

# RNA polymerase III-specific general transcription factor IIIC contains a heterodimer resembling TFIIF Rap30/Rap74

Nicholas M. I. Taylor<sup>1</sup>, Florence Baudin<sup>1,2</sup>, Gudrun von Scheven<sup>1</sup> and Christoph W. Müller<sup>1,\*</sup>

<sup>1</sup>European Molecular Biology Laboratory (EMBL), Structural and Computational Biology Unit, Meyerhofstrasse 1, 69117 Heidelberg, Germany and <sup>2</sup>UJF-EMBL-CNRS UMI 3265, Unit of Virus Host-Cell Interactions, 38042 Grenoble Cedex 9, France

Received April 9, 2013; Revised June 9, 2013; Accepted July 8, 2013

## ABSTRACT

Transcription of tRNA-encoding genes by RNA polymerase (Pol) III requires the six-subunit general transcription factor IIIC that uses subcomplexes  $\tau$ A and  $\tau$ B to recognize two gene-internal promoter elements named A- and B-box. The *Schizosaccharomyces pombe*  $\tau$ A subcomplex comprises subunits Sfc1, Sfc4 and Sfc7. The crystal structure of the Sfc1/Sfc7 heterodimer reveals similar domains and overall domain architecture to the Pol II-specific general transcription factor TFIIF Rap30/Rap74. The N-terminal Sfc1/Sfc7 dimerization module consists of a triple  $\beta$ -barrel similar to the N-terminal TFIIF Rap30/Rap74 dimerization module, whereas the C-terminal Sfc1 DNA-binding domain contains a winged-helix domain most similar to the TFIIF Rap30 C-terminal winged-helix domain. Sfc1 DNA-binding domain recognizes single and double-stranded DNA by an unknown mechanism. Several features observed for A-box recognition by  $\tau$ A resemble the recognition of promoters by bacterial RNA polymerase, where  $\sigma$  factor unfolds double-stranded DNA and stabilizes the non-coding DNA strand in an open conformation. Such a function has also been proposed for TFIIF, suggesting that the observed structural similarity between Sfc1/Sfc7 and TFIIF Rap30/Rap74 might also reflect similar functions.

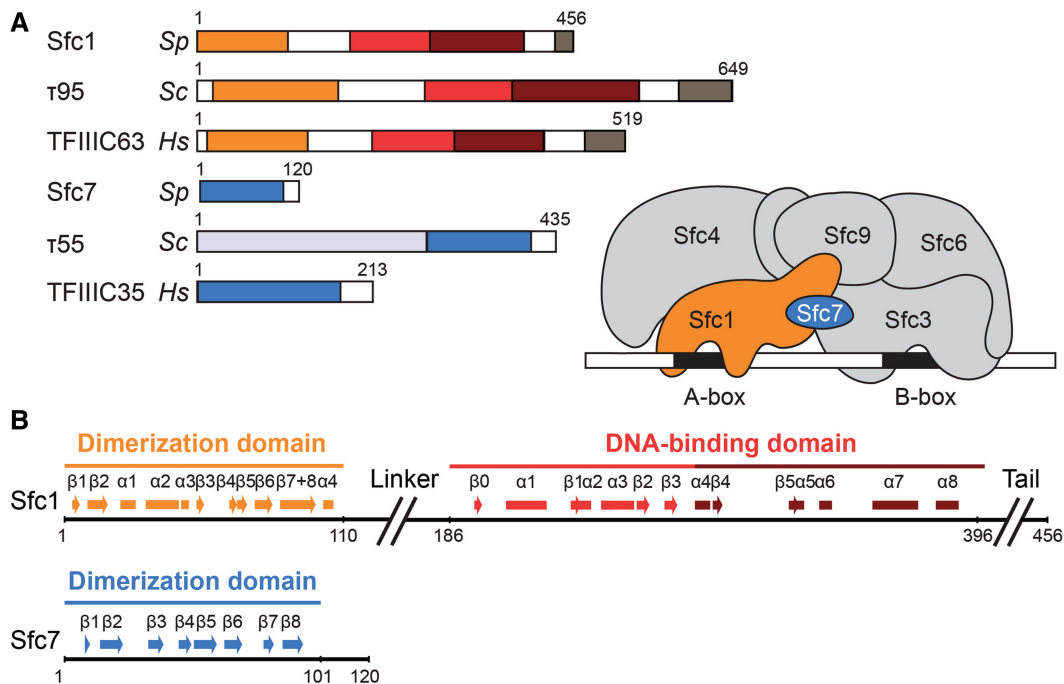
## INTRODUCTION

In most eukaryotes, three different RNA polymerases (Pol I, Pol II and Pol III) are responsible for the transcription of the genetic information. While Pol I synthesizes a

single rRNA precursor, Pol II synthesizes all mRNAs and most non-coding RNAs, and Pol III synthesizes a subset of small RNAs, most notably tRNAs, 5S rRNA and U6 snRNA. Pol III mainly uses type 2 promoters to transcribe tRNA genes. Type 2 promoters contain two gene-internal DNA-binding elements named A- and B-box that are recognized by the general transcription factor IIIC (TFIIIC) (1). Sequence comparisons have defined 11 and 9 nt consensus sequences for A- and B-box, respectively (2). During Pol III transcription initiation, TFIIIC binds in a first step the B-box of tRNA genes with high affinity, which in a second step then allows TFIIIC binding to the A-box that is less stringently conserved and only weakly bound (3). TFIIIC binding is followed by the recruitment of TFIIB and subsequently Pol III. As TFIIIC places TFIIB at a defined distance upstream of the TFIIIC-binding site, and TFIIB positioning determines the transcription start site, TFIIIC is involved in the selection of the transcription start site. The presence of TFIIIC in B-box-containing DNA regions has also been implicated in the spatial organization of the genomes from *Schizosaccharomyces pombe*, *Saccharomyces cerevisiae* and human, suggesting an additional role of TFIIIC as a chromosome-organizing clamp (4).

