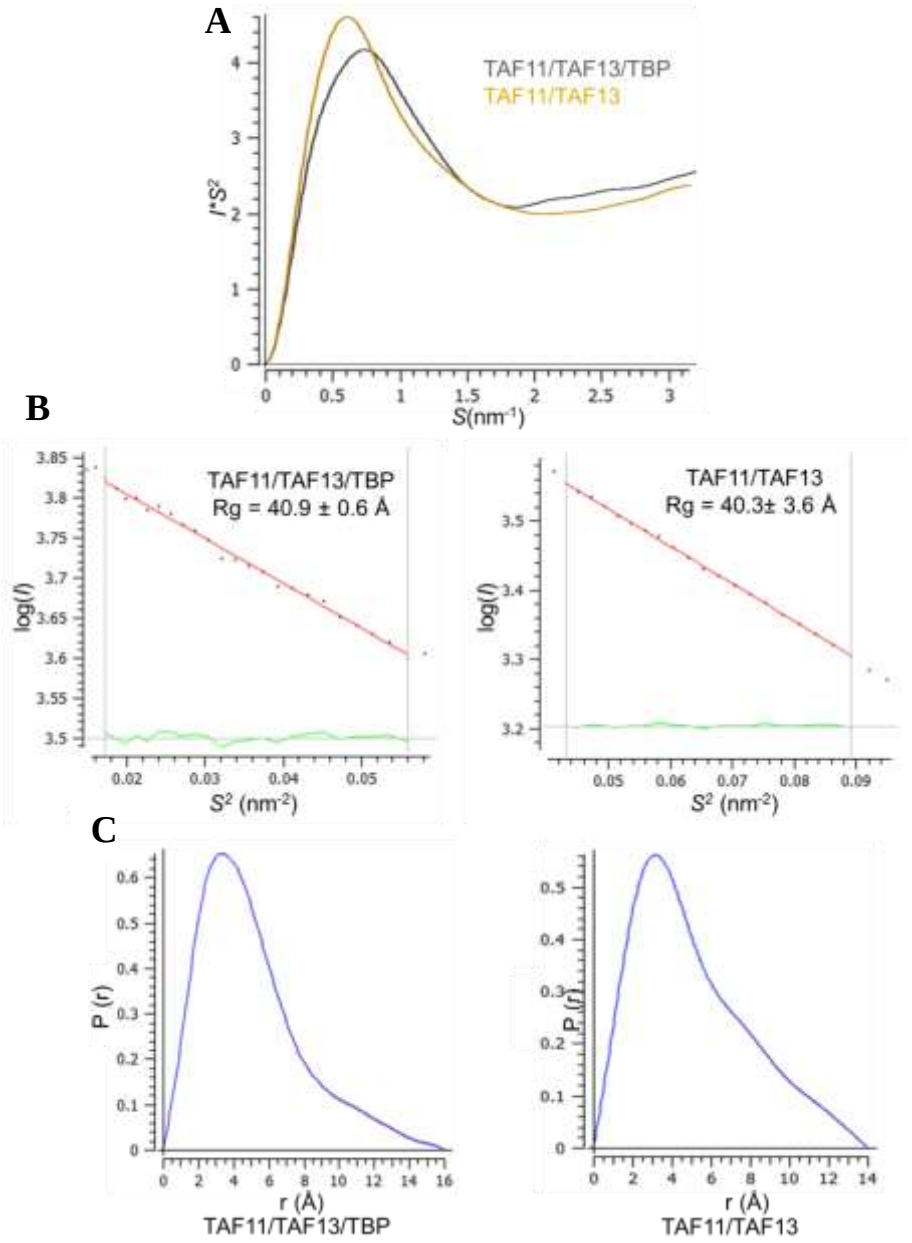


Figure 3–Figure Supplement 1

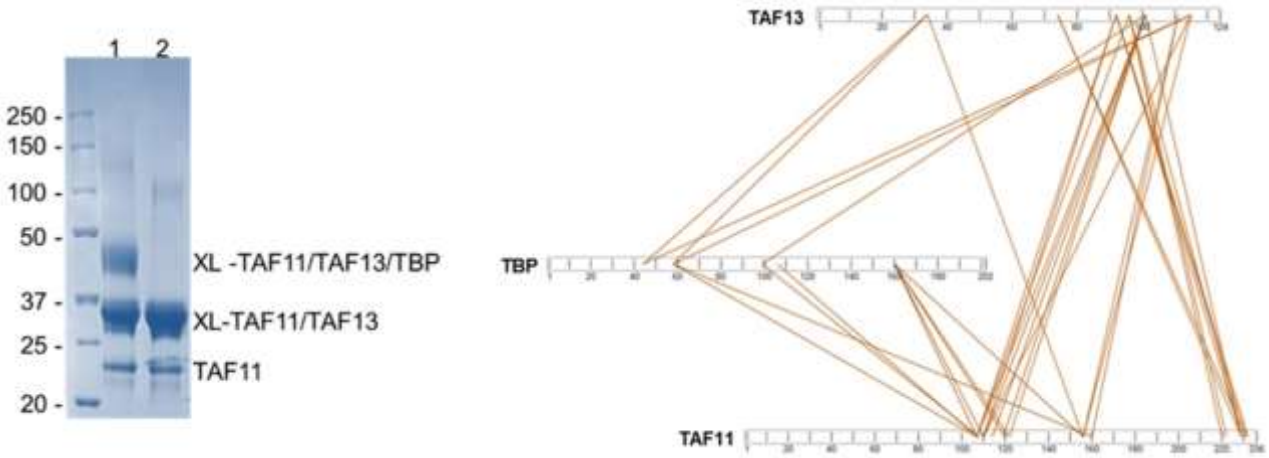


Figure 3–Figure Supplement 2

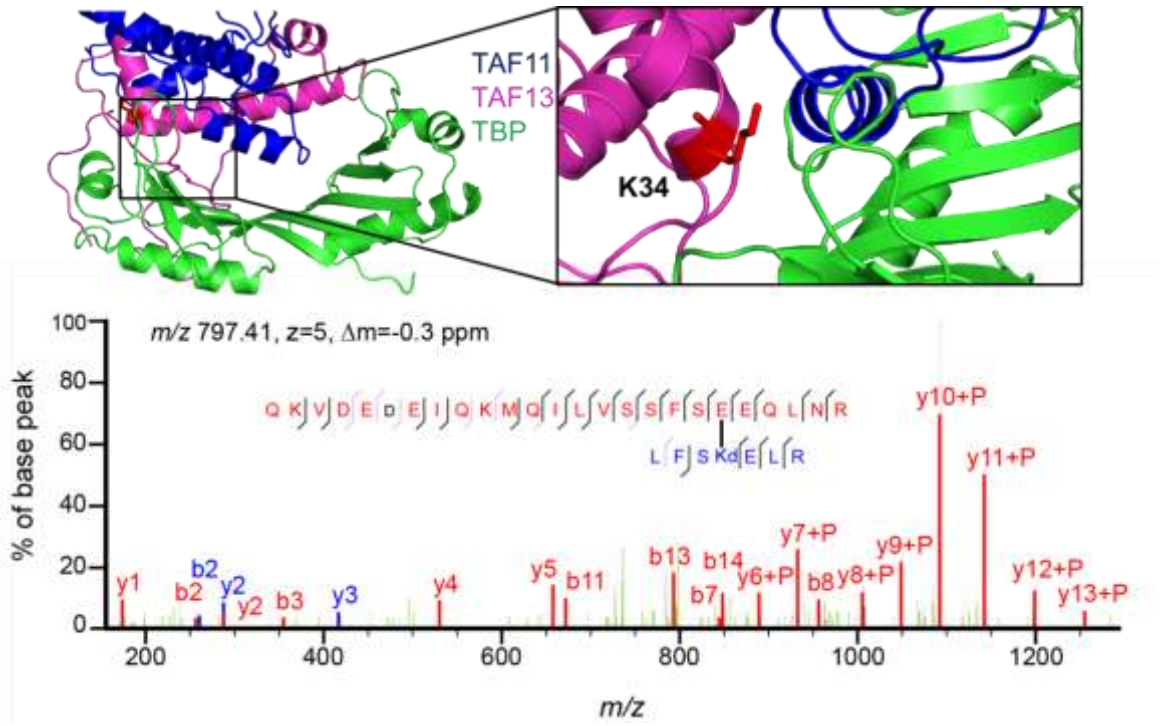


Figure 3–Figure Supplement 3

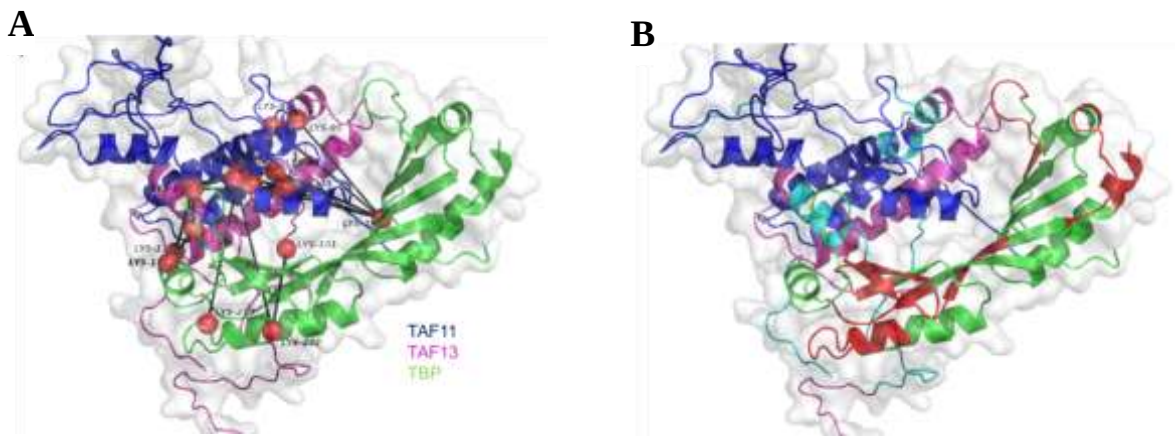


Figure 4–Figure Supplement 1

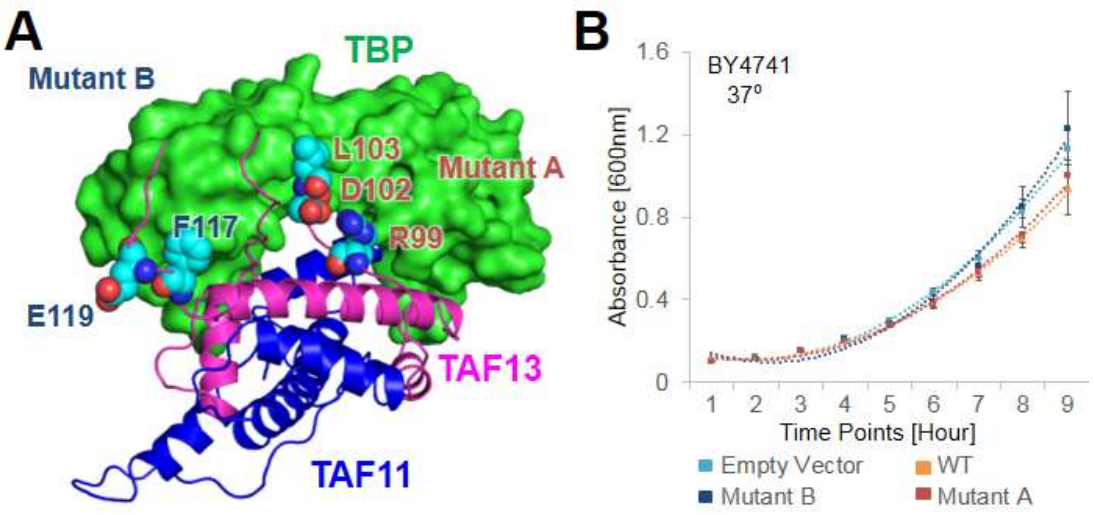


Figure 6–Figure Supplement 1

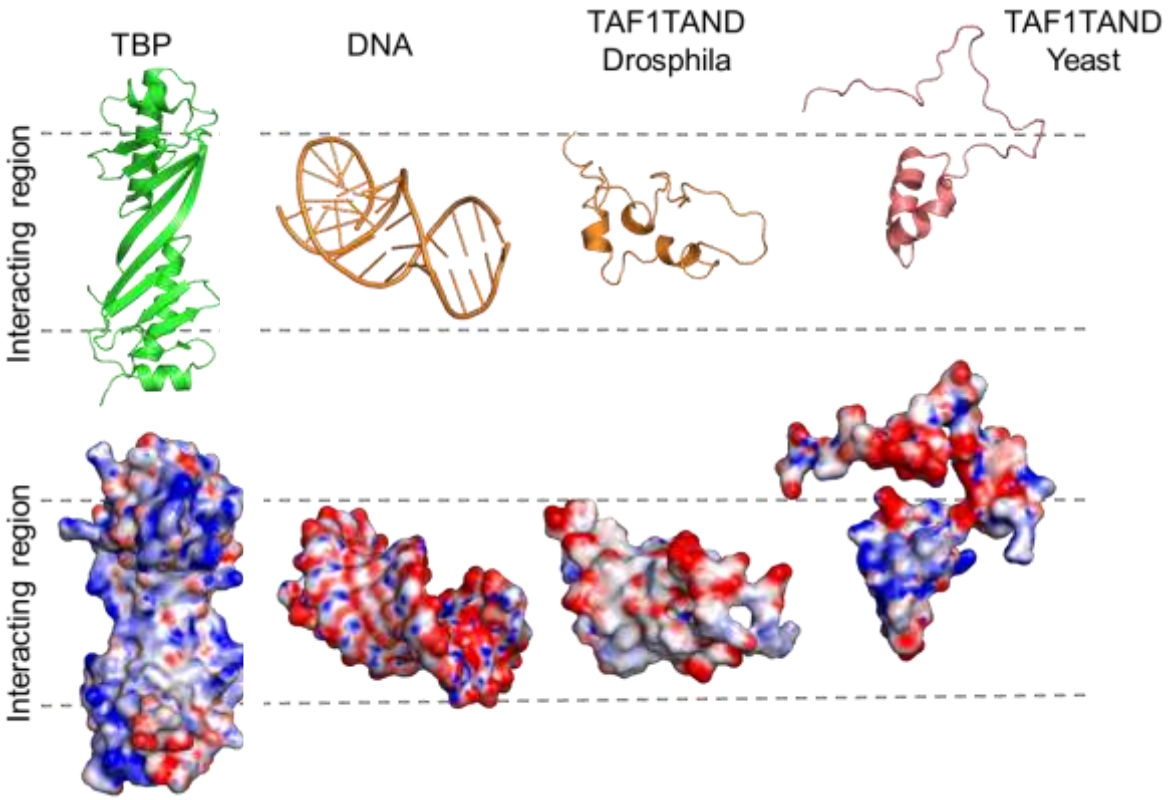


Table 1.

X-ray data collection and refinement statistics

	TFIIA^{s-c}
Data collection	
Space group	P65
Cell dimensions	
a, *b*, *c* (Å)	123.3, 123.3, 34.8
