HHS Public Access

Author manuscript

Int J Biol Macromol. Author manuscript; available in PMC 2024 December 31.

Published in final edited form as:

Int J Biol Macromol. 2023 December 31; 253(Pt 2): 126764. doi:10.1016/j.ijbiomac.2023.126764.

Structural basis for evolutionarily conserved interactions between TFIIS and Paf1C

Jie Gao[a,b,1], **Miki Jishage**[c,1], **Yuzhu Wang**[a,1], **Rui Wang**[a,d], **Meng Chen**[a], **Zhongliang Zhu**[a], **Jiahai Zhang**[a], **Yating Diwu**[a], **Chao Xu**[a], **Shanhui Liao**[a,*], **Robert G. Roeder**[c,*], **Xiaoming Tu**[a,*]

[a]MOE Key Laboratory for Membraneless Organelles and Cellular Dynamics, Hefei National Research Center for Interdisciplinary Sciences at the Microscale, School of Life Sciences, University of Science and Technology of China, Hefei, Anhui 230022, PR China

[b]Department of Ophthalmology, The Second Affiliated Hospital of Anhui Medical University, 678 Furong Road, Hefei, Anhui, PR China

[c]Laboratory of Biochemistry and Molecular Biology, The Rockefeller University, New York, NY 10065, USA

[d]Department of Anthropotomy and Histoembryology, Medical College, Henan University of Science and Technology, Luoyang, Henan 471023, PR China

Abstract

The elongation factor TFIIS interacts with Paf1C complex to facilitate processive transcription by Pol II. We here determined the crystal structure of the trypanosoma TFIIS LW domain in a complex with the LFG motif of Leo1, as well as the structures of apo-form TFIIS LW domains from trypanosoma, yeast and human. We revealed that all three TFIIS LW domains possess a conserved hydrophobic core that mediates their interactions with Leo1. Intriguingly, the structural study revealed that trypanosoma Leo1 binding induces the TFIIS LW domain to undergo a conformational change reflected in the length and orientation of α6 helix that is absent in the yeast and human counterparts. These differences explain the higher binding affinity of the TFIIS LW domain interacting with Leo1 in trypanosoma than in yeast and human, and indicate species-specific variations in the interactions. Importantly, the interactions between the TFIIS LW domain and an LFG motif of Leo1 were found to be critical for TFIIS to anchor the entire Paf1C complex. Thus, in addition to revealing a detailed structural basis for the TFIIS-Paf1C interaction, our studies also shed light on the origin and evolution of the roles of TFIIS and Paf1C complex in regulation of transcription elongation.

[*]Corresponding authors. ajsod@mail.ustc.edu.cn (S. Liao), roeder@mail.rockefeller.edu (R.G. Roeder), xmtu@ustc.edu.cn (X. Tu).
[1]These authors contributed equally to this work.

CRediT authorship contribution statement

J. Gao, S. Liao, R. Roeder and X. Tu designed the research. J. Gao, M. Jishage, Y. Wang, R. Wang, S. Liao, M. Chen, Z. Zhu, J. Zhang and Y. Diwu performed the experiments. J. Gao, S. Liao and X. Tu analyzed the data. J. Gao, S. Liao, R. Roeder and X. Tu wrote the paper. C. Xu discussed and gave advice on the manuscript.

Declaration of competing interest

The authors declare that they have no conflicts of interest with the contents of this article.

Appendix A. Supplementary data

Supplementary data to this article can be found online at https://doi.org/10.1016/j.ijbiomac.2023.126764.

Keywords

TFIIS; Paf1C complex; Leo1; LW domain; Transcription regulation; Protein structure

1. Introduction

Eukaryotic gene transcription regulation is extremely complex in space and time and defects in transcription processes are closely related to a variety of diseases. The regulation of transcription by RNA polymerase II (Pol II), which transcribes genes encoding proteins and some small RNAs, occurs in three major steps: transcription initiation, elongation and termination. In response to enhancer-bound transcriptional activators, a number of coactivators and general initiation and elongation factors mediate preinitiation complex formation and subsequent initiation and elongation by Pol II [1,2]. The efficiency of elongation by Pol II is regulated by a number of additional factors such as TFIIS, Spt4-Spt5, FACT, SPT6 and Paf1C [3]. Pol II achieves efficient nucleosome passage when the elongation factors DSIF, Paf1C, Rtf1, SPT6, and TFIIS are present [4–6]. Recently, a series of cryo-EM structures of Pol II-Paf1C-TFIIS complexes have been reported [4–6]. However, the detailed molecular mechanism of the interactions between TFIIS and Paf1C remains unclear.

TFIIS is one of the best-characterized transcription elongation factors [7]. TFIIS reactivates arrested elongation complexes that have fallen out of register with template DNA by generating a new 3′ end in the nascent mRNA [7]. It can stimulate the intrinsic transcript cleavage activity of backtracked Pol II to allow it to resume transcript elongation [3]. It promotes Pol II passage through nucleosome barriers [8,9] and DNA damage-induced lesions [2,10]. It is also implicated in transcription initiation through the formation of Pol II pre-initiation complex [11]. Typically, TFIIS is composed of three domains [12,13]. The N-terminus of TFIIS, termed domain I (or LW domain), consists of a four-helix bundle that is dispensable for its biological functions. The LW domain within domain I plays a role in the interactions between TFIIS and other elongation factors, enabling TFIIS to bind to a carrier protein that accomplishes nuclear import [13]. The other parts of TFIIS, domains II and III, are required for its cleavage-stimulating activity. Domain II forms a three-helix bundle and connects domain III, which forms a zinc ribbon composed of three antiparallel β-sheets, through an interdomain linker. Domain II and the linker are required for Pol II binding, while domain III is essential for stimulation of RNA cleavage [12].

RNA polymerase associated factor 1 complex (Paf1C), originally identified in yeast, is composed of five subunits: Paf1, Cdc73, Leo1, Ctr9 and Rtf1 [14]. The human Paf1 complex contains an additional subunit named Ski8/Wdr61 [15]. A number of studies have demonstrated that the defect of Paf1C is associated with several human diseases [16]. Paf1C carries out multiple functions during transcription and is therefore another important transcription elongation factor [17]. It plays an important role in the termination of Pol II transcription and the recruitment of proteins required for proper RNA 3′ end formation. Paf1C interacts with other transcription elongation factors to ensure the efficiency of transcription, and facilitates modification on the chromatin template [17].

Recent studies have yielded exciting new insights into the mechanisms by which Paf1C regulates transcription elongation, epigenetic modifications, and post-transcriptional steps in eukaryotic gene expression [18]. Leo1 is a subunit of Paf1C with roles in transcription elongation and chromatin modification [19]. The human Leo1 has an important role in transcriptional synergy beyond the direct interaction with TFIIS [20]. In yeast, the LW domain of TFIIS, which is mobile in the Pol II-TFIIS structure, interacts with Paf1C to increase its affinity for Pol II [6].

Some structurally relevant studies on the interactions between TFIIS and Paf1C have been reported recently. Previous study reported the cryo-EM structure of the Paf1C-TFIIS-Pol II complex from *S. cerevisiae* [6]. This finding offered a three-dimensional framework for further investigation of Paf1C's role in chromatin-mediated transcription. Furthermore, a study identified a common binary interaction module consisting of TFIIS N-terminal domains (TNDs) and natively unstructured TND-interacting motifs (TIMs) [21]. Meanwhile, the TNDs were reported to interact with the conserved TIMs of Leo1 from *H. sapiens* as well [21]. Farnung et al. then demonstrated that Paf1C and TFIIS enable efficient nucleosome passage. They determined the cryo-EM structure that shows the loss of H2A–H2B dimer and the backtracking of Pol II during nucleosomal traversal [4]. More recently, Farnung's team further reported cryo-EM structure of the Paf1C-TFIIS-Pol II complex from *H. sapiens*. This structure uncovers a direct role for Pol II and transcription elongation factors in nucleosome retention [5].

While the composition of the nuclear RNA polymerase complexes in *T. brucei* (a protozoan unicellular eukaryote causing sleeping sickness in human and nagana in cattle) has been studied extensively [22], studies of the transcription factors associated with them are very limited. A number of the basal transcription factors in eukaryotes are either absent or very divergent in trypanosoma [23]. The primary sequence of these transcription factors in *T. brucei* are extremely divergent from other eukaryotic counterparts [24]. In order to explore the structural basis and evolution of interactions between TFIIS and Paf1C from different eukaryotes, in this study, we determined the structures of the apo-form TbTFIIS2–2 LW domain and the TbTFIIS2–2 LW domain in a complex with Leo1 from *Trypanosoma brucei* (*T. brucei*), the ScTFIIS LW domain from *Saccharomyces cerevisiae* (*S. cerevisiae*) and the HsTFIIS LW domain from *Homo sapiens* (*H. sapiens*) and investigated the molecular mechanism of interactions between TFIIS and Leo1 from these three eukaryotes.

2. Materials and methods

2.1. Cloning, protein expression and purification

DNA fragments of genes (UniProt: C9ZJ96, P07273 and P23193) encoding TbTFIIS2–2 (residue 251–343/359), ScTFIIS (residue 1–81/93) and HsTFIIS (residue 1–78/96) LW domains were amplified from genomic DNA by PCR and cloned into the vector pET28-MHL, respectively. TbTFIIS2–$2_{251–359}$-$(Gly\text{-}Gly\text{-}Ser)_4$-$TbLeo1_{71–86}$ fusion clone was constructed by overlap extension PCR and was also cloned into the vector pET28-MHL. All proteins were expressed in *Escherichia coli* BL21 (DE3) cells and protein expression were induced at OD_{600} of 0.8, 16 °C for 20 h with 0.5 mM isopropyl beta-D-1-thiogalactoside (IPTG). Cells were harvested at 3600 g, 4 °C for 10 min, re-suspended in

lysis buffer (20 mM Tris, 400 mM NaCl, pH 7.8), and lysed by sonication. Lysate was cleared by centrifugation at 16000*g*, 4 °C for 30 min, and supernatant was collected. The recombinant TFIIS LW domain containing a N-terminal His-tag was purified by Ni^{2+}-NTA column (GE healthcare). Gel filtration and ion exchange chromatography were explored for further purification of TbTFIIS2–2, ScTFIIS and HsTFIIS LW domains. Corresponding fractions were collected and concentrated as required.

2.2. In vitro mutagenesis

A series of point mutations were introduced into the recombinant TbTFIIS2–2, ScTFIIS and HsTFIIS LW-pET28-MHL vectors. Site-specific mutations were introduced using two reverse and complementary primers containing the mutated codon with the plasmid encoding wild-type TbTFIIS2–2, ScTFIIS and HsTFIIS LW domains as the template, respectively. After PCR reaction, the PCR product was digested with *Dpn* I (TaKaRa, Dalian, China) overnight to remove methylated parental non-mutated plasmid, and then transformed into *E. coli* BL21 (DE3). Mutated TbTFIIS2–2, ScTFIIS and HsTFIIS LW domains were expressed and purified in the same way as wild-type proteins.

2.3. Crystallization, data collection and structure determination

The crystal was grown at 18 °C incubator by sitting drop vapor diffusion method by mixing 1 μl TbTFIIS2–$2_{251–359}$-$(Gly\text{-}Gly\text{-}Ser)_4$-$TbLeo1_{71–86}$ fusion protein (15 mg/mL) with 1 μl reservoir buffer (3.5 M NaCOOH, pH 7.0). The apo-form TbTFIIS2–2 LW domain (residue 251–359) was crystallized by mixing an equal volume of 15 mg/ml protein with crystallization buffer (1.5 M Lithium sulfate monohydrate, 0.1 M Tris, pH 8.5). The crystal of 20 mg/mL ScTFIIS LW domain (residue 1–81) was grown in buffer by mixing equal volume of protein and buffer containing 0.1 M Tris and 25 % PEG3350 (pH 8.5). Before flash-freezing crystals in liquid nitrogen, all crystals were soaked in a cryo-protectant consisting of 75 % reservoir solution plus 25 % glycerol. The diffraction data was collected on BL17U1 and BL18U1 at Shanghai Synchrotron Facility (SSRF) [25]. The initial model for apo-form TbTFIIS2–2 LW domain was solved by CRANK2 using X-ray diffraction data of SeMet-TbTFIIS2–$2_{251–359}$-$(Gly\text{-}Gly\text{-}Ser)_4$-$TbLeo1_{71–86}$ crystal [26]. The initial model for ScTFIIS LW domain was built by Robetta server [27]. All crystal structures were then solved by molecular replacing using PHENIX, and manually built with Coot and refined by PHENIX [28,29]. The statistics for data collection and structural refinement are listed in Table 1.