TFIIIC is a six-subunit protein complex of ~0.5 MDa mass consisting of two subcomplexes of similar size,  $\tau$ A and  $\tau$ B, which contact A-box and B-box promoter elements, respectively (5). The overall structural organization of the TFIIIC complex appears to be conserved from fungi to human, although conservation decreases when moving from the  $\tau$ A subcomplex close to the transcription start site towards the  $\tau$ B subcomplex further downstream of the start site (6). In the  $\tau$ A subcomplex, *S. cerevisiae* subunits  $\tau$ 131 (Tfc4),  $\tau$ 95 (Tfc1) and  $\tau$ 55 (Tfc7) correspond to *S. pombe* subunits Sfc4, Sfc1 and Sfc7, respectively, and to human subunits TFIIIC102, TFIIIC63 and TFIIIC35 (Figure 1A). In the  $\tau$ B subcomplex, *S. cerevisiae* subunits

\*To whom correspondence should be addressed. Tel: +49 6221 3878320; Fax: +49 6221 387519; Email: cmueller@embl.de



**Figure 1.** Overview of the TFIIC complex. (A) Domain representation of *S. pombe* subunits Sfc1 and Sfc7 and comparison with their *S. cerevisiae* and *Homo sapiens* orthologues. In Sfc1, τ95 and TFIIC63, the dimerization domain, winged-helix domain, WHI domain and the acidic tail are coloured in orange, red, dark red and grey, respectively. In Sfc7, τ55 and TFIIC35, the dimerization domain is coloured in blue, and the *S. cerevisiae* τ55-specific histidine phosphatase domain is in light blue. The right panel shows a schematic diagram of *S. pombe* TFIIC bound to type 2 promoter DNA as present in tRNA genes. The Sfc1/Sfc7 heterodimer is depicted in orange and blue, while all other TFIIC subunits are depicted in grey. (B) Domain structure of *S. pombe* Sfc1 and Sfc7. Secondary structure elements are indicated by bars ( $\alpha$ -helices) and arrows ( $\beta$ -sheets). The colour code corresponds to (A).

τ138 (Tfc3), τ91 (Tfc6) and τ60 (Tfc8) correspond to *S. pombe* subunits Sfc3, Sfc6 and Sfc9, respectively, and to human subunits TFIIC220, TFIIC110 and TFIIC90. In subcomplex τA, subunit τ95 is the only subunit with DNA-binding properties, suggesting it is involved in A-box recognition (7). Subunits τ95 and τ55 have both been cross-linked over the A-box of tRNA genes approaching the DNA from opposite sides (8). In addition, they have both been found to form a subcomplex of unknown function that occurs independently of holoTFIIC (9). The conserved C-terminal part of τ55 was shown to be sufficient for the τ95-τ55 interaction, whereas the non-conserved N-terminal histidine phosphatase domain is not required (10,11). In the τB subcomplex, subunits τ60 and τ91 both form WD40 propellers that are perpendicularly packed against each other (12). The third τB subunit, τ138 (Tfc3), directly contacts B-box DNA, possibly involving one or several winged-helix domains present in its N-terminal moiety.

The absence of detailed structural information on the τA complex of TFIIC has impaired our current understanding of Pol III pre-initiation complex (PIC) formation. We therefore structurally characterized the *S. pombe* TFIIC Sfc1/Sfc7 subcomplex and present crystal structures of two structural modules. The N-terminal moiety of Sfc1 (Sfc1-NT) and subunit Sfc7 heterodimerize through a triple  $\beta$ -barrel fold most similar to the dimerization domains of general transcription factor TFIIF Rap30/Rap74, whereas the C-terminal DNA-binding domain of Sfc1 (Sfc1-DBD) comprises a

winged-helix domain most similar to the winged-helix domain present in TFIIF Rap30 that is packed against a second domain with novel fold. The Sfc1/Sfc7 heterodimer and TFIIF Rap30/74 therefore share similar N-terminal dimerization modules and C-terminal winged-helix domains but also similar overall architectures. Further biochemical characterization identifies Sfc1-DBD as the central DNA-binding domain (DBD) of A-box DNA that recognizes single- and double-stranded DNA (ssDNA and dsDNA) and is autoinhibited by a ~60-residue C-terminal tail. Different DNA-binding modes of Sfc1-DBD to A-box DNA are principally possible. Comparison with primary  $\sigma$  factor domain 2 ( $\sigma$ 70 in *Escherichia coli*) suggests one possible DNA-binding mechanism, where Sfc1-DBD might unfold double-stranded A-box DNA and stabilise one DNA strand in an open conformation.

