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

Kapil Gupta1,2, Aleksandra A. Watson3, Tiago Baptista4-7, Elisabeth Scheer4-7, Anna L. Chambers1, Christine Koehler8, Juan Zou9, Ima Obong-Ebong10, Eaazhisai Kandiah2,11, Arturo Temblador2, Adam Round2, Eric Forest11, Petr Man12, Christoph Bieniossek2, Ernest D. Laue3, Edward A. Lemke8, Juri Rappsilber9, Carol V. Robinson10, Didier Devys4-7, Laszlo Tora4,7* and Imre Berger1*

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
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 TFIIID 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 the 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...
formation of a stable ternary complex formed by TAF11/TAF13 and TBP. We analysed the TAF11/TAF13/TBP complex biochemically and structurally utilizing a comprehensive, integrative approach. We report a novel interaction of the TAF11/TAF13 dimer binding to the concave DNA-binding groove of TBP, thus excluding TATA-box containing DNA. Using pull-down experiments with immobilized TAF1-TAND, we demonstrate competition between TAF11/TAF13 and TAF1-TAND for TBP binding. We identify a novel C-terminal TBP-binding domain (CTID) within TAF13 which is highly conserved from yeast to man. We probe key residues within this TAF13 CTID by mutagenesis in vitro and in vivo in cell growth experiments, revealing a key role of this domain for viability. We contrast the TAF11/TAF13 interaction with other TBP DNA-binding groove interactors including Mot1 and discuss the implications of our findings in the context of TFIID assembly. Based on our results, we propose a novel, functional state of TFIID in transcription regulation.
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, TFIIA αβ and TFIIA γ, with TFIIA αβ 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-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-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-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-c (Figure 1A, Figure 1–Figure Supplement 1, Table 1). The crystal structure revealed a virtually identical conformation of unliganded TFIIA s-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-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\textsuperscript{s-c}, TBP and AdMLP TATA-DNA. We observed that TFIIA\textsuperscript{s-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\textsuperscript{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 results provide 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-3-acetic 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
TAF11/TAF13/TBP complex would likewise represent such a sub-assembly. To this end we performed co-immunoprecipitations (co-IPs) from HeLa cell cytosolic and nuclear extracts using an anti-TAF11 antibody (Figure 5). We found dimeric TAF11/TAF13 complex in the cytosol representing the complete HF pair. We could not detect TBP in cytosolic co-IPs, however, our experiments evidenced TAF7 association with cytoplasmic TAF11/TAF13. The anti-TAF11 co-IP from nuclear extract, in contrast, contained all TFIID components. Surprisingly, normalized spectral abundance factor (NSAF) (Zybailov et al. 2007) analyses of several distinct anti-TFIID IPs (i.e. anti-TAF1 and anti-TAF7) from NE indicate that less than half of the nuclear TFIID specimens appear to contain stably bound TBP, implying considerable dynamics in TBP association with TFIID in the nucleus, possibly regulated by TBP containing TFIID submodules such as TAF11/TAF13/TBP.
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
regulated manner in the nucleus from preformed submodules (Bieniossek et al. 2013; Gupta et al. 2016; Trowitzsch et al. 2015). We analyzed TAF11-containing complexes by co-IP experiments from the cytoplasm and the nucleus of HeLa cells. In cytosolic co-IPs, we identified the TAF11/TAF13 histone fold pair, suggesting that this TFIID submodule may also be preformed in the cytoplasm (Figure 5). Note however that this cytoplasmic TAF11/TAF13 building block did not contain detectable amounts of TBP in our experiments, suggesting that trimeric TAF11/TAF13/TBP, within TFIID or as a discrete TFIID submodule, would be formed in the nucleus. Interestingly, we also identified TAF7 associated with TAF11/TAF13 in the cytoplasm, hinting at putative novel interactions between TAF11/TAF13 and TAF7. TAF1 forms stable complexes with both TAF7 and TBP (Gupta et al. 2016). We speculate that cytoplasmic TAF7/TAF11/TAF13 may represent an assembly intermediate towards a TAF1/TAF7/TAF11/TAF13/TBP module, which may integrate into a core-TFIID and TAF2/TAF8/TAF10 containing “8TAF” assembly (Trowitzsch et al. 2015) in the formation pathway to the complete nuclear holo-TFIID complex. In this TAF1/TAF7/TAF11/TAF13/TBP module, TBP would be tightly bound to either TAF1 or TAF11/TAF13, which could serve to ascertain that this putative TFIID submodule is efficiently blocked from any potentially detrimental interactions with DNA until holo-TFIID formation is completed.