2.4. NMR spectroscopy, data processing and structure calculation

To prepare the NMR samples, ^{15}N, ^{13}C-labeled ScTFIIS (longer version, residue 1–93) and HsTFIIS (residue 1–96) LW domains were expressed and purified in the same way as described above except for that super broth was replaced by M9 medium containing 2.5 g/L 99 % ^{13}C-glucose and 0.5 g/L 99 % $^{15}NH_4Cl$ as the sole carbon and nitrogen source, respectively. All buffer for NMR samples contain 25 mM phosphate (pH 6.8), 150 mM sodium chloride and 2 mM EDTA including 10 % D_2O. All NMR spectra including ^1H–^{15}N HSQC, CBCA(*CO*)NH, HBHA (*CO*)NH, CBCANH, HC(*CO*)NH, H(*CC*)ONH, ^{15}N-edited NOESY and ^{13}C-edited NOESY were acquired at 293 K on a Bruker DMX 600 spectrometer. The NMR data were processed with software of NMRpipe, NMRDraw

and SPARKY [30,31]. Dihedral angle restraints were calculated by TALOS program [32]. Structure was calculated using the program CYANA [33]. The final 20 lowest-energy structures were selected and analyzed by MOLMOL and PROCHECK online [34,35]. Finally, the structures of CYANA calculation were manually refined by Xplor-NIH [36–38].

2.5. NMR chemical shift perturbation

Chemical shift perturbation was performed to investigate the ability and the residues of ScTFIIS (longer version, residue 1–93) and HsTFIIS (residue 1–96) LW domains recognizing their corresponding peptides of Leo1. 0.3 mM ^{15}N-labeled ScTFIIS and HsTFIIS LW domains were titrated with peptide (containing LFG motif) of Leo1 to a molar ratio (TFIIS/Leo1: 1:1). Peptides stock solutions in identical buffer were titrated with a sample dilution of <10 % of total volume. ^{1}H–^{15}N HSQC spectra were acquired for every titration for analysis.

2.6. GST pull-down and coimmunoprecipitation

Different regions of the gene encoding TbLeo1 (residue 1–100, 1–200, 101–200, 49–100) were amplified by PCR from the *T. brucei* genome and cloned into pET-22b (+) and pGEX-4T1. A region of the TbTFIIS2–2 LW domain (residue 251–359) was amplified by PCR from the *T.brucei* genome and cloned into pGEX-4T1 vector. For *in vitro* binding assays, full-length and a series of truncated sequence of TbTFIIS2–2 and TbLeo1 were fused to a GST tag. The GST fusion proteins coupled to glutathione Sepharose 4B (GE Healthcare Life Sciences) were then used as affinity matrixes to absorb GST-tagged proteins in phosphate-buffered saline (PBS) containing 1 mM EDTA and 0.1 % Triton X-100 for 40 min at 4 °C. The beads were washed three times in ice-cold PBS containing 0.1 % Triton X-100 and finally incubated in 50 μl PBS containing 10 mM reduced glutathione (pH 8.0) for 10 min. The supernatants were transferred to new EP tubes, mixed with 50 μl 2 × SDS loading buffer and 1 mM DTT, then boiled and resolved by SDS-PAGE.

HeLa nuclear extract containing HsPaf1C or purified HsPaf1C complex was incubated with wild-type or mutated GST-HsTFIIS (residue 1–135) LW domains. Bound proteins were then visualized by immunoblot with related antibodies. GST pull-down and coimmunoprecipitation assays were performed as described [20].

2.7. Isothermal titration calorimetry (ITC)

Interactions between the TbTFIIS2–2 LW domain and the fragments of TbLeo1 were studied by ITC 200 (GE Healthcare). The data was analyzed by MicroCal LLC ITC software (MicroCal). Three peptides of TbLeo1 (residue 71–86, 77–95 and 85–100) and the peptide (residue 71–86) whose three conserved LFG residues were mutated to LAG or AAA were designed. Three peptides of ScLeo1 (residue 64–78, 71–85 and 86–100) and the peptide (residue 64–78) whose three conserved LFG residues were mutated to LAG were also designed. Meanwhile, three peptides of HsLeo1 (residue 2–16, 55–70 and 309–323) and the peptides (residue 64–78 and 71–85) whose three conserved LFG residues were mutated to LAG were designed as well. Meanwhile, different mutants of TFIIS LW domains were performed in the same way as with wild-type proteins. The volume of cell is about five times bigger than that of the syringe of ITC 200. For the binding experiment, 50 uM TFIIS LW

domain was titrated against 1mM peptides in 25 mM Tris, 1 mM EDTA and 1 % Glycerol buffer (pH 7.4) containing 200 mM NaCl at 25 °C. Sixteen injections of 2 μl of peptides were done at 120 s intervals. Titration of peptides to buffer was performed and subtracted to correct for the heat of dilution. The sequences of peptides are shown in Supplementary Table S1. The thermodynamic parameters of the ITC experiments are listed in Table 3.

2.8. Yeast two-hybrid (Y2H) screening

The DNA sequences of TbPaf1C subunits (TbCtr9, TbCdc73, TbLeo1, TbPaf1 and TbRtf1) and TbTFIIS2–2 were cloned into pGADT7 and pGBKT7 respectively, and co-transformed into the yeast strain AH109 containing the reporter genes *his3* and *lacZ.* Double transformants were plated on selective media (SD-Leu-Trp plate) and incubated for 2–3 days at 30 °C. The momoclonal yeast was then picked from the SD-Leu-Trp plate (2D plate) and resuspended with 100 ul sterilized water. 1 ul of yeast cells was then spotted onto 2D plate and SD-Leu-Trp-His (3D plate) and the cells were diluted to 1:10 and 1:100, spotted in a line on the 2D and 3D plates.

3. Results

3.1. TbTFIIS2–2, ScTFIIS and HsTFIIS all contain a conserved LW domain

Sequence alignments of transcription factor TFIIS from *T. brucei*, *S. cerevisiae* and *H. sapiens* were performed using ClustalW2 and the output files were processed using ESPript (http://espript.ibcp.fr/ESPript/ESPript/). Sequence analysis showed that they display low sequence identity and similarity (Fig. S1). TbTFIIS2–2 is different from other canonical TFIIS in architecture. It lacks the canonical domains II and III found in ScTFIIS and HsTFIIS but contains an additional PWWP domain in the N-terminus and a long disordered region in the middle (Fig. 1A). All TFIISs contain a domain I (LW domain). TbTFIIS2–2 contains this LW domain in its C-terminus, whereas ScTFIIS and HsTFIIS contain this LW domain in their N-termini (Fig. 1A). Sequence analysis indicated that although TFIISs show low sequence identity in general, the LW domains of TFIISs show relatively higher sequence identity (Fig. S1).

3.2. The interactions between TFIIS and Paf1C are mediated by the LW domain of TFIIS and a LFG motif in Leo1 subunit of Paf1C

The interactions between TFIIS and Paf1C have been revealed in different eukaryotes including yeast and human [6,20]. To investigate whether this interaction also exists in *T. brucei*, yeast two hybrid assay was used to observe the interactions between TFIIS and the subunits of Paf1C complex. TbTFIIS2–2 is able to bind to TbLeo1, and some other subunits as well, which suggested that the interactions between TFIIS and Paf1C might be prevalent in eukaryotes (Fig. S2).

It has been reported that the interactions between yeast TFIIS and Paf1C are mediated through the LW domain of TFIIS and the Leo1 subunit of Paf1C [6]. At the same time, sequence analysis indicated that although Leo1 subunits from different eukaryotes show low sequence identity in general (Fig. S3A), they all contain a highly conserved LFG motif (Fig. 1B). Therefore, in order to further determine whether this interaction in trypanosoma

is also mediated by the LW domain of TbTFIIS2–2, and whether the LFG motif of TbLeo1 might also be important for this interaction, a TbLeo1 fragment containing the LFG motif ($TbLeo1_{1–100}$) was expressed and analyzed for an interaction with the TbTFIIS2–2 LW domain by GST pull down. The results indicated that the N-terminal fragment of TbLeo1 containing the LFG motif ($TbLeo1_{1–100}$) is able to bind to the TbTFIIS2–2 LW domain (Fig. S3B). Furthermore, in order to verify that the LFG motif is essential for its interaction with TbTFIIS2–2 LW domain, three peptides corresponding to different fragments ($TbLeo1_{71–86}$ containing the LFG motif and $TbLeo1_{77–95}$ and $TbLeo1_{85–100}$ without the LFG motif) of TbLeo1, were synthesized and analyzed by ITC assays. ITC results indicated that only $TbLeo1_{71–86}$ containing LFG motif is able to interact with TbTFIIS2–2 LW domain, with a dissociation constant (K_d) of 0.15 μM (Fig. 2A), indicating that the LFG motif of TbLeo1 is essential for this interaction.

The LFG motif also exists in yeast Leo1 and human Leo1. In order to investigate whether this conserved LFG motif is also essential for TFIIS binding in yeast and human similar to the case in trypansoma, different fragments of ScLeo1 and HsLeo1 were also synthesized and tested for interactions with the homologous TFIIS LW domains by ITC. The results confirmed that only a fragment ($ScLeo1_{54–68}$) containing the LFG motif is able to bind to the ScTFIIS LW domain with a dissociation constant (K_d) of 3.30 μM (Fig. 2B). Human Leo1 contains three LFG motifs, and the ITC results indicated that the two fragments containing LFG motif ($HsLeo1_{2–16}$ and $HsLeo1_{55–70}$) are able to interact with HsTFIIS LW domain with dissociation constants (K_d) of 36.00 and 44.00 μM, respectively (Fig. 2C). Our results thus revealed that a highly conserved LFG motif in Leo1 is critical for its interactions with TFIIS LW domain in eukaryotes (Fig. 2D).

3.3. Crystal structure of the apo-form TbTFIIS2–2 LW domain from trypanosoma

To provide structural insights into the TbTFIIS2–2 LW domain, the apo-form TbTFIIS2–2 LW domain (residue 251–359 of TbTFIIS2–2) was expressed, purified and crystallized. The crystal structure of apo-form TbTFIIS2–2 LW domain was solved (Fig. 3A) to a resolution of 2.25 Å and deposited to Protein Data Bank (7XGW). The crystallographic statistics are summarized in Table 1.

In the apo-form structure, each asymmetrical unit contains six molecules of the TbTFIIS2–2 LW domain (Fig. 3A). The TbTFIIS2–2 LW domain adopts a fold formed by a six-helix bundle, composing of three long α-helices and three short α-helices. Among six molecules, the discrepancy comes from the different orientation and length of the α6 helix, suggesting that the comformation around this region is highly dynamic. Based on these results, the conformations of the apo-form TbTFIIS2–2 LW domain are categorized into two groups, group C1 and group C2 (Fig. 3A).

3.4. Crystal structure of the trypanosoma TbTFIIS2–2 LW domain in a complex with TbLeo1 and molecular mechanism of their interactions

To explore the molecular mechanism of the interactions between TFIIS and Leo1, a recombinant protein containing the LFG motif of TbLeo1 (residue 71–86 of TbLeo1) fused to the C-terminus of TbTFIIS2–2 LW domain (residue 251–359 of TbTFIIS2–2)

was expressed, purified and crystallized. X-ray crystallography was employed to study the structure of this complex. The detailed crystal structures and crystallographic statistics are summarized in Fig. 3B–E, Fig. S4 and Table 1. The structures have been deposited into Protein Data Bank (7FAX).

In this complex, α5 and α6 helices of TbTFIIS2–2 LW domain pack with each other by hydrophobic forces and hydrogen bonds in anti-parallel direction in the C-terminus (Fig. 3B–C). α6 helix is perpendicular to a hydrophobic interface and stabilizes the other side of α4 and α5 helices. TbTFIIS2–2 LW domain possesses the key and conserved hydrophobic residues forming a hydrophobic core to mediate Leo1 interaction. The interactions between TbTFIIS2–2 LW domain and TbLeo1, including hydrophobic and hydrogen bond interactions, are shown in Fig. 3D–E. The hydrophobic pocket accommodating the TbLeo1 is mainly composed of several interactive residues that include Leu304, Leu307, Ile308, Leu336, Trp339, Phe340, Leu343, Val347 and Leu351 of the TbTFIIS2–2 LW domain (Fig. 3D). The residues of TbLeo1 involved in recognizing TbTFIIS2–2 LW domain include Leu72, Leu75, Phe76 and Phe80 (Fig. 3D). Notably, the side chains of TbLeo1 Phe76 insert into a deep hydrophobic pocket of the TbTFIIS2–2 LW domain, resulting in specific recognition between TbTFIIS2–2 and TbLeo1, which is a key to lock TFIIS and Leo1 (Fig. 3D). In addition, the interactions are also mediated by several hydrogen bonds formed between the amine hydrogen atom of TbTFIIS2–2 Lys333 and the carbonyl oxygen of TbLeo1 Phe81, and between the hydrogen atom of TbTFIIS2–2 Arg355 and the carbonyl oxygen of TbLeo1 Phe80 (Fig. 3E).

To further verify the importance of these interactive residues of TbTFIIS2–2 LW domain involved in TbLeo1 recognition, L307, K333, L336, F340, R355 and W339 in the TbLeo1-binding pocket were mutated individually to alanine or glutamate to abolish hydrogen bond formation or weaken the hydrophobicity, and their binding affinities toward TbLeo1 were evaluated (Fig. S5A). As shown in Table 3, all TbTFIIS2–2 LW domain mutants weaken the interactions between the TbTFIIS2–2 LW domain and the $TbLeo1_{71–86}$ peptide, particularly F340A and W339E, which exhibit >60-fold reductions in binding affinities (K_d = 9.33 μM and weak binding, respectively), compared with the wild-type TbTFIIS2–2 LW domain (K_d = 0.15 μM) (Fig. 2A), indicating that these residues of the TbTFIIS2–2 LW domain are indeed important for TbLeo1 binding. Meanwhile, mutation of LFG residues to AAA in $TbLeo1_{71–86}$ was observed to abolish its binding to the TbTFIIS2–2 LW domain, further confirming that the LFG motif is critical for TbLeo1 to interact with TbTFIIS2–2 (Fig. S5A). Taken together, these results revealed that a hydrophobic pocket of the TbTFIIS2–2 LW domain and the LFG motif of TbLeo1 are essential in the specific interactions between TbTFIIS2–2 and TbLeo1.