## MATERIALS AND METHODS

### Production of recombinant proteins

*Schizosaccharomyces pombe* Sfc1 and Sfc7 constructs were cloned into the bacterial expression vectors pETM-11 and pCDF-MCN, respectively, thereby introducing a tobacco etch virus (TEV) protease-cleavable N-terminal His-tag into the Sfc1 protein. Full-length Sfc1 and Sfc7 and the Sfc1-NT/Sfc7 dimerization module were co-expressed overnight at 18°C in BL21 (DE3) RIPL *E. coli* cells. The Sfc1/Sfc7 complex was purified by Ni-NTA affinity

chromatography, followed by TEV protease digestion overnight and a second Ni-NTA affinity chromatography step. Subsequently, the sample was diluted and loaded onto an anion exchange column (Mono Q HR10/10, GE Healthcare) followed by gel filtration (HiLoad 26/600 Superdex S200, GE Healthcare). Finally, the sample was concentrated, and limited proteolysis studies were performed (13). The Sfc1/Sfc7 dimerization module was purified using essentially the same procedure. The Sfc1-DBD (residues 186–396) was purified similarly to the Sfc1/Sfc7 interaction module using a cation exchange column (Q Sepharose High performance, GE Healthcare) instead of an anion exchange column, and without performing a gel filtration step. Further details for all purification procedures are given in the Supplemental Methods.

### Crystal structure determinations

The Sfc1<sub>1-110(K48A,Q49A)</sub>/Sfc7<sub>1-101</sub> complex was crystallized by the hanging-drop vapour diffusion at 4°C by mixing 5 µl of protein solution, at a concentration of 8.5 g/l, with 5 µl of reservoir solution [0.1 M HEPES (pH 7), 1.4 M ammonium sulphate] and adding 0.5 µl of 1 M KCl. Crystals were harvested and subsequently derivatized by soaking them in mother liquor (without KCl) supplemented with 1 mM para-chloro-mercuribenzenesulfonate before back-soaking them in mother liquor (without KCl) containing 30% glycerol and flash-frozen in liquid nitrogen. A single-wavelength anomalous diffraction data set of the Hg-derivatized crystals was collected at European Synchrotron Radiation Facility (ESRF) beamline ID23-1. The data were processed using programs XDS (14) and Scala (15). SHELXD (16) as implemented in HKL2MAP (17) was used to solve the substructure, and final phasing was performed using program SHARP (18). Density modification was done using SOLOMON (19), followed by automatic building of an initial model with ARP/wARP (20). The model was modified and extended manually using COOT (21), refinement with phenix.refine (22) and model validation with MolProbity (23). Selenomethionine-containing Sfc1-DBD was prepared by inhibiting the methionine biosynthesis pathway as described (24). Crystals grew at 20°C using hanging-drop vapour diffusion. 1 µl of protein solution, at a concentration of 36 g/l, was mixed with 2 µl of reservoir solution containing 0.2 M MgCl<sub>2</sub>, 0.1 M Tris-HCl (pH 8.5) and 20% (w/v) PEG 3350. The crystals were harvested and cryo-protected with mother liquor supplemented with 20% glycerol, before being flash-frozen in liquid nitrogen. A multiple-wavelength anomalous diffraction (MAD) data set was collected at ESRF beamline BM30A, and data processing was performed using program XDS. The crystal structure was determined as described for the Sfc1<sub>1-110(K48A,Q49A)</sub>/Sfc7<sub>1-101</sub> complex except that XSCALE was used instead of Scala, and the MAD method was used for experimental phasing. The SeMet data set collected at the peak of the absorption spectrum was used during the refinement.

### DNA-binding studies

Oligonucleotides containing the *S. pombe* A-box DNA site, the tRNA gene and unspecific DNA sequences,

used in the DNA-binding assays, are given in the Supplemental Methods. High pressure liquid chromatography-purified oligonucleotides were labelled with <sup>32</sup>P by T4 polynucleotide kinase and gel-purified on denaturing 15% urea-PAGE. Radioactively labelled <sup>32</sup>P oligonucleotides were first annealed by heating in water at 95°C for 1 min, cooled down to 25°C and incubated in buffer D [20 mM Tris-HCl (pH 7.5), 200 mM KCl, 2 mM MgCl<sub>2</sub>, 10 mM DTT]. For the electrophoretic mobility shift assays (EMSA) <sup>32</sup>P-labelled DNA was incubated with Sfc1-DBD constructs in buffer D. After incubation at 25°C, the samples were loaded on a native 6% acrylamide gel and run at 160 V in Tris-glycine buffer. For the competition experiments, <sup>32</sup>P-labelled A-box was first bound to 1 µM Sfc1-DBD. Then, increasing concentrations of Sfc1<sub>434-456</sub> peptide and control peptides (1.10<sup>-6</sup> to 3.10<sup>-4</sup> M) were added to challenge the complex. After incubation at 25°C for 30 min, the reactions were loaded on a native 6% acrylamide gel. A blank reaction and a control without peptide were also carried out. After migration, the native gel was dried and autoradiographed for 2 h with X-ray film (Biomax, MR-film, Kodak).

For filter-binding assays, proteins were serially diluted (from 100 µM to 0.1 nM) in a final volume of 180 µl of buffer D, followed by addition of 20 µl of renatured <sup>32</sup>P-labelled DNA (~5000 cpm). After incubation at 25°C, the reactions were filtered through a nitrocellulose filter and the radioactivity counted. For each series, a blank reaction and the total DNA input were determined. Because the amount of labelled DNA is negligible, an apparent dissociation coefficient ( $K_D$ ) can be estimated from the concentration value at 50% DNA retention (25).