α, β, γ (°)	90, 90, 120
Resolution (Å)	53.4-2.4
Last resolution bin (Å)	2.52-2.38
$R_{measure}$ (%)	12.9 (64.8)
$I / \sigma I$	11.5 (2.72)
Completeness (%)	99.8 (99.9)
Multiplicity	6.7 (6.8)
Refinement	
Resolution (Å)	40.36-2.38 (2.44-2.38)
No. reflections	
Work set	11859
Free set	601
R_{work}	0.18 (0.27)
R_{free}	0.24 (0.36)
No. atoms	
Protein	1689
Water	50
r.m.s deviations	
Bond lengths (Å)	0.0223
Bond angles (°)	2.088

**Values in parentheses are for highest resolution shell.*

1253 **Table 2.**

1254 Data collection and refinement statistics SAXS

	TAF11/TAF13/TBP	TAF11/TAF13
Data collection parameters		
Beamline	ESRF-BM29	ESRF-BM29
Beam size at sample	~700 µm x 700 µm	~700 µm x 700 µm
Wavelength (Å)	0.931	0.931
S range (Å-1)	0.003-0.497	0.003-0.497
Concentration range (mg ml-1)	0.3-7.11	0.53-7.48
Temperature (°C)	4	4
Beamline	ESRF-BM29	ESRF-BM29
Beam size at sample	~700 µm x 700 µm	~700 µm x 700 µm
Wavelength (Å)	0.931	0.931
Structural parameters†		
I(0) (arbitrary units) [from P(r)]	49.63	43.65
Rg (Å) [from P(r)]	41	41.2
I(0) (arbitrary units) (from Guinier)	50.21 ± 0.33	44.15 ± 0.08
Rg (Å) (from Guinier)	40.9 ± 0.6	40.3± 3.6
Dmax (Å)	160	140
Porod volume estimate (Å3)	120110	89850
Molecular mass Mr [from porod volume]	70.65 kDa	52.91 kDa
I(0) (arbitrary units) [from P(r)]	49.63	43.65
Rg (Å) [from P(r)]	41	41.2
I(0) (arbitrary units) (from Guinier)	50.21 ± 0.33	44.15 ± 0.08
Rg (Å) (from Guinier)	40.9 ± 0.6	40.3± 3.6
Dmax (Å)	160	140
Porod volume estimate (Å3)	120110	89850
Molecular mass Mr [from porod volume]	70.65 kDa	52.91 kDa

1255 † *Reported for experimental merged data.*

1256

Table 3.