3.5. Conformational change of TbTFIIS2–2 LW domain upon binding to TbLeo1

The structures of the apo-form TbTFIIS2–2 LW domain (two conformation groups) were compared with that of the TbTFIIS2–2 LW domain in a complex with TbLeo1. The overall root mean square deviation (RMSD) of the Cα atoms between the two structures is only 0.41–0.55 Å (calculated by PyMOL) (Fig. 3A–B and Fig. 3F), suggesting that the global conformation of TbTFIIS2–2 LW domain is unchanged upon TbLeo1 binding.

The most significant difference between two structures lies in the α6 helix. The structural comparison indicated that TbLeo1 binding induces the TbTFIIS2–2 LW domain to undergo a conformational change that is reflected in the length and orientation of the α6 helix (Fig. 3F). In the apo form, the α6 helix is relatively short and its C-terminus is orientated toward the core five-helix bundle. Upon TbLeo1 binding, the α6 helix becomes longer and rotates about 90 degrees, leading to orientation of its C-terminus away from the core five-helix bundle. This conformational change further exposes and enlarges the hydrophobic pocket, which therefore facilitates the binding of TbLeo1 to the TbTFIIS2–2 LW domain. The two conformation groups of the apo-form TbTFIIS2–2 LW domain based on the difference of the length and orientation of α6 helix and the conformational change in the same region upon TbLeo1 binding suggest that the α6 helix of TbTFIIS2–2 LW domain is highly dynamic and favorable to TbLeo1 binding.

3.6. Crystal structure of the yeast ScTFIIS LW domain and solution structure of the human HsTFIIS LW domain

The ScTFIIS LW domain (residue 1–81 of ScTFIIS) from *S. cerevisiae* and the HsTFIIS LW domain (residue 1–96 of HsTFIIS) from *H. sapiens* were also expressed and purified. The structures of ScTFIIS and HsTFIIS LW domains were solved by X-ray crystallography and NMR spectrometry, respectively.

The crystal structure and structural statics parameters of ScTFIIS LW domain are shown in Fig. 4A–B and Table 1. The structure has been deposited into Protein Data Bank (7FAW). Similar to the TbTFIIS2–2 LW domain, the ScTFIIS LW domain also adopts a fold formed by a five-helix bundle, comprised of three long α-helices and two short α-helices, and the central five-helix bundle is connected by short loops. This structure is consistent with the previously reported structure of the ScTFIIS LW domain (residue 1–79 of ScTFIIS) [39]. The solution structure of the HsTFIIS LW domain was determined by a set of NMR spectra. The atomic co-ordinates for the 20 lowest-energy structures of the HsTFIIS LW domain have been deposited in the Protein Data Bank with the PDB ID code 7CNF. The solution structure and structural statics parameters of the HsTFIIS LW domain are summarized in Fig. 4C–D and Table 2, which show that the HsTFIIS LW domain also has a conformation consisting of a five-helix bundle that is similar to the TbTFIIS2–2 and ScTFIIS LW domains.

Comparison of the three structures of the TbTFIIS2–2, ScTFIIS and HsTFIIS LW domains indicated that they all display the highly conserved global fold. The TFIIS LW domains all consist of five or six α-helices, and all possess a conserved hydrophobic pocket (Fig. 4E). On the other hand, some distinct differences in the C-terminus were observed. The TbTFIIS2–2 LW domain contains one α-helix α6 in its C-terminus. However, the solution structure of the HsTFIIS LW domain shows that it possesses a disordered loop in its C-terminal counterpart. The crystal structure of ScTFIIS LW domain lacks the C-terminal counterpart. Therefore, in order to investigate whether the ScTFIIS LW domain contains one α-helix or a disordered loop in its C-terminus, a longer version of the ScTFIIS LW domain (residue 1–93 of ScTFIIS) was expressed, purified and studied by NMR spectrometry. The secondary structure of the ScTFIIS LW domain (longer version) was predicted by the

consensus chemical shift index (CSI) based on $^{1}H_{\alpha}$, $^{13}C_{\alpha}$ and $^{13}C_{\beta}$ chemical shifts [40]. The result indicated that the ScTFIIS LW domain contains a C-terminal loop instead of α-helix [41]. Therefore, different from TbTFIIS2–2, the C-termini of ScTFIIS and HsTFIIS LW domains both possess a disordered loop instead of an α-helix (Fig. 4C). Furthermore, the comparison of apo-TbTFIIS2–2 LW domain and TbTFIIS2–2 LW-TbLeo1 complex revealed that C-terminal loop in all TFIIS LW domains are dynamic (Fig. 3F and Fig. 4C). In the crystal structure of the TbTFIIS2–2 LW domain in a complex with TbLeo1, α6 in the TbTFIIS2–2 LW domain is involved in the interaction with TbLeo1. The structural differences of the C-termini in the TFIIS LW domains from different eukaryotes might be indicative of differential interaction affinities for Leo1.

3.7 The Leo1-binding surface of TFIIS LW domain in yeast and human

To further investigate whether the molecular interactions between TFIIS and Leo1 in yeast and human are similar to that in trypanosoma, NMR chemical shift perturbation was applied. The crystal structure reveals that the ScTFIIS LW domain lacks a counterpart of the TbTFIIS2–2 LW C-terminal α6 helix that is involved in the interaction with TbLeo1. Therefore, in order to investigate whether or not the C-terminal counterpart of the ScTFIIS LW domain participates in the ScLeo1 interaction, a longer version of the ScTFIIS LW domain (residue 1–93 of ScTFIIS) containing the presumptive C-terminal counterpart was expressed and purified for NMR study. NMR chemical shift perturbation of an ^{15}N-labeled ScTFIIS LW domain titrated with an increasing molar ratio of ScLeo1 was performed. The overlaid HSQC spectra of ^{15}N-labeled ScTFIIS LW domain showed a number of residues of the ScTFIIS LW domain with obvious chemical shifts after the addition of ScLeo1 (Fig. 5A). These residues include Glu43, Lys45, Gly47, Glu49, Val50, Lys54, Ile69, Ser71, Tyr72 and Asp74, suggesting their potential involvement in the ScLeo1 interaction. The residues are mainly located at the hydrophobic pocket formed by α3, α4 and α5, the loop that links α3 and α4, and the loop that links α4 and α5 (Fig. 5B–C).

In order to further verify the importance of these ScTFIIS LW domain residues for ScLeo1 binding, the conserved residues (L41, K66, I69 and W72) in the potential binding pocket were mutated individually to alanine and ITC assays were performed to evaluate their binding affinities for ScLeo1. As shown in Table 3 and Fig. S5B, all ScTFIIS LW domain mutants weaken the interactions between the ScTFIIS LW domain and the ScLeo1 peptide – particularly the W339A mutant, which exhibits significant reduction (weak binding) in binding affinity compared with wild-type ScTFIIS LW domain (Fig. 2B), confirming that these residues are critical for ScTFIIS to interact with ScLeo1.

ITC assays indicated that the HsTFIIS LW domain is able to interact with the $HsLeo1_{2–16}$ and $HsLeo1_{55–70}$ fragments. To further define the interaction surface of the HsTFIIS LW domain with $HsLeo1_{2–16}$ and $HsLeo1_{55–70}$, NMR chemical shift perturbation of an ^{15}N-labeled HsTFIIS LW domain, titrated with an increasing molar ratio of HsLeo1, was performed. The overlaid HSQC spectra of the ^{15}N-labeled HsTFIIS LW domain showed a number of HsTFIIS LW domain residues with obvious chemical shifts after the addition of $HsLeo1_{2–16/55–70}$ (Fig. 6A–B). These residues include Gly47, Ala52, Leu69, Ser72, Tyr73, Lys75, Leu76, Asp78 and Gly79, suggesting their potential involvement in the interaction.

The residues are mainly located at the hydrophobic pocket formed by α3, α4 and α5 and at the loop between α5 and C-terminus (Fig. 6C–E).

In the same way, the conserved residues (L41, K67, I70 and W73) in the potential binding pocket of HsTFIIS LW domain were mutated individually to alanine and ITC assays were performed to evaluate their effects on the binding affinities for $HsLeo1_{2–16}$ and $HsLeo1_{55–70}$, respectively. As shown in Table 3 and Fig. S5C, all HsTFIIS LW domain mutants impair the interactions between HsTFIIS LW domain and $HsLeo1_{2–16}$/$HsLeo1_{55–70}$ peptides, in particular the W339A mutation that abolishes the interactions, confirming these residues are critical for HsTFIIS to interact with HsLeo1.

Taken together, these results revealed that TFIIS LW domains from different eukaryotes all interact with Leo1 through a conserved hydrophobic pocket formed by some key conserved residues (Fig. 6E and Fig. 7A).

3.8. Key and conserved Leo1 residues involved in the Leo1 interaction with TFIIS

Previous studies showed that Leo1 interacts with the TFIIS LW domain through its LFG motif, and that the Phe residue in ‘LFG’ is the most important residue for the interactions between TbTFIIS2–2 and TbLeo1 in trypanosoma (Fig. 2A–D and Fig. 3D). This Phe acts like a wedge to lock the N-terminus of TbLeo1 to a groove located in TbTFIIS2–2 LW domain (Fig. 3D). Furthermore, the structural analysis of the complex revealed that the ‘LFG’ Phe actually plays a key role, while the ‘LFG’ Leu and Gly residues are not involved in the interaction (Fig. 3D). Therefore, in order to further verify whether the Phe residue in ‘LFG’ is essential for the interactions not only in trypanosoma but also in yeast and human, Leo1 fragments (residue 71–86 of TbLeo1, residue 54–68 of ScLeo1, residue 2–16 and 55–70 of HsLeo1) with ‘LFG’ mutated to ‘LAG’ were synthesized and analyzed for their binding affinities with TFIIS by ITC assays. The results showed that all these single-point mutations abolished the binding of Leo1 to TFIIS, suggesting that the Phe in ‘LFG’ is a key and conserved residue for Leo1 binding to TFIIS in all three eukaryotes (Fig. S6).

3.9. The molecular interactions between TFIIS and Leo1 are conserved in eukaryotes

Structural analyses, combined with ITC assays, *in vitro* mutagenesis and NMR chemical shift perturbations, revealed that in trypanosoma, yeast and human, TFIIS LW domains all interact with Leo1 through a conserved hydrophobic pocket formed by some key and conserved residues that include Leu307, Lys333, Leu336, and Trp339 in trypanosoma and their counterparts in yeast and human (Fig. 4E and Fig. 7A). The interactions were mediated by the highly conserved LFG motif in Leo1 from these three eukaryotes (Fig. 7B). In order to further explore whether this interaction mechanism is prevalent in all eukaryotes, sequence alignments from additional eukaryotes including *C. elegans*, *A. thaliana* and *M. musculus* were carried out. The analysis revealed that all eukaryotic TFIIS LW domains have corresponding conserved sites involved in Leo1 interaction (Fig. 4E and Fig. 7A), and that all eukaryotic Leo1 proteins possess a highly conserved LFG motif (Fig. 7B), suggesting that the molecular mechanism of the interactions between TFIIS and Leo1 is conserved in eukaryotes.

3.10. Species-specific variations in TFIIS LW domain and Leo1 interaction affinities

ITC assays revealed dissociation constants (K_d) for interactions between TFIIS LW domains containing the C-terminus (TbTFIIS2–$2_{251–359}$, ScTFIIS$_{1–93}$ and HsTFIIS$_{1–96}$) and Leo1 fragments (TbLeo1, ScLeo1 and HsLeo1–1/2) of 0.15 μM, 1.39 μM and 36.00/44.00 μM, respectively (Fig. 8A–D). These data indicate a significantly stronger interaction in trypanosoma than in yeast and human. Moreover, structural and NMR chemical shift perturbation studies showed that the C-terminal α6 in TbTFIIS2–2 is involved in the interaction, whereas the C-terminal counterpart (loop instead of helix) in ScTFIIS and HsTFIIS is not important in the interaction. In order to further investigate the importance of the C-termini of TFIIS in the interactions, truncated TFIIS LW domains without the C-termini (TbTFIIS2–$2_{251–343}$, ScTFIIS$_{1–81}$ and HsTFIIS$_{1–78}$) were constructed, expressed and purified. The dissociation constants (K_d) between these truncated TFIIS LW domains without their C-termini and the Leo1 fragments (TbLeo1, ScLeo1 and HsLeo1–1/2) were 51.70 μM, 3.30 μM and 116.00/57.60 μM, respectively (Fig. 8A–D). The TbTFIIS2–2 LW domain without the C-terminal α6 exhibits >300-fold reduction in TbLeo1-binding affinity compared with TbTFIIS2–2 LW domain containing the C-terminal α6. This reduced binding affinity is similar to those of ScTFIIS and HsTFIIS LW domains interacting with Leo1. However, the ScTFIIS and HsTFIIS LW domains without the C-terminal loops exhibit only approximately 1–3-fold reductions in binding affinities compared with ScTFIIS and HsTFIIS LW domains containing the C-terminal loops. The above results indicated that the C-terminal α6 of TbTFIIS2–2 LW domain is far more important than the C-terminal loop of ScTFIIS and HsTFIIS LW domains in the interactions with Leo1. This result is consistent with our structural analysis of apo-form TbTFIIS2–2 LW domain and TbTFIIS2–2 LW domain in complex with TbLeo1, which suggested that the α6 helix of TbTFIIS2–2 LW domain is highly dynamic and that the conformational change in this region facilitates the TbLeo1 binding. On the other hand, this can explain the stronger affinity of the TFIIS LW domain interacting with Leo1 in trypanosoma than in yeast and human.