### Isothermal titration calorimetry

Isothermal titration calorimetry (ITC) measurements were performed at 25°C using a VP-ITC Microcal calorimeter (Microcal, Northampton, USA). Protein was dialysed overnight against ITC buffer [50 mM HEPES (pH 7.5), 150 mM NaCl, 2 mM β-mercaptoethanol]. The lyophilized peptide was re-suspended in ITC buffer. The titration experiments consisted of injecting 10 µl of a 200 µM peptide solution into 2 ml of a 10 µM protein solution at time intervals of around 5 min.

## RESULTS

### Sfc1/Sfc7 heterodimer consists of two separate modules

The *S. pombe* Sfc1/Sfc7 heterodimer was co-expressed in *E. coli* and purified to homogeneity. Limited proteolysis to investigate the domain organization of the Sfc1/Sfc7 heterodimer resulted in three different species: Sfc1 ( $M_r = 52.7$  kDa) was rapidly cleaved into two protease resistant fragments, one similar in size to Sfc7 ( $M_r = 13.3$  kDa) and one ~25 kDa in size, whereas Sfc7 was resistant to proteolysis. Sfc7 and the smaller Sfc1 fragment remained stably associated during ion-exchange chromatography, unlike the second larger Sfc1 fragment. Electrospray mass spectrometry in combination



with peptide mass fingerprinting allowed the assignment of the smaller Sfc1 fragment interacting with Sfc7 to residues 1–107 (Sfc1-NT) and the larger Sfc1 fragment to residues 186–396 (corresponding to Sfc1-DBD). In contrast, the residues that connect Sfc1-NT with Sfc1-DBD and the C-terminal extension following Sfc1-DBD are readily degraded (Figure 1B). The Sfc1-NT/Sfc7 heterodimer did not interact with the larger fragment during gel filtration, indicating that the Sfc1/Sfc7 heterodimer consists of two separate globular modules connected by a protease sensitive linker.

### Dimerization modules of Sfc1/Sfc7 and TFIIF Rap30/Rap74 are similar

Well-diffracting crystals were obtained for a Sfc1<sub>1-110(K48A,Q49A)</sub>/Sfc7<sub>1-101</sub> construct, where two non-conserved Sfc1 residues (K48 and Q49) were mutated into alanine following the suggestion of the surface entropy reduction prediction (SERp) server (26). The structure was solved at 2.4 Å by single-wavelength anomalous diffraction using para-chloro-mercuribenzenesulfonate as heavy-atom derivative (Table 1). The Sfc1-NT/Sfc7 dimerization module consists of three interwoven  $\beta$ -barrels containing 15  $\beta$ -strands and four  $\alpha$ -helices (Figure 2A and C). The interwoven structure suggests that the subunits fold co-operatively and need to be simultaneously expressed. Indeed, in this study, Sfc1<sub>1-110(K48A,Q49A)</sub> and Sfc7<sub>1-101</sub> (as well as full-length Sfc1 and Sfc7) were co-expressed in *E. coli* because it had been observed that the expression of isolated full-length Sfc1 resulted in insoluble protein. The crystal structure also confirmed that the mutated residues K48 and Q49 lay at the surface of the Sfc1 dimerization domain and do not interfere with Sfc1–Sfc7 heterodimerization.

**Table 1.** Crystallographic statistics for the Sfc1-NT/Sfc7 complex

Sfc1-NT/Sfc7 complex (Hg-derivative)	
Data collection	
Space group	P6 <sub>3</sub> 22
Cell dimensions	a, b = 85.3
a, b, c (Å)	c = 154.1
Wavelength (Å)	1.00685
Resolution (Å)	42.6–2.40 (2.53–2.40)
R <sub>sym</sub> (%)	10.7 (88.8)
<I/σ(I)>	14.5 (2.1)
Completeness (%)	99.9 (99.6)
Multiplicity	12.1 (9.4)
Refinement	
Resolution (Å)	33.32–4.0
No. reflections	24,503
R <sub>work</sub> /R <sub>free</sub> (%)	18.1/22.4
No. atoms	
Protein	1534
Other	3
Water	42
B-factor (Å <sup>2</sup> )	
Protein	56.1
Other	92
Water	49.7
R.m.s. deviations	
Bond lengths (Å)	0.011
Bond angles (°)	1.318

According to program Dali (27), the structure shares the highest similarity with the TFIIF Rap30/Rap74 dimerization domains (28). Sfc1 and Sfc7 resemble Rap30 and Rap74, respectively ( $Z$ -score = 9.3, rmsd<sub>154CA</sub> = 4.7 Å) (Figure 2B). Because of the pseudo 2-fold symmetry of the triple-barrel fold, Sfc1 and Sfc7 also superimpose with Rap74 and Rap30, respectively, although with lower  $Z$ -score. Despite the readily apparent structural similarity between Sfc1 and Rap30 and Sfc7 and Rap74, Sfc1 helices  $\alpha$ 1 to  $\alpha$ 3 and  $\beta$ -strands  $\beta$ 3 and  $\beta$ 4 replace Rap30 strands  $\beta$ 3 and  $\beta$ 4 (approximately residues 20–68), whereas the ‘arm domain’ of Rap74 is absent in Sfc7. The Sfc1-NT/Sfc7 dimerization domain can also be superimposed with the dimerization modules of Pol I subunits A49/A34.5 ( $Z$ -score = 7.7, rmsd<sub>142CA</sub> = 4.2 Å) and mouse RNase H2B/H2C ( $Z$ -score = 7.8, rmsd<sub>134CA</sub> = 3.7 Å).