Peptide deuteration level changes upon complex formation in HDX-MS

TAF11			
Residues Numbers	*Sequence*	*Δ%D (15 Sec)†*	*Δ%D (120 Sec)†*
93-96	EKKQ	-8.465	-7.309
105-109	KMQIL	-7.014	
150-153	VVIA		-7.295

TAF13			
Residues Numbers	*Sequence*	*Δ%D (15 Sec)†*	*Δ%D (120 Sec)†*
14-31	NEEIGGGAEGGQGKRKRL	-7.507	-7.192
32-35	FSKE	-8.794	
36-40	LRCMM	-7.0	
86-88	IVF	-8.598	
86-89	IVFL	-10.366	
97-104	FARVKDLL		-7.223
116-124	AFDEANYGS	10.89	8.721

TBP			
Residues Numbers	*Sequence*	*Δ%D (15 Sec)†*	*Δ%D (120 Sec)†*
157-167*	IVPQLQNIVST		9.038
167-174	TVNLGCKL	-9.675	-10.171
193-197	FAAVI	-14.068	-12.267
197-208	IMRIREPRTTAL	-7.785	
199-208	RIREPRTTAL	-9.661	
209-215	IFSSGKM	-8.655	-8.75
232-244	KYARVVQKLGFPA	-21.337	-18.382
233-242	YARVVQKLGF	-7.128	
247-252	LDFKIQ		-7.271
250-256	KIQNMVG	-8.913	
259-266	DVKFPIRL	-11.932	-10.109
259-268	DVKFPIRLEG	-13.79	-11.479
260-266	VKFPIRL	-10.393	
260-268	VKFPIRLEG	-16.066	-13.002
281-287	PELFPGL		-7.811
285-289	PGLIY	-8.061	-7.339
326-335	PILKGFRKTT	-7.077	-7.902

† *Peptides exhibiting changes in deuteration level ≥ 7% are shown.*

* *TBP peptide with increasing deuteration level upon complex formation.*

1262 **Table 4.** Cross-links observed by BS3 CLMS

TAF11-TBP			
TAF11 Residue	*TBP Residue*	*No. of matches*	*Highest Score*
K122	K160	5	9.416
K110	K160	2	8.818
K107	K106	2	8.314
K160	K58	1	8.085
K107	K160	2	7.926
K107	K58	1	7.684
K156	K160	1	6.928
K108	K99	1	6.124
K120	K160	1	4.145
(9 unique links)			

TAF13-TBP			
TAF13 Residue	*TBP Residue*	*No. of matches*	*Highest Score*
K34	K44	1	10.578
K34	K58	1	7.856
K101	K99	1	7.584
K115	K58	1	7.519
K115	K44	1	5.742
(5 unique links)			

TAF11-TAF13			
TAF11 Residue	*TAF13 Residue*	*No. of matches*	*Highest Score*
K220	K96	2	12.795
K110	K101	2	12.014
K229	K96	3	11.053
K114	K101	5	10.98
K229	K101	6	10.909
K156	K34	3	9.627
K107	K92	2	9.321
K222	K96	2	9.308
K156	K111	10	9.226
K160	K111	2	8.371
K156	K115	1	8.326
K232	K92	1	8.292
K107	K96	1	7.981
K110	K115	1	7.943
K119	K101	2	7.934
K122	K101	1	7.867
K229	K92	4	6.872
K110	K92	1	6.817
K107	K101	1	6.437
K231	K92	1	5.77
K229	S74	1	3.42
K231	S74	1	3.42
(22 unique links)			

1263

Table 5.

Interaction surfaces in TBP complexes

Interactor	*Interface (Å^2)*
TBP Dimer	3010.2
DNA	4020.1
TAF1-TAND *(D. melanogaster)*	3287.8
TAF1-TAND (Yeast)	7483.1†
Mot1 (*E. cuniculi*)	4300.0
TAF11/TAF13	3305.2

Calculated with PyMol v1.8.2.0 (www.pymol.org)

† *includes TAND1 and TAND2*