In the nucleus, IPs utilizing antibodies against several different TFIID specific TAFs co-precipitated all known TFIID subunits, although with variable stoichiometry. Strikingly, stoichiometry analyses carried out by NSAF calculations of our nuclear anti-TAF IPs indicated that TBP was only present in less than half of the TFIID specimens, when compared to TAF1 or TAF7 for example, suggesting that TFIID-type complexes may exist which do not contain TBP. TBP is thought to be highly mobile structurally in the context of holo-TFIID, with recent cryo-EM studies predicting large-scale migration of TBP within
distinct TFIID conformational states (Cianfrocco et al. 2013; Louder et al. 2016). Our observations point to an additional level of compositional TBP and/or TAF dynamics in TFIID formation, raising the interesting possibility that the accretion of TBP in holo-TFIID may be regulated by partially assembled nuclear TFIID building blocks including for instance the TAF11/TAF13/TBP complex we identified.
**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\(^\text{\beta}\), TBP\(^\text{\gamma}\) and TFIIA\(^{\text{\alpha\beta\gamma}}\) 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\(^{\text{\alpha}}\) (AA2-59; Uniprot accession number P52655), \(\beta\) (AA302-376; Uniprot accession number P52655) and \(\gamma\) (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-
cleavable maltose-binding protein (MBP) tag using sequence and ligation independent cloning (SLIC) (Li and Elledge 2007).

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

Preparation of double-stranded TATA-box containing promoter DNA substrate. AdMLP) TATA-DNA was prepared from synthetic oligonucleotides d(ctgctataaaaggctg) and d(cagcctttatatagcag) (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\(^{\gamma\text{-c}}\). 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 \(\gamma\) chain and the \(\alpha\) chain. Similarly, a second linker (L2) with sequence SRAVDGELFDT was introduced connecting the \(\alpha\) chain with the \(\beta\) chain giving rise to a single-chain TFIIA\(^{\gamma\text{-c}}\) construct \(\gamma\)-L1-\(\alpha\)-L2-\(\beta\) encompassing 240 amino acid residues in total. The gene encoding for TFIIA\(^{\gamma\text{-c}}\) was cloned in a pET28a plasmid resulting in frame with a C-terminal hexa-histidine tag. TFIIA\(^{\gamma\text{-c}}\) 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 40,000g for 45 min, and
loaded on a Talon affinity column. After 10 column volumes of washing with Binding Buffer, TFIIA<sup>c</sup> 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 20 mM Tris-HCl pH 7.4, 150 mM NaCl, 0.5 mM EDTA and 1 mM DTT and loaded onto a Heparin column. TFIIA<sup>c</sup> was eluted with HS Buffer (20 mM Tris-HCl pH 7.4, 1 M NaCl, 0.5 mM EDTA and 1 mM DTT) applying a gradient, and polished by SEC with a Superdex75 column equilibrated in SEC Buffer 2 (20 mM Tris-HCl pH 7.4, 150 mM NaCl, 0.5 mM EDTA and 1 mM DTT).

Purified TFIIA<sup>c</sup> protein was aliquoted and stored in Storage Buffer (20 mM Tris pH 8.0, 1 mM DTT, 0.5 mM EDTA and 150 mM NaCl) at −80°C.