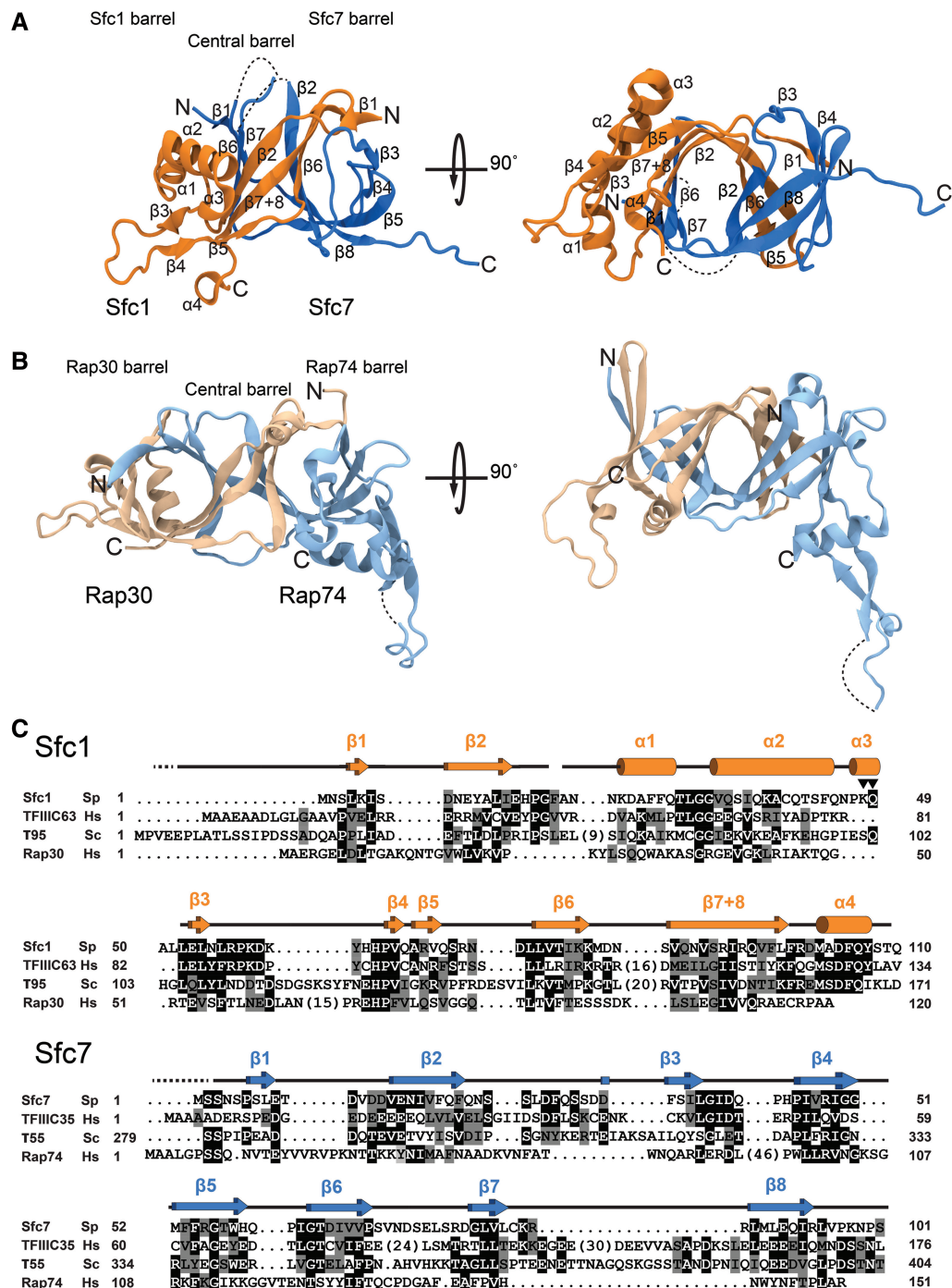
### Sfc1-DBD and TFIIF Rap30 contain similar winged-helix domains

The crystal structure of the C-terminal domain of Sfc1 (Sfc1-DBD, residues 186–396) was solved at 1.45 Å resolution by MAD using selenomethionine-substituted protein (Table 2). The Sfc1-DBD structure forms a compact domain that can be further divided into two subdomains (Figure 3A and D). The first domain (residues 186–283) adopts a winged-helix domain fold with canonical  $\alpha$ 1- $\beta$ 1- $\alpha$ 2- $\alpha$ 3- $\beta$ 2- $\beta$ 3 topology. One additional strand ( $\beta$ 0) pairs strand  $\beta$ 2 of the winged-helix domain in an anti-parallel manner. The third helix  $\alpha$ 3 is not continuous, but split into two short helices  $\alpha$ 3a and  $\alpha$ 3b. The second domain (residues 284–396) consists of five  $\alpha$ -helices and two  $\beta$ -strands and adopts a novel fold with no significant similarity to any known structure. Subsequently, we will refer to this domain as winged-helix-interacting (WHI) domain. Structural comparison using program DALI identifies the winged-helix domain of Rap30 (29) as the structure with the closest similarity to the winged-helix domain of Sfc1 ( $Z$ -score = 6.1, rmsd<sub>61CA</sub> = 2.7 Å) among the many winged-helix domains present in the PDB database (Figure 3C). Sfc1 therefore resembles TFIIF Rap30 in its N-terminal dimerization domain and in its C-terminal DBD.