3.11. The interactions between the TFIIS LW domain and Leo1 are critical for TFIIS binding to the Paf1C complex

To evaluate the functional importance of the interactions between the TFIIS LW domain and Leo1, GST pull-down and coimmunoprecipitation assays were carried out. Either HeLa nuclear extract containing HsPaf1C or purified HsPaf1C complex was incubated with wild-type or mutated (WT/L41A/L65A/I70A/W73A) GST-HsTFIIS LW domains (Fig. 9A–C). Bound proteins were then visualized by immunoblots. The mutated GST-HsTFIIS (L65A) LW domain was used as a negative control. GST pull-down and coimmunoprecipitation assays showed that all these mutations abolished GST-HsTFIIS binding to the HsPaf1C complex. Since these mutation sites are essential for the interactions between HsTFIIS LW domain and HsLeo1, these results indicate that the interactions between the HsTFIIS LW domain and HsLeo1 are critical for HsTFIIS to bind to the HsPaf1C complex.

4. Discussion

Transcriptional regulation is extremely important for many physiological events in eukaryotes, and regulation is manifested at both initiation and elongation steps. TFIIS

is a transcription elongation factor that can reactivate the backtracking of stalled Pol II to facilitate resumption of transcription elongation and therefore is crucial for productive transcription. Paf1C carries out multiple functions during transcription mediated by Pol II and is therefore another important transcription elongation factor [17]. The cooperative roles of TFIIS and Paf1C in transcription regulation have been indicated in different kinds of eukaryotes including human, yeast and plant [6,20,42]. In particular, previous biochemical studies indicated a strong synergy between TFIIS and Paf1C and an underlying mechanism involving direct TFIIS-Paf1C interactions and cooperative binding to Pol II [20]. Leo1, as a subunit of Paf1C, plays an especially important role through its direct TFIIS interaction in the cooperative binding of TFIIS and Paf1C to Pol II. Although recent cryo-EM structural studies of the Paf1C-TFIIS-Pol II complex from *H. sapiens* and *S. cerevisiae* displayed Paf1C-TFIIS interactions [4–6], the detailed structural basis and molecular mechanism of the interactions between Paf1C and TFIIS remain unclear due to the absence of the structure showing the direct interactions between Paf1C and TFIIS from *H. sapiens* and the low resolution of the structure of Paf1C complex from *S. cerevisiae*.

Of relevance to the present study, *T. brucei* counterparts to the eukaryotic elongation factors TFIIS and PAF1C have been identified [43–45]. Meanwhile, it is recently reported that TFIIS2–2 can interact with Paf1C and several Pol II subunits in *T. brucei* [46]. While the TFIIS counterpart TbTFIIS2–2 in *T. brucei* is not a typical TFIIS, it possesses a domain I (LW domain) similar to that of other TFIIS members. Unlike other TFIIS members, TbTFIIS2–2 possesses the LW domain in its C-terminus instead of in the N-terminus. In this study, TbTFIIS2–2 was found to interact with TbPaf1C subunit TbLeo1 through its LW domain. Our discovery in parasites demonstrated that this cooperation between TFIIS and Paf1C should be prevalent among eukaryotes.

Intriguingly, we identified a highly conserved motif (LFG) in TbLeo1 that is critical for TbLeo1 binding to the LW domain of TbTFIIS2–2. Furthermore, in three different eukaryotes including trypansoma, yeast and human, we revealed that the interactions between TFIIS and Paf1C are all mediated by the LW domain of TFIIS and a LFG motif in Leo1, suggesting these interactions might be conserved in eukaryotes. We further showed that these interactions are critical for TFIIS to bind to Paf1C complex, indicating their functional importance.

More importantly, our structural studies on the apo-form TbTFIIS2–2 LW domain and the TbTFIIS2–2 LW domain in a complex with TbLeo1 from *T. brucei*, the ScTFIIS LW domain from *S. cerevisiae* and the HsTFIIS LW domain from *H. sapiens* revealed a conserved molecular mechanism for the interactions between TFIIS and Leo1. TFIIS LW domains all adopt a conserved fold composed of a compact five-helix or six-helix bundle. Structural analyses, supported by NMR chemical shift perturbation, ITC, and site-directed mutagenesis analyses, revealed that TFIIS LW domains all interact with Leo1 through a conserved hydrophobic pocket formed by some key and conserved residues in trypanosoma, yeast and human. Moreover, the interactions are all mediated by a highly conserved LFG motif in Leo1 from these three eukaryotes.

Further sequence analysis on other eukaryotes besides trypanosoma, yeast and human indicated that despite relatively low sequence similarities of TFIIS and Leo1 among different eukaryotes, all eukaryotic TFIIS LW domains examined have corresponding conserved sites involved in Leo1 interactions and all eukaryotic Leo1 proteins examined possess a highly conserved LFG motif. These discoveries reveal that despite their evolution from lower to higher eukaryotes through divergent pathways, reflected by their low primary sequence similarities, TFIIS and Paf1C retain the most important and conserved motifs and residues through which they interact with each other.

Interestingly, the binding affinity between the TFIIS LW domain and Leo1 is much stronger in trypanosoma than that in yeast and human. Structural analysis showed that the TbTFIIS2–2 LW domain possesses a specific and highly dynamic α6 helix that interacts with TbLeo1. This interaction involves a specific hydrogen bond that provides a better anchor on TbLeo1. The comparison of the structures between the apo-form TbTFIIS2–2 LW domain and the TbTFIIS2–2 LW domain in a complex with TbLeo1 indicated that Tbleo1 binding induces the TbTFIIS2–2 LW domain to undergo a conformational change that is reflected in the length and orientation of helix α6, which will facilitate the TbLeo1 binding. In yeast and human, the corresponding C-terminus forms a flexible loop instead of a helix and is not important for Leo1 binding. These differences explain the higher binding affinity of the TFIIS LW domain interacting with Leo1 in trypanosoma than in yeast and human. Different from higher eukaryotes, the genome of *T. brucei* is mostly transcribed polycistronically by RNA polymerase II. In addition, unlike bacterial operons, genes in a polycistronic unit are not functionally related and most gene regulation occurs mainly at the post-transcriptional level in *T. brucei* [47]. Therefore, *T. brucei* possesses polycistronic units containing dozens to hundreds of genes, which requires a strong affinity among TFIIS, Paf1C and RNA polymerase to stabilize a long chain of pre-mRNA. The stronger affinity between TbTFIIS2–2 and TbLeo1, compared with that from other eukaryotes such as yeast and human, may explain how transcription elongation with such a long transcription distance is maintained in trypanosoma.

In conclusion, a molecular mechanism of the interactions between TFIIS and Paf1C complex has been revealed, and also may contribute to our understanding of the origin and evolution of transcription factor TFIIS and Paf1C complex in transcription regulation among eukaryotes.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We are grateful to the staffs at beamline BL17U1 at Shanghai Synchrotron Radiation Facility for their assistances in data collection. This work was supported by the Chinese National Natural Science Foundation (32071220), and the Key Project of Natural Science Research of Universities in Anhui Province (2022AH050743). Work at The Rockefeller University was supported by NIH grant CA273709 to R.G.R.

Data availability

Data will be made available on request.

References

[1]. Roeder RG, 50+ years of eukaryotic transcription: an expanding universe of factors and mechanisms, Nat. Struct. Mol. Biol 26 (9) (2019) 783–791. [PubMed: 31439941]

[2]. Noe Gonzalez M, Blears D, Svejstrup JQ, Causes and consequences of RNA polymerase II stalling during transcript elongation, Nat. Rev. Mol. Cell Biol 22 (1) (2021) 3–21. [PubMed: 33208928]

[3]. Antosz W, Deforges J, Begcy K, Bruckmann A, Poirier Y, Dresselhaus T, Grasser KD, Critical role of transcript cleavage in arabidopsis RNA polymerase II transcriptional elongation, Plant Cell 32 (5) (2020) 1449–1463. [PubMed: 32152189]

[4]. Farnung L, Ochmann M, Garg G, Vos SM, Cramer P, Structure of a backtracked hexasomal intermediate of nucleosome transcription, Mol. Cell 82 (17) (2022) 3126–3134.e7. [PubMed: 35858621]

[5]. Filipovski M, Soffers JHM, Vos SM, Farnung L, Structural basis of nucleosome retention during transcription elongation, Science (New York, N.Y.) 376 (6599) (2022) 1313–1316. [PubMed: 35709268]

[6]. Xu Y, Bernecky C, Lee CT, Maier KC, Schwalb B, Tegunov D, Plitzko JM, Urlaub H, Cramer P, Architecture of the RNA polymerase II-Paf1C-TFIIS transcription elongation complex, Nat. Commun 8 (2017) 15741.

[7]. Kim B, Nesvizhskii AI, Rani PG, Hahn S, Aebersold R, Ranish JA, The transcription elongation factor TFIIS is a component of RNA polymerase II preinitiation complexes, Proc. Natl. Acad. Sci. U. S. A 104 (41) (2007) 16068–16073. [PubMed: 17913884]

[8]. Guermah M, Palhan VB, Tackett AJ, Chait BT, Roeder RG, Synergistic functions of SII and p300 in productive activator-dependent transcription of chromatin templates, Cell 125 (2) (2006) 275–286. [PubMed: 16630816]

[9]. Kireeva ML, Hancock B, Cremona GH, Walter W, Studitsky VM, Kashlev M, Nature of the nucleosomal barrier to RNA polymerase II, Mol. Cell 18 (1) (2005) 97–108. [PubMed: 15808512]

[10]. Xu J, Chong J, Wang D, Strand-specific effect of Rad26 and TFIIS in rescuing transcriptional arrest by CAG trinucleotide repeat slip-outs, Nucleic Acids Res. 49 (13) (2021) 7618–7627. [PubMed: 34197619]

[11]. Allen BL, Taatjes DJ, The mediator complex: a central integrator of transcription, Nat. Rev. Mol. Cell Biol 16 (3) (2015) 155–166. [PubMed: 25693131]

[12]. Awrey DE, Shimasaki N, Koth C, Weilbaecher R, Olmsted V, Kazanis S, Shan X, Arellano J, Arrowsmith CH, Kane CM, Edwards AM, Yeast transcript elongation factor (TFIIS), structure and function. II: RNA polymerase binding, transcript cleavage, and read-through, J. Biol. Chem 273 (35) (1998) 22595–22605. [PubMed: 9712888]

[13]. Ling Y, Smith AJ, Morgan GT, A sequence motif conserved in diverse nuclear proteins identifies a protein interaction domain utilised for nuclear targeting by human TFIIS, Nucleic Acids Res. 34 (8) (2006) 2219–2229. [PubMed: 16648364]

[14]. Chu X, Qin X, Xu H, Li L, Wang Z, Li F, Xie X, Zhou H, Shen Y, Long J, Structural insights into Paf1 complex assembly and histone binding, Nucleic Acids Res. 41 (22) (2013) 10619–10629. [PubMed: 24038468]

[15]. Mueller CL, Porter SE, Hoffman MG, Jaehning JA, The Paf1 complex has functions independent of actively transcribing RNA polymerase II, Mol. Cell 14 (4) (2004) 447–456. [PubMed: 15149594]

[16]. Jaehning JA, The Paf1 complex: platform or player in RNA polymerase II transcription? Biochim. Biophys. Acta 1799 (5–6) (2010) 379–388. [PubMed: 20060942]

[17]. Tomson BN, Arndt KM, The many roles of the conserved eukaryotic Paf1 complex in regulating transcription, histone modifications, and disease states, Biochim. Biophys. Acta 1829 (1) (2013) 116–126. [PubMed: 22982193]

[18]. Francette AM, Tripplehorn SA, Arndt KM, The Paf1 complex: a keystone of nuclear regulation operating at the interface of transcription and chromatin, J. Mol. Biol 433 (14) (2021), 166979.