**X-ray crystallography.** Large crystals of TFIIA<sup>c</sup> were obtained by vapor diffusion at room temperature from a protein solution concentrated to 15 mg/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<sup>c</sup> 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<sup>c</sup> 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<sub>free</sub> values of 0.18 and 0.24, respectively. Refinement statistics are provided in Table 1. TFIIA<sup>c</sup> 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\textsuperscript{c} (4 μM) or TBP\textsuperscript{fl} (4 μM), respectively, and TFIIA\textsuperscript{c} (6μM) in EMSA Reaction Buffer (10mM Tris pH 8.0, 60mM KCl, 10mM MgCl\textsubscript{2}, 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% polyacrylamide 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 $1.76 \times 10^6$ mbar and analyser pressure of $2 \times 10^4$ 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
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 ($R_g$) were calculated from the Guinier approximation, to compute the hydrated particle volume using the Porod invariant and to determine the maximum particle size ($D_{\text{max}}$) 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
purified by injecting on a SuperdexS200 10/300 column. Peak fractions containing purified cross-linked samples were concentrated using pin concentrators (Amicon) and separated by SDS-PAGE using a NuPAGE 4-12% bis-Tris gel (Thermo Fisher Scientific).

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

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

**DiAzKs cross-linking**: An unnatural amino acid, DiAzKs, was introduced at K34 position of TAF13 using genetic code expansion (GCE) we implemented recently in our 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...
HPLC and UPLC pumps. A Peltier-cooled box (4 °C) contained two Rheodyne automated valves, a desalting cartridge (Trap Acquity UPLC Protein BEH C18 2.1x5mm, Waters) and a UPLC column (Acquity UPLC BEH C18 1.7µm 1x100mm, Waters). HDX-MS reactions were carried out using either TAF11/TAF13 or TBP alone or in complex at a concentration of 20 µM. Deuteration was initiated by a 5-fold dilution of the protein samples (10 µl) with the same buffer in D2O (40 µl). The proteins were deuterated for 15 sec or 2 min at 4 °C. The exchange was quenched using 50 µl of 200 mM glycine-HCl, pH 2.5 at 4 °C. The proteins or complexes were digested online with immobilized porcine pepsin (Sigma) and recombinant nepenthesin-1. The peptides were desalted for 6 min using a HPLC pump (Agilent Technologies) with 0.1% formic acid in water, at a flow rate of 100 µl/min. Desalted peptides were separated using a UPLC pump (Agilent Technologies) at 50 µl/min for 10 min with 15–70% gradient B (Buffer A: 0.1% formic acid in water; Buffer B: 0.1% formic acid in 95% acetonitrile), followed by 1 min at 100% B. The peptide masses were measured using an electrospray-TOF mass spectrometer (Agilent 6210) in the 300–1300 m/z range. Each deuteration experiment was conducted in triplicate. The Mass Hunter (Agilent Technologies) software was used for data acquisition. The HD Examiner software (Sierra Analytics) was used for HDX-MS data processing. Identification of peptides generated by the digestion was done as described previously (Giladi et al. 2016). Different proteases (pepsin, nepenthesin-1, nepenthesin-2, rhizopuspepsin) or their combinations were tested for protein digestion with pepsin-nepenthesin-1 pair providing the best digestion parameters and sequence coverage.

**Integrative multiparameter-based model building and refinement.** Initial models of the two component structures (TAF11/TAF13, TBP) were taken from the PDB (1BH8 and 1CDW) (Birck et al. 1998; Nikolov et al. 1996). 1BH8 was extended to include a helix structure missing from the complete histone-fold domain as described before (Birck et al. 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
experiments. Three (ts strain(s)) or two (degron strain) independent experiments were performed and data was normalized against the first time point taken. Average absorbance was plotted against time, standard errors of mean (SEM) were calculated over each data point.