### Sfc1-DBD binds DNA with an extended basic region

Inspection of the Sfc1-DBD electrostatic surface potential identifies an extended basic region on one side of the Sfc1-DBD that could interact with DNA (Figure 3B), whereas the opposite side appears rather uncharged. The basic interaction surface of Sfc1-DBD shows two separate positively charged patches. One continuous larger patch runs diagonally across the surface of Sfc1-DBD and predominantly comprises basic residues of helices  $\alpha$ 2 and  $\alpha$ 3 of the winged-helix domain. The second smaller patch comprises residue R269 of the winged-helix domain and residues K294 and K296 of the WHI domain (Figure 3B). Almost all residues forming both basic patches are conserved between TFIIC Sfc1 orthologues (Figure 3D).





**Figure 2.** Structure of the Sfc1/Sfc7 dimerization domains. (A) Two orthogonal views of the Sfc1-NT/Sfc7<sub>1-101</sub> dimerization domains. Sfc1-NT is depicted in orange, whereas Sfc7<sub>1-101</sub> is depicted in blue. Secondary structure elements and N- and C-termini are labelled. Disordered loops are represented by dashed lines in (A) and (B). (B) Two orthogonal views of the TFIIIF Rap30/Rap74 dimerization domains. Rap30 is depicted in light orange, and Rap74 is depicted in light blue. The orientation is the same as for the Sfc1/Sfc7 dimerization domains in (A). (C) Sequence alignment of the N-terminal moieties of *S. pombe* Sfc1 (*Sp*), human TFIIIC63 (*Hs*), *S. cerevisiae* τ95 (*Sc*) and human TFIIIF Rap30 (*Hs*), and of *S. pombe* Sfc7 (*Sp*), human TFIIIC35 (*Hs*), the C-terminal part of *S. cerevisiae* τ55 (*Sc*) and human TFIIIF Rap74 (*Hs*). The α-helices and β-sheets are depicted as cylinders and arrows, respectively. Sequence conservation of identical and similar residues is represented by dark and light grey shading, respectively. Arrowheads indicate Sfc1 K48 and Q49 residues that were mutated to alanine to reduce surface entropy.

In a next step, we tested binding of Sfc1-DBD to dsDNA and ssDNA DNA oligonucleotides using electrophoretic mobility shift assays (EMSA) (Figure 4A) and filter binding experiments (Figure 4B).

Sfc1-DBD binds 25bp dsDNA as well as ssDNA. Consistent with a minor contribution of A-box binding to the overall binding affinity of TFIIIC (30), competition experiments and filter binding assays showed only μmolar

**Table 2.** Crystallographic statistics for Sfc1-DBD

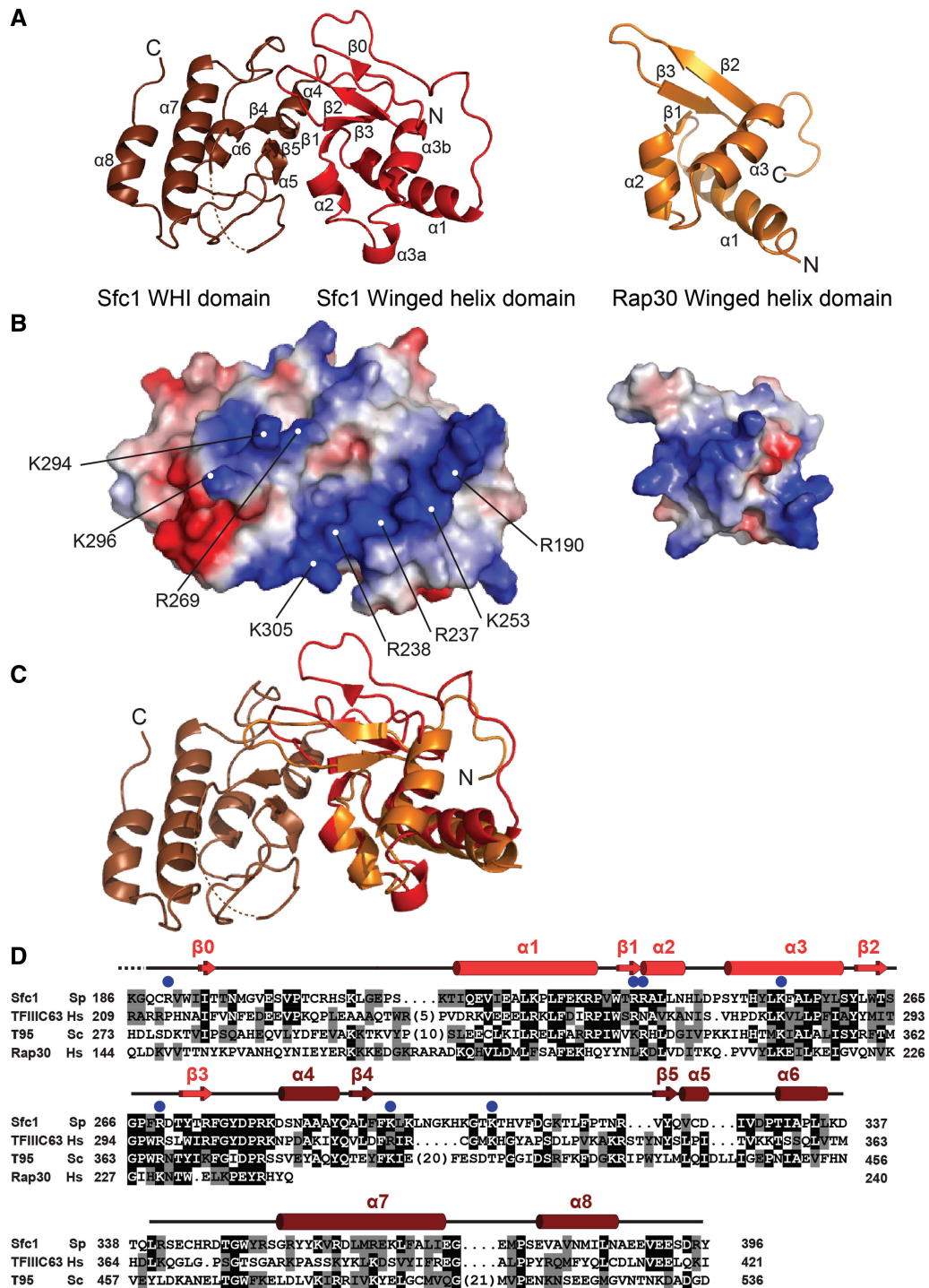
	SeMet Sfc1-DBD		
	Peak	Inflection point	Remote
Data collection			
Space group		P2 <sub>1</sub> 2 <sub>1</sub> 2 <sub>1</sub>	
Cell dimensions			
a, b, c (Å)		35.4, 67.6, 88.1	
Wavelength (Å)	0.980062	0.980199	0.978558
Resolution (Å)	50.0-1.45 (1.49-1.45)	50.0-1.45 (1.49-1.45)	50.0-1.45 (1.49-1.45)
R <sub>sym</sub> (%)	4.4 (28.9)	4.8 (30.0)	4.3 (28.7)
<I/σ(I)>	19.23 (2.85)	19.17 (2.71)	20.89 (2.81)
Completeness (%)	95.8 (74.0)	96.5 (75.9)	96.5 (76.2)
Multiplicity	3.01 (1.74)	3.28 (1.87)	3.27 (1.87)
Refinement			
Resolution (Å)	44.0-1.45		
No. reflections	69,339		
R <sub>work</sub> /R <sub>free</sub> (%)	15.2/17.0		
No. atoms			
Protein	1720		
Water	258		
B-factor (Å <sup>2</sup> )			
Protein	17.7		
Water	29.4		
R.m.s. deviations			
Bond lengths (Å)	0.012		
Bond angles (°)	1.339		