[19]. Tiwari V, Kulikowicz T, Wilson DM 3rd, Bohr VA, LEO1 is a partner for Cockayne syndrome protein B (CSB) in response to transcription-blocking DNA damage, Nucleic Acids Res. 49 (11) (2021) 6331–6346. [PubMed: 34096589]

[20]. Kim J, Guermah M, Roeder RG, The human PAF1 complex acts in chromatin transcription elongation both independently and cooperatively with SII/TFIIS, Cell 140 (4) (2010) 491–503. [PubMed: 20178742]

[21]. Cermakova K, Demeulemeester J, Lux V, Nedomova M, Goldman SR, Smith EA, Srb P, Hexnerova R, Fabry M, Madlikova M, Horejsi M, De Rijck J, Debyser Z, Adelman K, Hodges HC, Veverka V, A ubiquitous disordered protein interaction module orchestrates transcription elongation, Science (New York, N.Y.) 374 (6571) (2021) 1113–1121. [PubMed: 34822292]

[22]. Palenchar JB, Bellofatto V, Gene transcription in trypanosomes, Mol. Biochem. Parasitol 146 (2) (2006) 135–141. [PubMed: 16427709]

[23]. Lee JH, Nguyen TN, Schimanski B, Gunzl A, Spliced leader RNA gene transcription in Trypanosoma brucei requires transcription factor TFIIH, Eukaryot. Cell 6 (4) (2007) 641–649. [PubMed: 17259543]

[24]. Srivastava A, Badjatia N, Lee JH, Hao B, Gunzl A, An RNA polymerase II-associated TFIIF-like complex is indispensable for SL RNA gene transcription in Trypanosoma brucei, Nucleic Acids Res. 46 (4) (2018) 1695–1709. [PubMed: 29186511]

[25]. Wang QS, Pan QY, Liu K, Xu CY, Sun B, Zhou FH, Cui Y, Xu Yu Q, Earnest T, He JH, The Macromolecular Crystallography Beamline of SSRF 26 (2015) 10102.

[26]. Pannu NS, Waterreus WJ, Skubak P, Sikharulidze I, Abrahams JP, de Graaff RA, Recent advances in the CRANK software suite for experimental phasing, Acta Crystallogr. D Biol. Crystallogr 67 (Pt 4) (2011) 331–337. [PubMed: 21460451]

[27]. Kim DE, Chivian D, Baker D, Protein structure prediction and analysis using the Robetta server, Nucleic Acids Res. 32 (Web Server issue) (2004), W526–31. [PubMed: 15215442]

[28]. Adams PD, Grosse-Kunstleve RW, Hung L-W, Ioerger TR, McCoy AJ, Moriarty NW, Read RJ, Sacchettini JC, Sauter NK, Terwilliger TC, PHENIX: Building New Software for Automated Crystallographic Structure Determination 58(11), 2002, pp. 1948–1954.

[29]. Emsley P, Cowtan K, Coot: Model-building Tools for Molecular Graphics 60(12), 2004, pp. 2126–2132.

[30]. Delaglio F, Grzesiek S, Vuister GW, Zhu G, Pfeifer J, Bax A, NMRPipe: A Multidimensional Spectral Processing System Based on UNIX Pipes 6(3), 1995, pp. 277–293.

[31]. Lee W, Tonelli M, Markley JLJB, NMRFAM-SPARKY: Enhanced Software for Biomolecular NMR Spectroscopy 31(8), 2014, pp. 1325–1327.

[32]. Shen Y, Delaglio F, Cornilescu G, Bax A, TALOS+: a hybrid method for predicting protein backbone torsion angles from NMR chemical shifts, J. Biomol. NMR 44 (4) (2009) 213–223. [PubMed: 19548092]

[33]. Guntert P, Mumenthaler C, Wuthrich K, Torsion angle dynamics for NMR structure calculation with the new program DYANA, J. Mol. Biol 273 (1) (1997) 283–298. [PubMed: 9367762]

[34]. Koradi R, Billeter M, Wuthrich K, MOLMOL: a program for display and analysis of macromolecular structures, J. Mol. Graph 14 (1) (1996) (51–5, 29–32). [PubMed: 8744573]

[35]. Laskowski RA, MacArthur MW, Moss DS, Thornton JM, PROCHECK: a program to check the stereochemical quality of protein structures, J. Appl. Crystallogr 26 (2) (1993) 283–291.

[36]. Schwieters CD, Bermejo GA, Clore GM, Xplor-NIH for molecular structure determination from NMR and other data sources, Protein Sci. 27 (1) (2018) 26–40. [PubMed: 28766807]

[37]. Schwieters CD, Kuszewski JJ, Tjandra N, Clore GM, The Xplor-NIH NMR molecular structure determination package, J. Magn. Reson 160 (1) (2003) 65–73. [PubMed: 12565051]

[38]. Tian Y, Schwieters CD, Opella SJ, Marassi FM, A practical implicit solvent potential for NMR structure calculation, J. Magn. Reson. (San Diego, Calif.) 243 (2014) 54–64.

[39]. Booth V, Koth CM, Edwards AM, Arrowsmith CH, Structure of a conserved domain common to the transcription factors TFIIS, elongin A, and CRSP70, J. Biol. Chem 275 (40) (2000) 31266–31268. [PubMed: 10811649]

[40]. Wishart DS, Sykes BD, The 13C Chemical-Shift Index: a simple method for the identification of protein secondary structure using 13C chemical-shift data, J. Biomol. NMR 4 (2) (1994) 171–180. [PubMed: 8019132]

[41]. Gao J, Zhang J, Tu X, Liao S, (1)H, (13)C and (15)N resonance assignments of TFIIS LW domain from Saccharomyces cerevisiae, Biomol. NMR Assign 16 (1) (2022) 87–89. [PubMed: 35060010]

[42]. Antosz W, Pfab A, Ehrnsberger HF, Holzinger P, Kollen K, Mortensen SA, Bruckmann A, Schubert T, Langst G, Griesenbeck J, Schubert V, Grasser M, Grasser KD, The composition of the arabidopsis RNA polymerase II transcript elongation complex reveals the interplay between elongation and mRNA processing factors, Plant Cell 29 (4) (2017) 854–870. [PubMed: 28351991]

[43]. Ouna BA, Nyambega B, Manful T, Helbig C, Males M, Fadda A, Clayton C, Depletion of trypanosome CTR9 leads to gene expression defects, PLoS One 7 (4) (2012), e34256.

[44]. Wang R, Gao J, Zhang J, Zhang X, Xu C, Liao S, Tu X, Solution structure of TbTFIIS2–2 PWWP domain from Trypanosoma brucei and its binding to H4K17me3 and H3K32me3, Biochem. J 476 (2) (2019) 421–431. [PubMed: 30626613]

[45]. Uzureau P, Daniels JP, Walgraffe D, Wickstead B, Pays E, Gull K, Vanhamme L, Identification and characterization of two trypanosome TFIIS proteins exhibiting particular domain architectures and differential nuclear localizations, Mol. Microbiol 69 (5) (2008) 1121–1136. [PubMed: 18627464]

[46]. Staneva DP, Carloni R, Auchynnikava T, Tong P, Rappsilber J, Jeyaprakash AA, Matthews KR, Allshire RC, A systematic analysis of Trypanosoma brucei chromatin factors identifies novel protein interaction networks associated with sites of transcription initiation and termination, Genome Res. 31 (11) (2021) 2138–2154. [PubMed: 34407985]

[47]. Clayton CE, Life without transcriptional control? From fly to man and back again, EMBO J. 21 (8) (2002) 1881–1888. [PubMed: 11953307]

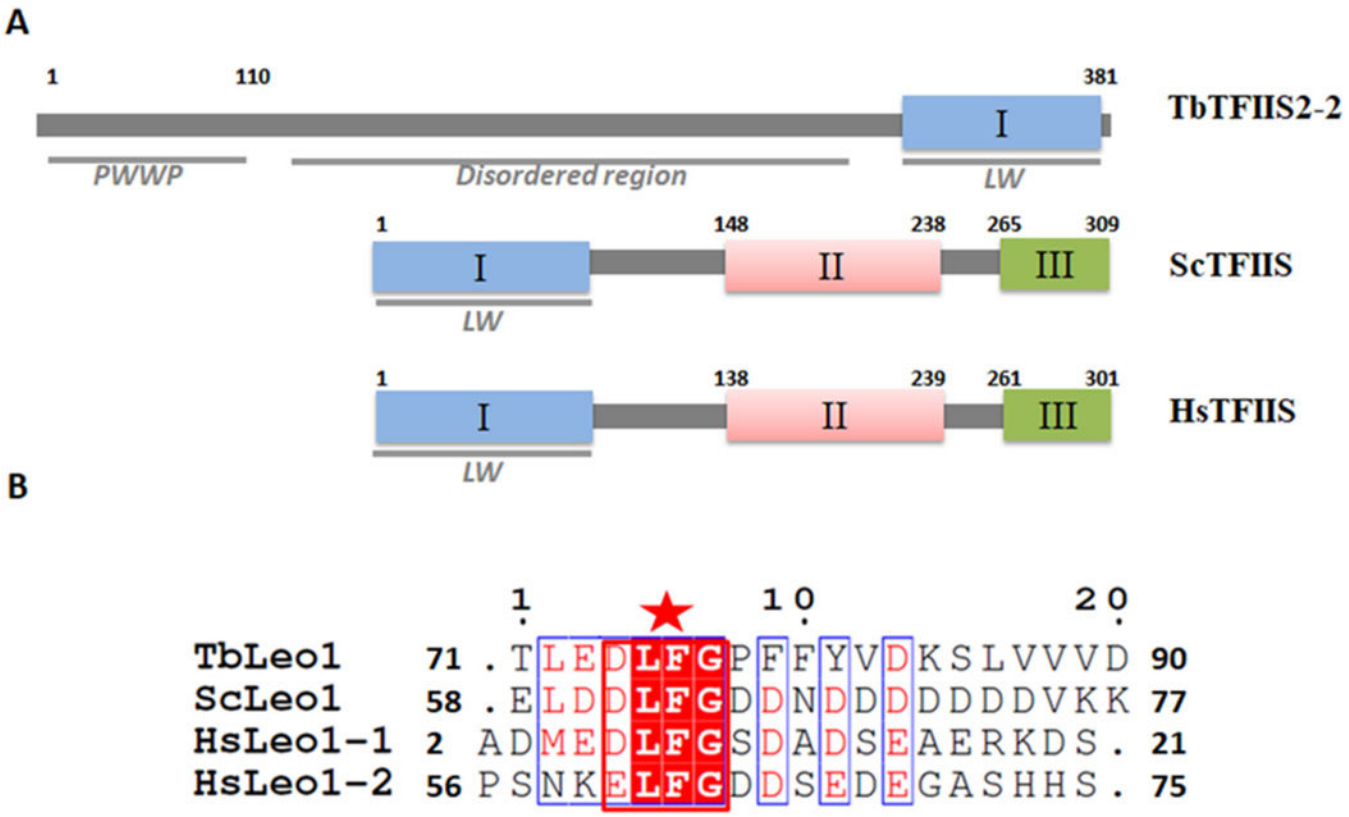

Fig. 1.

Domain architectures of TFIIS and multiple sequence alignments of Leo1. (A) Domain architectures of HsTFIIS, ScTFIIS and TbTFIIS2–2. (B) Multiple sequence alignments of the N termini of Leo1 containing the LFG motif from different eukaryotes. The LFG motif is marked with red colour.

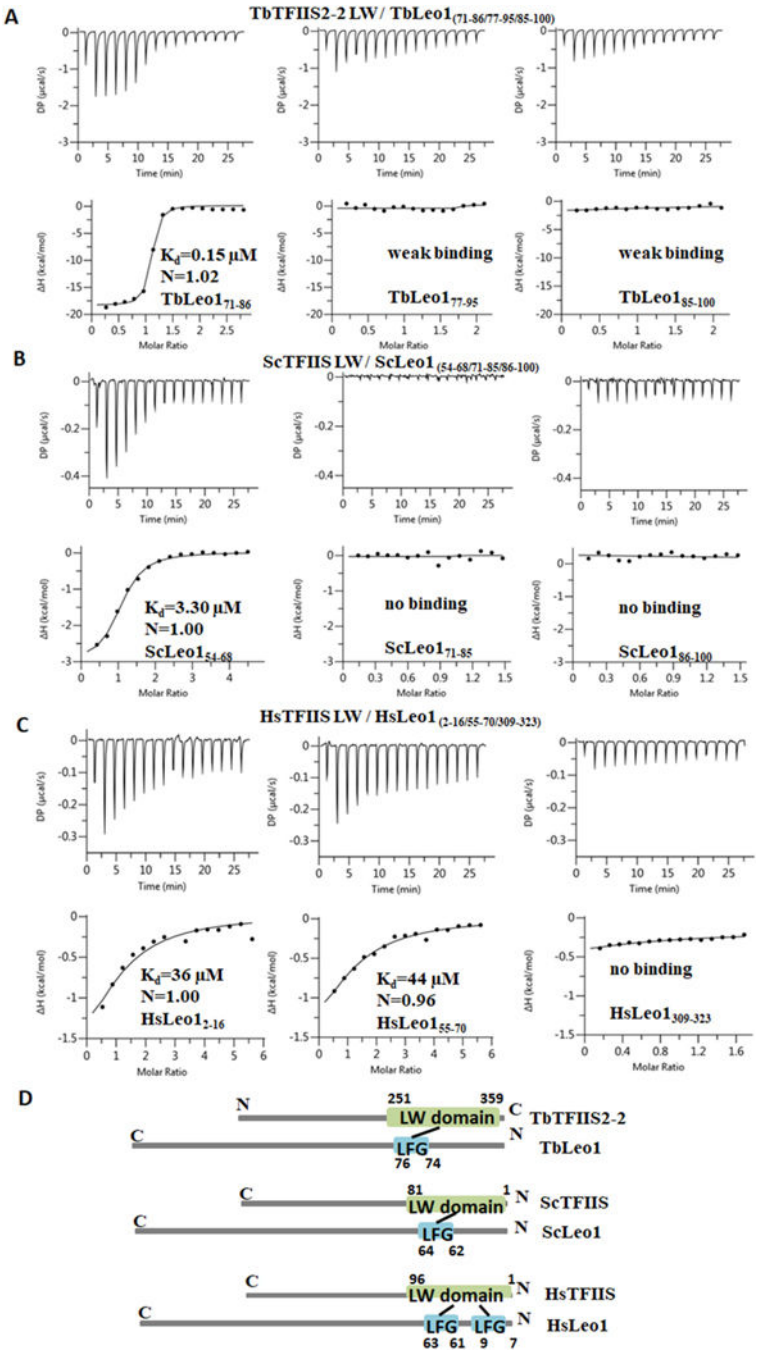

Fig. 2.