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

**TFIID immunoprecipitation experiments.** HeLa cells were grown in suspension culture. 10^{11} cells were harvested by centrifugation and a nuclear extract was prepared according to a modified protocol (Dignam et al. 1983): Briefly, nuclei were prepared by resuspending the pellets in 4 packed cell volume (PCV) of 50 mM Tris-HCl, pH 7.9; 1 mM EDTA; 1 mM DTT and proteinase inhibitors and opening the cells with a Dounce-homogenizer. Nuclei were collected by centrifugation and lysed in 4 PCV of 50 mM Tris-HCl, pH 7.9; 25% glycerol; 500 mM NaCl; 0.5 mM EDTA; 1 mM DTT and protease inhibitors by powerful strokes. The lysate was centrifuged at 50,000 g for 20 minutes. The supernatant was filtered and proteins precipitating in 30% (w/v) (NH_{4})_{2}SO_{4} were pelleted. They were resuspended in 50 mM Tris-HCl, pH 7.9; 20% glycerol; 100 mM KCl; 5 mM MgCl_{2}; 1 mM DTT and dialysed against the same buffer.

For immunoprecipitation, 200 μl protein G-Sepharose (Pharmacia) was incubated with approximately 50 μg of the different antibodies (as indicated). Washed antibody-bound beads were then mixed with 4 mg of pre-cleared HeLa cell nuclear extract and incubated overnight at 4°C. Antibody-protein G Sepharose bound protein complexes were washed three times 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 Fisher 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\(^f\)) and core (TBP\(^c\)), and a single-chain version of TFIIA (TFIIA\(^{sc}\)) were purified to homogeneity as shown by SDS-PAGE (left). TFIIA subunits \(\alpha\)(AA2–59), \(\beta\)(AA302–376) and \(\gamma\)(AA2–110) (Bleichenbacher et al. 2003) were connected with linkers L1(-DGKNTANSANTNTV-) and L2(-SRAVDGELFDT-) as indicated (middle, top). TFIIA\(^{sc}\) 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 1\(\sigma\) (middle, bottom). TFIIA\(^{sc}\) 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
TAF11/TAF13/TBP ternary complex. The second, smaller peak corresponds to excess TAF11/TAF13. (B) Native mass spectrometry of the TAF11/TAF13/TBP complex confirms 1:1:1 stoichiometry. Collision induced dissociation (CID) results in TBP monomer and TAF11/TAF13. Experimental masses are provided (inset). Calculated masses: 20659 Da (TBP); 40691 Da (TAF11/TAF13); 61351 Da (TAF11/TAF13/TBP). Mass spectra of the TAF11/TAF13/TBP complex and TBP dimer are shown in Figure 2–Figure Supplement 1. (C) Small-angle X-ray scattering (SAXS) data of TAF11/TAF13/TBP and TAF11/TAF13. Fit of experimental scattering (dots) and scattering calculated from ab initio models (solid lines) is provided in a log plot (left). SAXS envelopes of TAF11/TAF13 (yellow), TAF11/TAF13/TBP (grey) and an overlay are shown (right). TBP (PDB ID 1CDW) and TAF11/TAF13 (PDB ID 1BH8) were placed into the overlay. A distinct extension (panhandle), likely arising from unstructured tail regions in TAF11/TAF13, enabled alignment of the SAXS envelopes. Kratky plot, P(r) and Guinier analysis for corresponding complexes are provided in Figure 2–Figure Supplement 2. (D) TAF11/TAF13/TBP and TAF11/TAF13 were analysed by hydrogen-deuterium exchange/mass spectrometry (HDX-MS) (Rajabi et al. 2015). Changes in the deuteration level of selected peptides in TAF11, TAF13 or TBP are depicted in diagrams (top row). Peptides protected upon ternary complex formation are coloured in red in cartoon representations of TAF11/TAF13 and TBP (bottom row). One peptide in TBP (grey, far right) becomes more accessible, hinting at disassembly of a TBP dimer when TAF11/TAF13 binds (Figure 2–Figure Supplement 1). All peptides implicated in TAF11/TAF13 binding map to the concave DNA-binding surface of TBP. (N) and (C) indicate N- and C-terminal TBP stirrups, respectively.