binding affinities of Sfc1-DBD for dsDNA and ssDNA (Figure 4A and B). The competition experiments also suggested that dsDNA and ssDNA use overlapping binding sites as dsDNA is able to replace ssDNA and *vice versa*. Binding to dsDNA and ssDNA was also tested for Sfc1-DBD R237A/R238A and R269A/K294A mutant proteins that modify the large and small positively charged patches, respectively. Both mutant proteins are still able to bind ssDNA and dsDNA with similar reduced affinities (Figure 4A and B), suggesting that both patches are involved in ssDNA and dsDNA binding and are required for the full DNA-binding capacity of Sfc1-DBD. Accordingly, the winged-helix domain (residues 186–283) lacking the WHI domain still binds dsDNA and ssDNA, although with ~10-fold reduced affinity compared to entire Sfc1-DBD (Figure 4B and C). Also consistent with the results of endogenous TFIIC binding to the A-site (29), Sfc1-DBD possesses little sequence specificity for A-box containing DNA oligonucleotides, as it binds with similar affinity to B-box dsDNA (data not shown) and to coding and non-coding ssDNA (Figure 4B). To probe the contribution of individual Sfc1-DBD residues more quantitatively, we tested the binding affinities for dsDNA of a wide range of structure-based site-directed mutants (Figure 4C). Several point mutants had an inhibitory effect on dsDNA binding decreasing the affinities by more than 2-fold. These included mutants R237A in helix  $\alpha$ 2, R269A in the 'wing' connecting strands  $\beta$ 1 and  $\beta$ 2, single mutant K294A in the loop connecting strands  $\beta$ 4 and  $\beta$ 5 of the WHI domain and double mutant R269A/K294A (Figure 4C). The observed low-affinity DNA binding with little sequence specificity of Sfc1-DBD well agrees with the DNA-binding characteristics observed for endogenous TFIIC purified from *S. cerevisiae*. Total

deletion of the A-box only decreased the equilibrium constant for TFIIC by a factor 2- to 5-fold (30), whereas ssDNA can interfere with TFIIC binding to the A-box and higher temperature increases the stability of the TFIIC-DNA complex (31) (see 'Discussion' section later in the text).

#### The acidic tails of Sfc1 inhibits its DNA binding activity

We observed a stretch of acidic residues in the C-terminal extension following Sfc1-DBD that is conserved between different Sfc1-orthologs. To further explore the function of these acidic residues, we compared the DNA-binding affinity of Sfc1-DBD and construct Sfc1<sub>111-396</sub> that contains 75 additional N-terminal residues with longer constructs Sfc1<sub>186-456</sub> and Sfc1<sub>111-456</sub> comprising the C-terminal 60-residue extension including the conserved acidic residues. Interestingly, the constructs including this extension did no longer bind dsDNA (Figure 5A), suggesting an auto-inhibitory activity in this region. The Sfc1<sub>1-396</sub>/Sfc7 heterodimer lacking the C-terminal tail also binds dsDNA, whereas the Sfc1/Sfc7 heterodimer that includes the C-terminal tail did not bind DNA, demonstrating that the C-terminal acidic tail of Sfc1 also inhibits DNA binding in the context of the Sfc1/Sfc7 heterodimer. Subsequently, we tested whether the most conserved region of acidic residues in the C-terminal tail might be sufficient for this inhibition. Indeed, a synthetic peptide comprising only 22 C-terminal residues of Sfc1 (Sfc1<sub>435-456</sub>) was able to bind to Sfc1-DBD with an affinity of  $3.0 \pm 0.8 \mu\text{M}$  as shown by isothermal calorimetry (data not shown). The same peptide was also able to disrupt binding of Sfc1-DBD to DNA, when it was titrated in during an EMSA (Figure 5B), suggesting that Sfc1