Interactions between the TFIIS LW domains and the Leo1 fragments analyzed by ITC. (A) Interactions between the TbTFIIS2–2 LW domain and the different fragments of TbLeo1. (B) Interactions between the ScTFIIS LW domain and the different fragments of ScLeo1. (C) Interactions between the HsTFIIS LW domain and the different fragments of HsLeo1. (D) Schematic depiction of the conserved interactions between the TFIIS LW domains and the Leo1 LFG motifs.

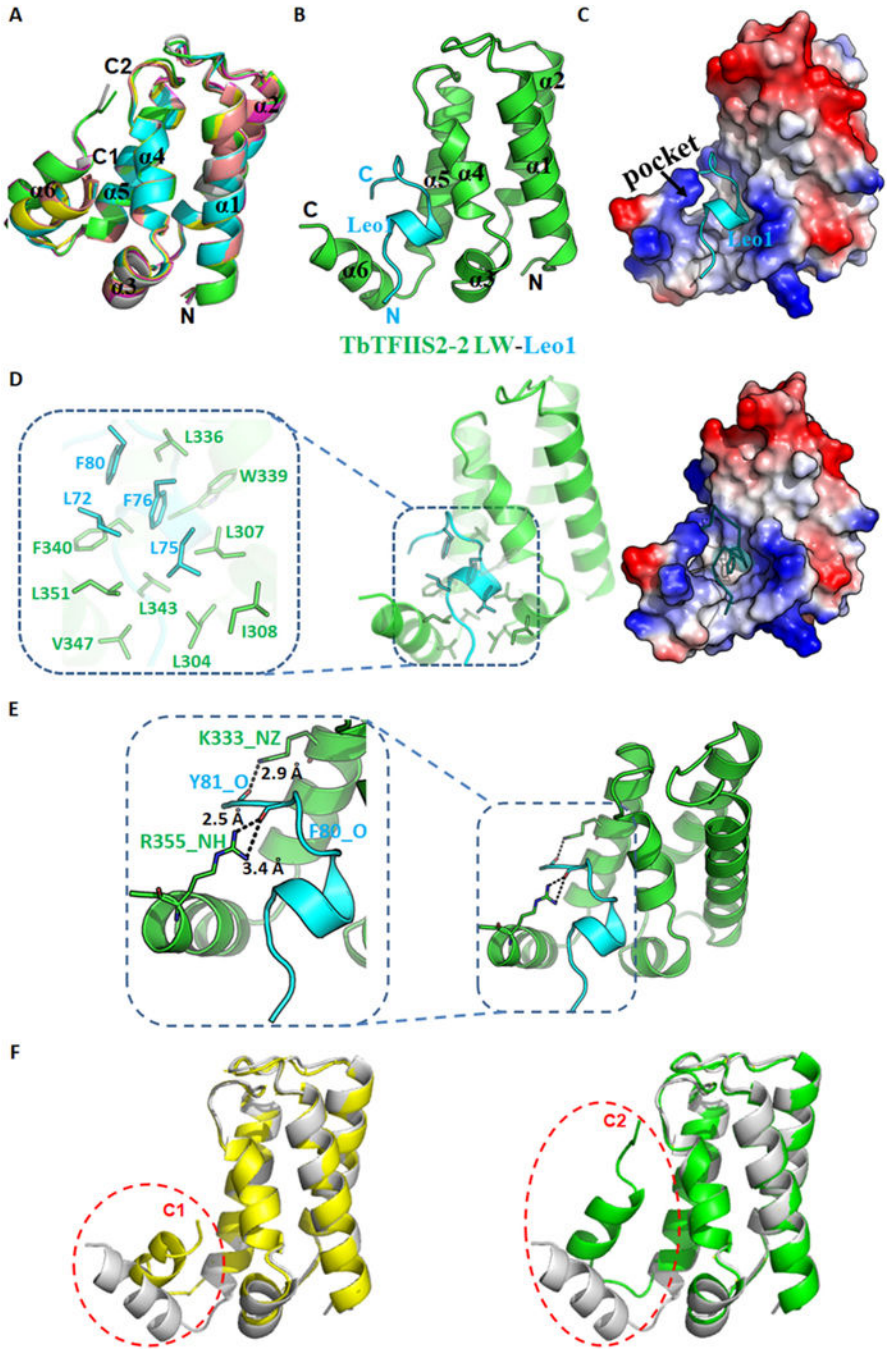

Fig. 3.

Crystal structure of the trypanosoma apo-form TbTFIIS2–2 LW domain, and the TbTFIIS2–2 LW domain in a complex with TbLeo1, and the molecular mechanism of their interactions. (A) Crystal structure of the apo-form TbTFIIS2–2 LW domain. Six molecules of the TbTFIIS2–2 LW domain are marked with different colors. (B) Overall structure of the TbTFIIS2–2 LW domain in a complex (green cartoon) with a TbLeo1 fragment (cyan cartoon). (C) The surface electrostatic potential of the TbTFIIS2–2 LW domain bound to TbLeo1. Residue 71–83 of TbLeo1 (cyan cartoon) are visible in the structure. (D) Detailed

hydrophobic interactions between the TbTFIIS2–2 LW domain and TbLeo1. (E) The details of hydrogen bond interactions between the TbTFIIS2–2 LW domain and TbLeo1. (F) Structural comparison between the apo-form TbTFIIS2–2 LW domain (C1: yellow cartoon and C2: green cartoon) and the TbTFIIS2–2 LW domain in a complex with TbLeo1 (gray cartoon). The red dotted ellipse indicates the different areas.

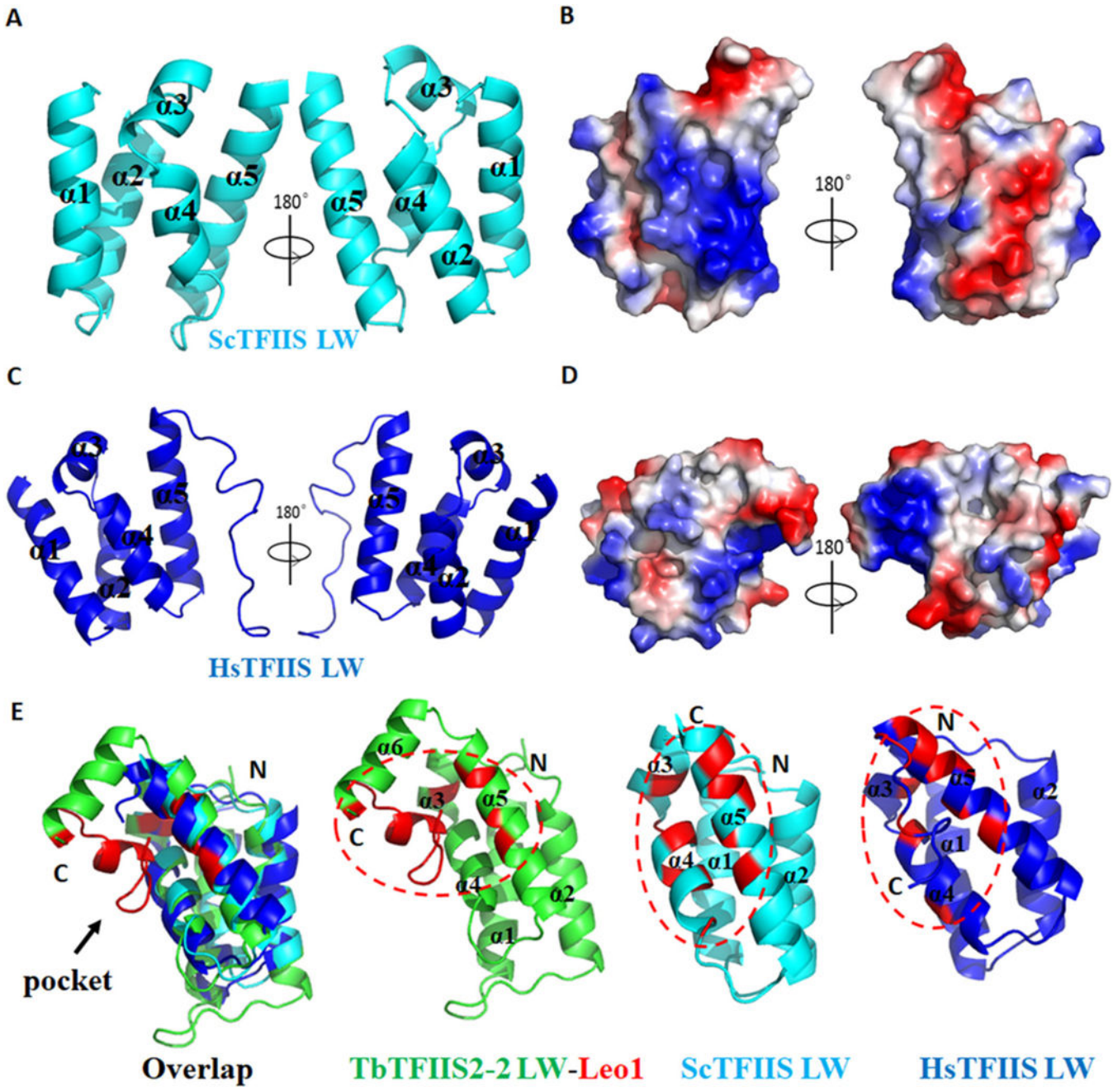

Fig. 4.

The structures of the yeast ScTFIIS and human HsTFIIS LW domains. (A) The crystal structure of the ScTFIIS LW domain (cyan cartoon). (B) The surface electrostatic potential of the ScTFIIS LW domain. (C) The solution structure of the HsTFIIS LW domain (blue cartoon). (D) The surface electrostatic potential of the HsTFIIS LW domain. (E) Structural comparison of the TbTFIIS2–2 LW domain with the ScTFIIS and HsTFIIS LW domains. Binding interfaces are colored in red in the ribbon representation. The red dotted ellipse indicates the binding areas. TbTFIIS2–2: green; TbLeo1: red; ScTFIIS: cyan; HsTFIIS: blue.

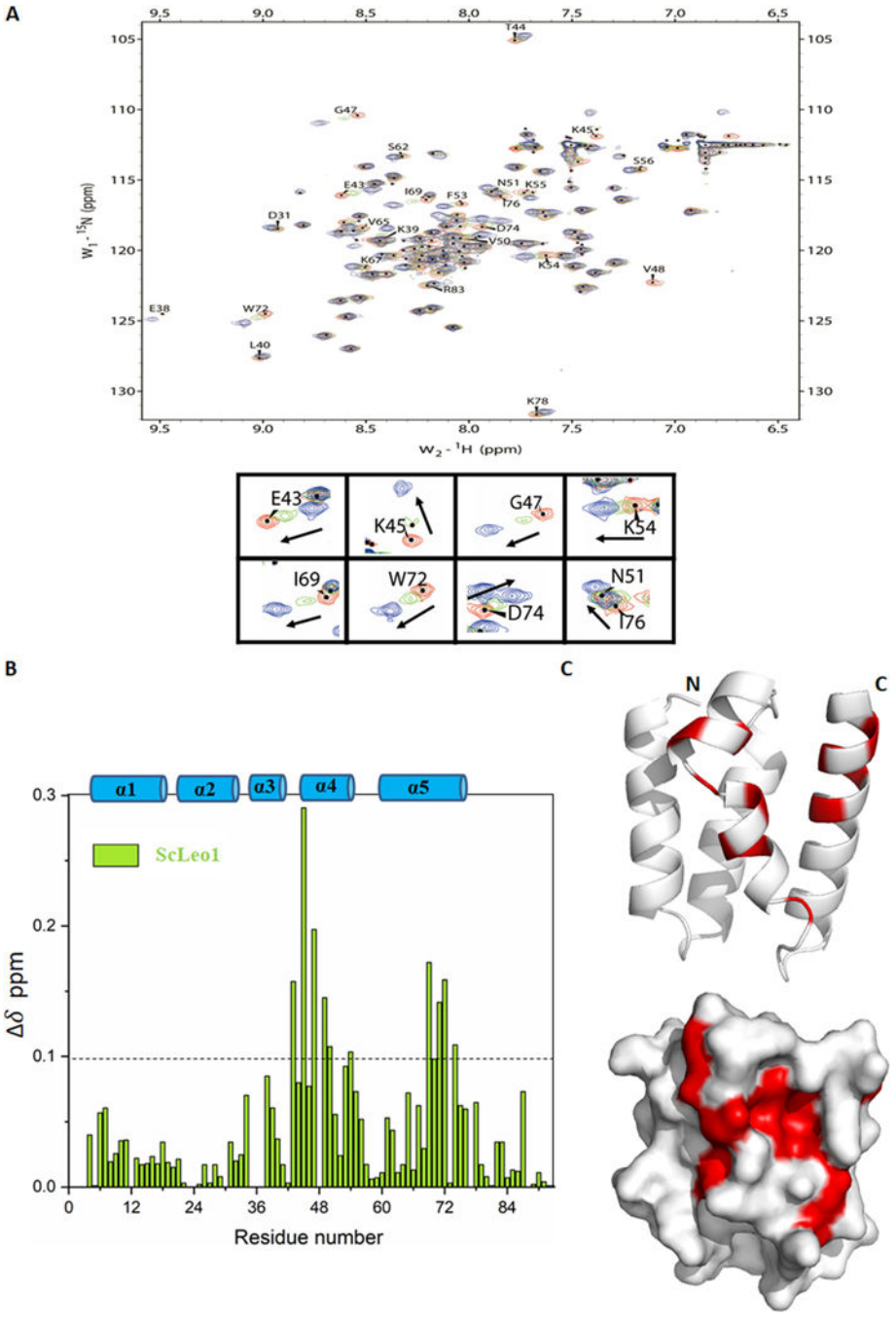

Fig. 5.