Figure 3: Architecture of TAF11/TAF13/TBP complex. TAF11/TAF13/TBP complex architecture was determined by using an integrative multi-parameter approach, We utilized the crystal structure of TBP (Nikolov et al. 1996) as well as the crystal structure of the
TAF11/TAF13 dimer (Birck et al. 1998) combined with our native MS, SAXS, AUC and HDX-MS results as well as distance constraints from CLMS experiments (Figure 3–Figure Supplement 1, Figure 3–Figure Supplement 2). The structure of the TAF11/TAF13/TBP ternary complex is shown in a cartoon representation in stereo (top) and as a space filling model (devoid of unstructured regions) in three views (bottom). Three axes (x, y, z, drawn as arrows) illustrate the special relation between the views. TAF11 is colored in blue, TAF13 in magenta and TBP in green. This model satisfies >90 % of the experimental constrains (Figure 3–Figure Supplement 3, Tables 3 and 4).

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

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

**Figure 1–Figure Supplement 1: Crystal structure of TFIIA^{sc} at 2.4 Å resolution.**

(A) Superimposition of human TFIIA from a TFIIA/TBP/DNA ternary complex (green, PDB ID: 1NVP) and uncomplexed TFIIA^{sc} in our crystal (blue) is shown in a cartoon representation in a stereo view. The structures are virtually identical (rms deviation 0.753 Å). C, N are C- and N-terminal end of single-chain; L1, L2 are Linker regions. (B) Crystal packing of TFIIA^{sc}, highlighting the critical role of engineered linkers (yellow spheres) in mediating the majority of crystal contacts. The unit cell is indicated (grey lines). (C) A close up highlights the interactions mediated by the linker region between symmetry related molecules in the crystal lattice. Symmetry related molecules are colored in grey or green, respectively. Hydrogen bonds (distance cut-off 3.7Å) are shown as black dashed lines. Amino acid residue numbers are indicated. A section of the 2Fo-Fc electron density map 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<sup>s-c</sup>.

Interaction analysis of TAF11/TAF13/TBP complex with TFIIA<sup>s-c</sup> by SEC. Preformed TAF11/TAF13/TBP complex and TFIIA<sup>s-c</sup> 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<sup>s-c</sup> 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 (Diazipine-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**
Cartoon representation (top row) and electrostatic surface potential (bottom row) of TBP (left, PDB ID: 1TBP), TATA-DNA (PDB ID: 1CDW), and TAF1-TAND from *Drosophila* (PDB ID: 1TBA) and Yeast (PDB ID: 4B0A). The DNA binding surface of TBP is shown (left), as well as the corresponding complementary surfaces of the interactors (right). Red indicates regions with negative charge, blue indicates regions with positive charge. The shape and surface charge similarity between *Drosophila* TAF1-TAND and TATA-DNA is evident (‘TATA-box mimicry’), but is less pronounced in TAF1-TAND from yeast. Yeast TAF1-TAND binding is more extensive involving also the convex surface of TBP (Table 5).


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


47
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
The authors declare no competing financial interest.
FIGURES

Figure 1

A

B

C

D

E

The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY 4.0 International license.
Figure 2

A

B

C

D

TAF11/TAF13
TAF11/TAF13/TBP

\[ S (\AA^{-1}) \]

TAF11/TAF13

\[ \chi^2 = 1.12 \]

\[ \chi^2 = 2.01 \]

Rel Denat level (%)
Figure 3
Figure 5

![Graph showing NE anti-TAF1, NE anti-TAF11, NE anti-TAF7, and Cytoplasmic anti-TAF11 IP values for TAF proteins.](image-url)
Figure 6
**Figure 7**

![Diagram showing TFIID complexes in different configurations: Canonical, Rearranged, Inhibited TFIID, and Activated TFIID.](image-url)

- TFIID with TBP, TAF1, TAF13, and TAF11 in different positions.
- Promoter DNA indicated.

---

56
Figure 1–Figure Supplement 1

A

B

C

TFIIA<sup>S-C</sup>  
TFIIA

L1  
N

L2  
C

R156  
E161  
V158  
T165  
P113  
L162

not peer-reviewed) is the author/funder. It is made available under a CC-BY 4.0 International license.
Figure 1–Figure Supplement 2

[Graph showing absorption at 280 nm against elution volume and fractions]

IN M 1 2 3 4 5 6 7 8 9 10 11

- TAF11
- TFIIA^S-C
- TBP
- TAF13
Figure 2–Figure Supplement 1
Figure 2–Figure Supplement 2

A

B

C

TAF11/TAF13/TBP
Rg = 40.9 ± 0.6 Å

TAF11/TAF13
Rg = 40.3 ± 3.6 Å

TAF11/TAF13/TBP

TAF11/TAF13

P(r)

r (Å)

r (Å)
Figure 3–Figure Supplement 1
Figure 3–Figure Supplement 2

[Diagram showing proteins TAF11, TAF13, and TBP with labeled residues and mass spectroscopy data.]

m/z 797.41, z=5, Δm=-0.3 ppm

% of base peak

200 400 600 800 1000 1200

m/z

y1 b2 y2 b3 y3 y4 y5 b11 b13 b14 b7 y6+P y8+P y9+P y7+P y10+P y11+P y12+P y13+P
Figure 3–Figure Supplement 3

A

B

TAF11
TAF13
TBP
Figure 4–Figure Supplement 1

A

B

BY4741

37°C

Absorbance [600nm]

Time Points [Hour]

Empty Vector

WT

Mutant B

Mutant A

1 2 3 4 5 6 7 8 9

64

bioRxiv preprint first posted online Jul. 26, 2017; doi: http://dx.doi.org/10.1101/168716. The copyright holder for this preprint (which was not peer-reviewed) is the author/funder. It is made available under a CC-BY 4.0 International license.
Figure 6–Figure Supplement 1
Table 1.
X-ray data collection and refinement statistics

<table>
<thead>
<tr>
<th>Data collection</th>
<th>TFIIA&lt;sup&gt;c&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Space group</td>
<td>P65</td>
</tr>
<tr>
<td>Cell dimensions</td>
<td></td>
</tr>
<tr>
<td>$a$, $b$, $c$ (Å)</td>
<td>123.3, 123.3, 34.8</td>
</tr>
<tr>
<td>$\alpha$, $\beta$, $\gamma$ (°)</td>
<td>90, 90, 120</td>
</tr>
<tr>
<td>Resolution (Å)</td>
<td>53.4-2.4</td>
</tr>
<tr>
<td>Last resolution bin (Å)</td>
<td>2.52-2.38</td>
</tr>
<tr>
<td>$R_{\text{measure}}$ (%)</td>
<td>12.9 (64.8)</td>
</tr>
<tr>
<td>$I/\sigma I$</td>
<td>11.5 (2.72)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>99.8 (99.9)</td>
</tr>
<tr>
<td>Multiplicity</td>
<td>6.7 (6.8)</td>
</tr>
</tbody>
</table>

| Refinement            |                    |
| Resolution (Å)        | 40.36-2.38 (2.44-2.38) |
| No. reflections       |                    |
| Work set              | 11859              |
| Free set              | 601                |
| $R_{\text{work}}$     | 0.18 (0.27)        |
| $R_{\text{free}}$     | 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.*
# Table 2.

Data collection and refinement statistics SAXS

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

† Reported for experimental merged data.
### Table 3.
Peptide deuteration level changes upon complex formation in HDX-MS