The Leo1-binding surface of the TFIIS LW domain from yeast. (A) NMR chemical shift perturbation of the ScTFIIS LW domain (longer version) titrated with unlabeled ScLeo1 at different molar ratios (Leo1/TFIIS) of 0 (red), 0.3 (green) and 1 (blue). (B) The chemical shift changes (Δδ)of the ScTFIIS LW domain (longer version) between the first and last titration point. The horizontal solid line represents the calculated average chemical shift perturbation. (C) Cartoon and surface representation of the ScTFIIS LW domain (longer

version). Residues with significant chemical shift changes upon the addition of ScLeo1 are colored in red.

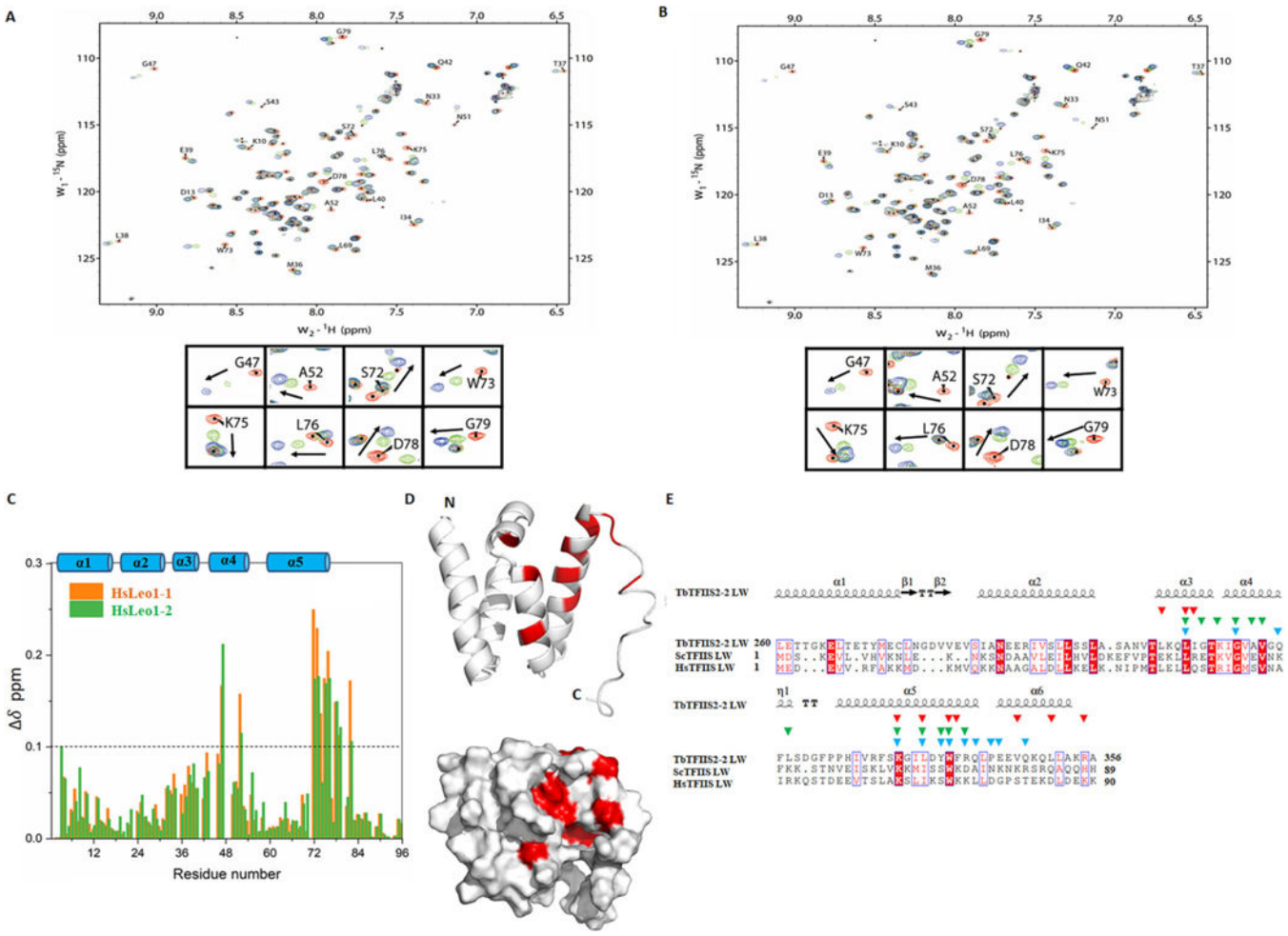

Fig. 6.

The Leo1-binding surface of the TFIIS LW domain from human, and the important residues for the TFIIS LW domain to interact with Leo1 from trypanosoma, yeast and human.

(A) NMR chemical shift perturbation of the HsTFIIS LW domain titrated with unlabeled HsLeo1–1 at different molar ratios (Leo1/TFIIS) of 0 (red), 0.3 (green) and 1 (blue).

(B) NMR chemical shift perturbation of the HsTFIIS LW domain titrated with unlabeled HsLeo1–2 at different molar ratios (Leo1/TFIIS) of 0 (red), 0.3 (green) and 1 (blue). (C) The chemical shift changes (Δδ) of the HsTFIIS LW domain between the first and last titration point. The horizontal solid line represents the calculated average chemical shift perturbation. (D) Cartoon and surface representation of the HsTFIIS LW domain. Residues with significant chemical shift changes upon the addition of HsLeo1 are colored in red.

(E) Multiple sequence alignments of the TFIIS LW domains from different eukaryotes, the secondary structure is above the sequence. The red, green and blue triangles indicate the residues which might be important for the TFIIS LW domain to interact with Leo1 from trypanosoma, yeast and human, respectively.

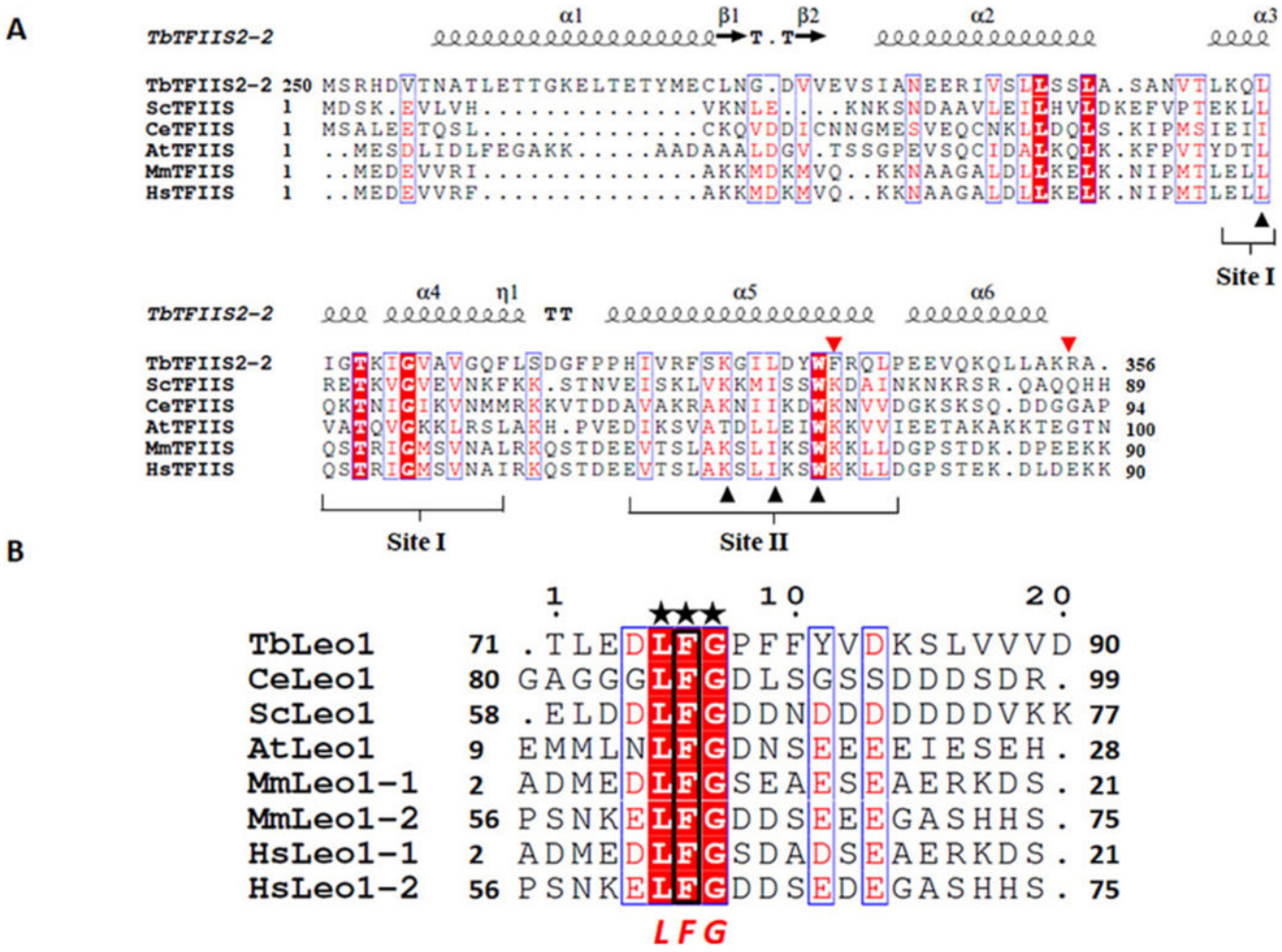

Fig. 7.

Multiple sequence alignments of the TFIIS LW domains and the N-terminal Leo1 from different eukaryotes. (A) Multiple sequence alignments of the TFIIS LW domains from different eukaryotes, the secondary structure is included at the top of the sequences. The bottom black triangle indicates the conserved residues critical for Leo1 interaction in the TFIIS LW domains from three species (trypanosoma, yeast, and human), and the upper red triangle indicates the residues important for TbLeo1 interaction specific in the TbTFIIS2–2 LW domain. (B) Multiple sequence alignments of the N-terminal Leo1 from different eukaryotes. Black boxes and five-pointed stars indicate conserved residues (leucine, phenylalanine and glycine).