<table>
<thead>
<tr>
<th>Residues Numbers</th>
<th>Sequence</th>
<th>Δ%D (15 Sec)†</th>
<th>Δ%D (120 Sec)†</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>TAF11</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>93-96</td>
<td>EKKQ</td>
<td>-8.465</td>
<td>-7.309</td>
</tr>
<tr>
<td>105-109</td>
<td>KMQIL</td>
<td>-7.014</td>
<td>-7.295</td>
</tr>
<tr>
<td>150-153</td>
<td>VVIA</td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>TAF13</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>14-31</td>
<td>NEEIGGAEGGQGKRKRL</td>
<td>-7.507</td>
<td>-7.192</td>
</tr>
<tr>
<td>32-35</td>
<td>FSKE</td>
<td>-8.794</td>
<td></td>
</tr>
<tr>
<td>36-40</td>
<td>LRCMM</td>
<td>-7.0</td>
<td></td>
</tr>
<tr>
<td>86-88</td>
<td>IVF</td>
<td>-8.598</td>
<td></td>
</tr>
<tr>
<td>86-89</td>
<td>IVFL</td>
<td>-10.366</td>
<td></td>
</tr>
<tr>
<td>97-104</td>
<td>FARVKDLDL</td>
<td></td>
<td>-7.223</td>
</tr>
<tr>
<td>116-124</td>
<td>AFDEANYGS</td>
<td>10.89</td>
<td>8.721</td>
</tr>
<tr>
<td><strong>TBP</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>157-167*</td>
<td>IVPQLQNVST</td>
<td>9.038</td>
<td></td>
</tr>
<tr>
<td>167-174</td>
<td>TVNLGCKL</td>
<td>-9.675</td>
<td>-10.171</td>
</tr>
<tr>
<td>193-197</td>
<td>FAAV</td>
<td>-14.068</td>
<td>-12.267</td>
</tr>
<tr>
<td>197-208</td>
<td>IMRIREPTTL</td>
<td>-7.785</td>
<td></td>
</tr>
<tr>
<td>199-208</td>
<td>RIREPTTL</td>
<td>-9.661</td>
<td></td>
</tr>
<tr>
<td>209-215</td>
<td>IFSSGKM</td>
<td>-8.655</td>
<td>-8.75</td>
</tr>
<tr>
<td>232-244</td>
<td>KYARVVQKLGFP</td>
<td>-21.337</td>
<td>-18.382</td>
</tr>
<tr>
<td>233-242</td>
<td>YARVVQKLG</td>
<td>-7.128</td>
<td></td>
</tr>
<tr>
<td>247-252</td>
<td>LDFKIQ</td>
<td></td>
<td>-7.271</td>
</tr>
<tr>
<td>250-256</td>
<td>KIQNMG</td>
<td>-8.913</td>
<td></td>
</tr>
<tr>
<td>259-266</td>
<td>DVKFPIRL</td>
<td>-11.932</td>
<td>-10.109</td>
</tr>
<tr>
<td>259-268</td>
<td>DVKFPIRLEG</td>
<td>-13.79</td>
<td>-11.479</td>
</tr>
<tr>
<td>260-266</td>
<td>VKFPIRL</td>
<td>-10.393</td>
<td></td>
</tr>
<tr>
<td>281-287</td>
<td>PELFPGL</td>
<td>-7.811</td>
<td></td>
</tr>
<tr>
<td>326-335</td>
<td>PILKGFRRKT</td>
<td>-7.077</td>
<td>-7.902</td>
</tr>
</tbody>
</table>

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

* TBP peptide with increasing deuteration level upon complex formation.
**Table 4.** Cross-links observed by BS3 CLMS

<table>
<thead>
<tr>
<th>TAF11 Residue</th>
<th>TBP Residue</th>
<th>No. of matches</th>
<th>Highest Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>K122</td>
<td>K160</td>
<td>5</td>
<td>9.416</td>
</tr>
<tr>
<td>K110</td>
<td>K160</td>
<td>2</td>
<td>8.818</td>
</tr>
<tr>
<td>K107</td>
<td>K106</td>
<td>2</td>
<td>8.314</td>
</tr>
<tr>
<td>K160</td>
<td>K58</td>
<td>1</td>
<td>8.085</td>
</tr>
<tr>
<td>K107</td>
<td>K160</td>
<td>2</td>
<td>7.926</td>
</tr>
<tr>
<td>K107</td>
<td>K58</td>
<td>1</td>
<td>7.684</td>
</tr>
<tr>
<td>K156</td>
<td>K160</td>
<td>1</td>
<td>6.928</td>
</tr>
<tr>
<td>K108</td>
<td>K99</td>
<td>1</td>
<td>6.124</td>
</tr>
<tr>
<td>K120</td>
<td>K160</td>
<td>1</td>
<td>4.145</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(9 unique links)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TAF13 Residue</th>
<th>TBP Residue</th>
<th>No. of matches</th>
<th>Highest Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>K34</td>
<td>K44</td>
<td>1</td>
<td>10.578</td>
</tr>
<tr>
<td>K34</td>
<td>K58</td>
<td>1</td>
<td>7.856</td>
</tr>
<tr>
<td>K101</td>
<td>K99</td>
<td>1</td>
<td>7.584</td>
</tr>
<tr>
<td>K115</td>
<td>K58</td>
<td>1</td>
<td>7.519</td>
</tr>
<tr>
<td>K115</td>
<td>K44</td>
<td>1</td>
<td>5.742</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(5 unique links)</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>TAF11 Residue</th>
<th>TAF13 Residue</th>
<th>No. of matches</th>
<th>Highest Score</th>
</tr>
</thead>
<tbody>
<tr>
<td>K220</td>
<td>K96</td>
<td>2</td>
<td>12.795</td>
</tr>
<tr>
<td>K110</td>
<td>K101</td>
<td>2</td>
<td>12.014</td>
</tr>
<tr>
<td>K229</td>
<td>K96</td>
<td>3</td>
<td>11.053</td>
</tr>
<tr>
<td>K114</td>
<td>K101</td>
<td>5</td>
<td>10.98</td>
</tr>
<tr>
<td>K229</td>
<td>K101</td>
<td>6</td>
<td>10.909</td>
</tr>
<tr>
<td>K156</td>
<td>K34</td>
<td>3</td>
<td>9.627</td>
</tr>
<tr>
<td>K107</td>
<td>K92</td>
<td>2</td>
<td>9.321</td>
</tr>
<tr>
<td>K222</td>
<td>K96</td>
<td>2</td>
<td>9.308</td>
</tr>
<tr>
<td>K156</td>
<td>K111</td>
<td>10</td>
<td>9.226</td>
</tr>
<tr>
<td>K160</td>
<td>K111</td>
<td>2</td>
<td>8.371</td>
</tr>
<tr>
<td>K156</td>
<td>K115</td>
<td>1</td>
<td>8.326</td>
</tr>
<tr>
<td>K232</td>
<td>K92</td>
<td>1</td>
<td>8.292</td>
</tr>
<tr>
<td>K107</td>
<td>K96</td>
<td>1</td>
<td>7.981</td>
</tr>
<tr>
<td>K110</td>
<td>K115</td>
<td>1</td>
<td>7.943</td>
</tr>
<tr>
<td>K119</td>
<td>K101</td>
<td>2</td>
<td>7.934</td>
</tr>
<tr>
<td>K122</td>
<td>K101</td>
<td>1</td>
<td>7.867</td>
</tr>
<tr>
<td>K229</td>
<td>K92</td>
<td>4</td>
<td>6.872</td>
</tr>
<tr>
<td>K110</td>
<td>K92</td>
<td>1</td>
<td>6.817</td>
</tr>
<tr>
<td>K107</td>
<td>K101</td>
<td>1</td>
<td>6.437</td>
</tr>
<tr>
<td>K231</td>
<td>K92</td>
<td>1</td>
<td>5.77</td>
</tr>
<tr>
<td>K229</td>
<td>S74</td>
<td>1</td>
<td>3.42</td>
</tr>
<tr>
<td>K231</td>
<td>S74</td>
<td>1</td>
<td>3.42</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(22 unique links)</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Interactor</td>
<td>Interface (Å²)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------------------------</td>
<td>----------------</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TBP Dimer</td>
<td>3010.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DNA</td>
<td>4020.1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TAF1-TAND (D. melanogaster)</td>
<td>3287.8</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TAF1-TAND (Yeast)</td>
<td>7483.1†</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mot1 (E. cuniculi)</td>
<td>4300.0</td>
<td></td>
<td></td>
</tr>
<tr>
<td>TAF11/TAF13</td>
<td>3305.2</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Calculated with PyMol v1.8.2.0 ([www.pymol.org](http://www.pymol.org))

† includes TAND1 and TAND2