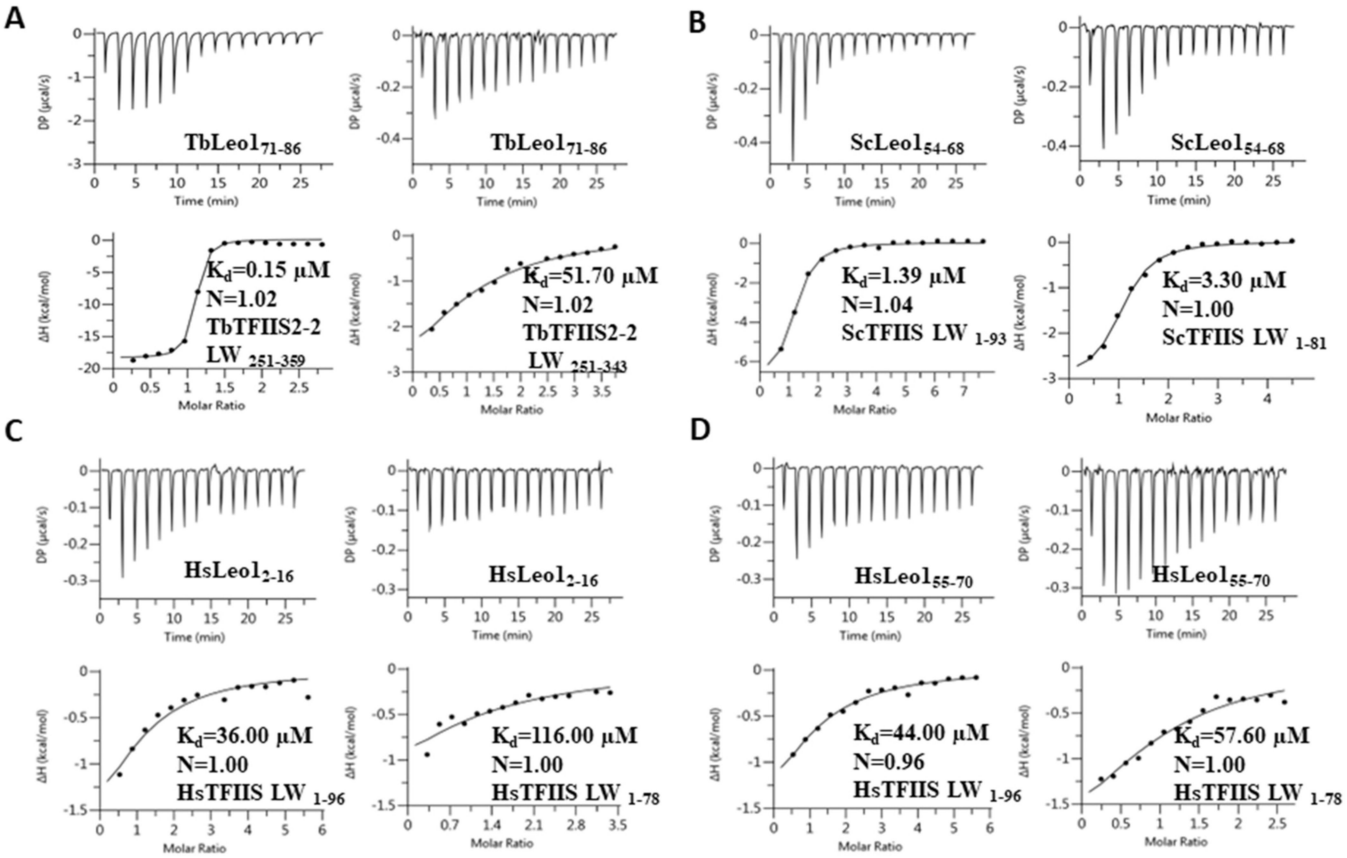

Fig. 8.

The different importance of the C-termini of the TFIIS LW domains in the interactions with Leo1. (A) ITC assays showed that the interactions between the TbTFIIS2–2 LW domain with (residue 251–359) or without the C-terminus (residue 251–343) and the TbLeo1 fragment. (B) ITC assays showed the interactions between the ScTFIIS LW domain with (residue 1–93) or without the C-terminus (residue 1–81) and the ScLeo1 fragment. (C)-(D) ITC assays showed that the interactions between the HsTFIIS LW domain with (residue 1–96) or without the C-terminus (residue 1–78) and the HsLeo1–1/2 fragments.

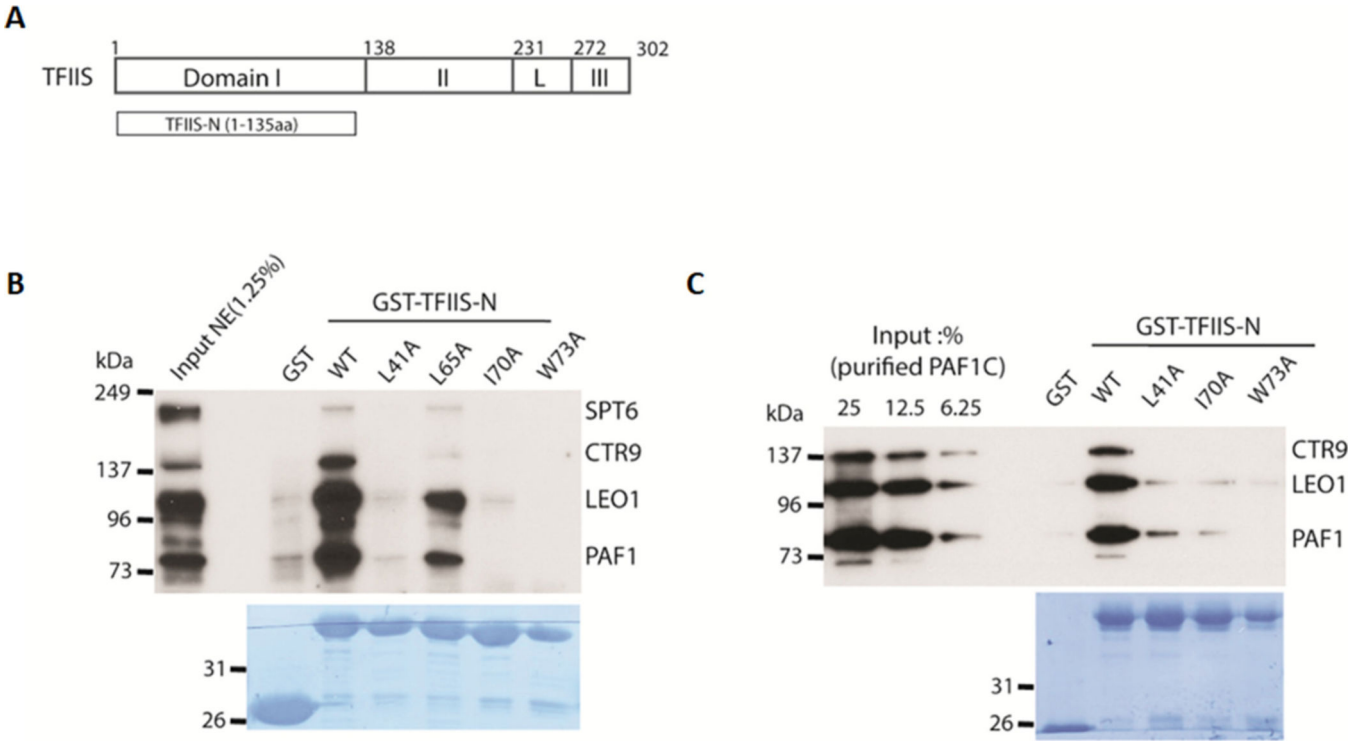

Fig. 9.

The interactions between the TFIIS LW domain and Leo1 are critical for TFIIS binding to the PAF1C complex. (A) Domain architectures of HsTFIIS. HeLa nuclear extract containing HsPaf1C (B) or purified Paf1C complex (C) was incubated with the wild-type or mutated (WT/L41A/L65A/I70A/W73A) GST-HsTFIIS N-terminal domain (LW domain, residue 1–135), respectively. Bound proteins were then visualized by immunoblot with related antibodies.

Table 1

Crystallographic data and refinement statistics.

	TbTFIIS2–2 LW-TbLeo1	apo-TbTFIIS2–2 LW	ScTFIIS LW
PDB Accession code	7FAX	7XGW	7FAW
Data collection Wavelength	0.9785	0.9791	0.9791
Space group	I4	P 32 2 1	P 2_1 2_1 2_1
Cell dimensions a,b,c (Å)	79.92, 79.92, 38.84	111.72, 111.72, 93.9	32.01, 91.57, 118.07
α,β,γ (°)	90, 90, 90	90, 90, 120	90, 90, 90
Resolution (Å)	50.00–1.60 (1.66–1.60)	48.38–2.25 (2.32–2.25)	118.07–2.44 (2.50–2.44)
R_{merge}	0.077(0.538)	0.101(1.111)	0.104(0.621)
$I/\sigma I$	75.7(3.4)	19.5(2.9)	20.2(4.8)
Completeness	100.0(100.0)	99.4(96.8)	100.0(100.0)
CC(1/2)	0.990(0.867)	0.999(0.798)	0.999(0.958)
Redundancy	12.8(13.2)	19.1(18.6)	14.0(14.4)
Refinement Resolution (Å)	21.97–1.80	42.27–2.25	36.18–2.44
No. of Unique reflections	11,535	32,323	13,624
R_{work}/R_{free}	0.2149/0.2384	0.1993/0.2299	0.2030/0.2350
Number of atoms/B-factor (Å2)			
Proteins	800/13.59	4452/55.0	2049/42.27
Peptide	110/10.31	N/A	N/A
Solvent	30/27.99	78/46.4	48/42.95
RMSD bonds (Å)/angles (○)	0.005/0.620	0.004/0.800	0.005/0.710
Ramachandran plot favored/allowed/outliers (%)	100/0/0	99/1/0	99.59/0.41/0

Statistics for the highest-resolution shell are shown in parentheses.

Table 2

NMR structural statistics of the HsTFIIS LW domain.

NMR restraints in the structure calculation	
Intraresidue	259
Sequential (|i - j| =1)	606
Medium-range (1 < |i - j| <5)	576
Long-range (|i - j| >/=5)	134
Hydrogen bonds	20
Total distance restraints	1595
Dihedral angle restraints	126
CYANA calculation	
Residual violations	
CYANA target functions, Å	0.29
NOE upper distance constrain violation Maximum, Å	0.14 ± 0.04
Number > 0.2 Å	0 ± 0
Dihedral angle constrain violations Maximum, °	1.66 ± 0.57
Number > 5 °	0 ± 0
Vander Waals violations Maximum, Å	0.26 ± 0.03
Number > 0.2 Å	1 ± 0.2
Average structural rmsd to the mean coordinates (Å)	
Secondary structure backbone[a]	0.64
All backbone atoms[b]	0.66
All heavy atoms[b]	1.17
XPLOR-NIH refinement	
TOTAL	
Energy	2097.60 ± 25.61
RMSD	0.354 ± 0.020
viols	0.8 ± 0.7
tDB Energy	3046.53 ± 27.81
viols eefxpot	0.8 ± 0.7
Energy	−1351.87 ± 11.19
NOE Energy	16.44 ± 2.78
RMSD	0.027(0.002
viols	0 ± 0
ANGL Energy	299.48 ± 6.86
RMSD	0.825 ± 0.009
viols	0 ± 0
BOND Energy	65.17 ± 1.42
RMSD	0.006 ± 0.000
viols	0 ± 0
CDIH Energy	2.17 ± 0.73

NMR restraints in the structure calculation	
RMSD	0.492 ± 0.088
viols	0 ± 0
IMPR Energy	19.67 ± 1.65
RMSD viols 0.0(0.0)	0.421 ± 0.018
Ramachandran statistics, % of all residues	
Most favored regions	92.0
Additional allowed regions	5.7
Generously allowed regions	2.3
Disallowed regions	0.0
Total number of residues	96

[a]Includes residues in secondary structure: 2–18, 21–33, 38–44, 46–56, 60–77.

[b]Obtained for residue M1-D78 since no long NOEs were identified for amino acids 79–96.

Table 3

The thermodynamic parameters of the ITC experiments.

Peptide(Tb)	**TbTFIIS2–2$_{251–359}$ LW**	**K_d μmol/l**	**N**
TbLeo1$_{71–86}$	Wild type	0.15	1.02
TbLeo1$_{71–86}$	TbTFIIS2–2$_{251–343}$ LW	51.70	1.02
TbLeo1$_{77–95}$	Wild type	weak binding	
TbLeo1$_{85–100}$	Wild type	weak binding	
TbLeo1$_{71–86}$ LFG(75–77)AAA	Wild type	no binding	
TbLeo1$_{71–86}$ F56A	Wild type	no binding	
TbLeo1$_{71–86}$	L307A	0.41	0.98
TbLeo1$_{71–86}$	K333A	0.52	0.90
TbLeo1$_{71–86}$	L336A	1.11	0.99
TbLeo1$_{71–86}$	W339E	weak binding	
TbLeo1$_{71–86}$	F340A	9.33	0.98
TbLeo1$_{71–86}$	R355A	0.25	1.06
Peptide(Sc)	ScTFIIS$_{1–81}$ LW	K_d μmol/l	N
ScLeo1$_{54–68}$	Wild type	3.30	1.00
ScLeo1$_{54–68}$	ScTFIIS LW$_{1–93}$	1.39	1.04
ScLeo1$_{71–85}$	Wild type	no binding	
ScLeo1$_{86–100}$	Wild type	no binding	
ScLeo1$_{54–68}$ F63A	Wild type	no binding	
ScLeo1$_{54–68}$	L41A	37.00	1.03
ScLeo1$_{54–68}$	K66A	5.10	0.95
ScLeo1$_{54–68}$	I69A	44.00	0.99
ScLeo1$_{54–68}$	W72A	weak binding	
Peptide(Hs)	HsTFIIS$_{1–96}$ LW	K_d μmol/l	N
HsLeo1$_{2–16}$	Wild type	36.00	1.00
HsLeo1$_{2–16}$	HsTFIIS$_{1–78}$ LW	116.00	1.00
HsLeo1$_{55–70}$	Wild type	44.00	0.96
HsLeo1$_{55–70}$	HsTFIIS$_{1–78}$ LW	57.60	1.00
HsLeo1$_{309–323}$	Wild type	no binding	
HsLeo1$_{2–16}$ F8A	Wild type	no binding	
HsLeo1$_{55–70}$ F62A	Wild type	weak binding	
HsLeo1$_{2–16}$	L41A	no binding	
HsLeo1$_{2–16}$	K67A	122.00	1.00
HsLeo1$_{2–16}$	I70A	no binding	
HsLeo1$_{2–16}$	W73A	no binding	
HsLeo1$_{55–70}$	L41A	no binding	
HsLeo1$_{55–70}$	K67A	weak binding	

Peptide(Tb)	**TbTFIIS2–2$_{251–359}$ LW**	**K_d μmol/l**	**N**
HsLeo1$_{55–70}$	I70A	no binding	
HsLeo1$_{55–70}$	W73A	no binding	