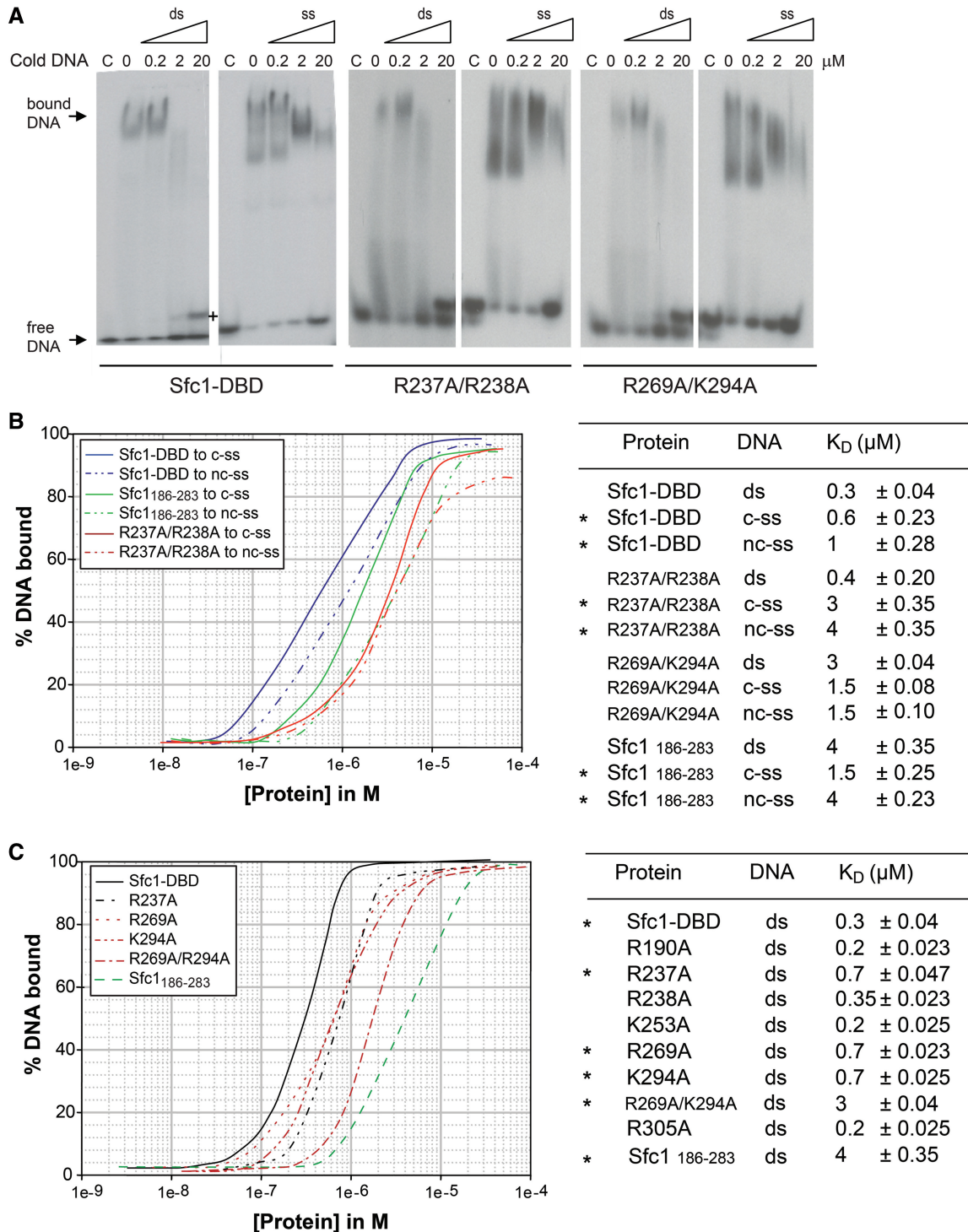


**Figure 3.** Structure of Sfc1-DBD. (A) Ribbon representation of Sfc1-DBD. The winged-helix domain is depicted in bright red, and the WHI domain is depicted in dark red. The right panel shows the winged-helix domain of TFIIIF Rap30 in the same orientation. (B) Electrostatic surface charge representation of Sfc1-DBD. Residues that contribute forming the small and large basic patches are labelled. The right panel shows the electrostatic surface charge of the TFIIIF Rap30 winged-helix domain. (C) Superposition of the winged-helix domains of Sfc1-DBD and Rap30. (D) Sequence alignment of *S. pombe* Sfc1-DBD (*Sp*) with corresponding regions of *H. sapiens* TFIIIC63 (*Hs*), *S. cerevisiae*  $\tau 95$  (*Sc*) and human TFIIIF Rap30 (*Hs*). Secondary structure elements present in Sfc1-DBD are depicted as cylinders ( $\alpha$ -helices) and arrows ( $\beta$ -sheets). Blue dots indicate conserved basic Sfc1-DBD residues mutated for DNA-binding studies.

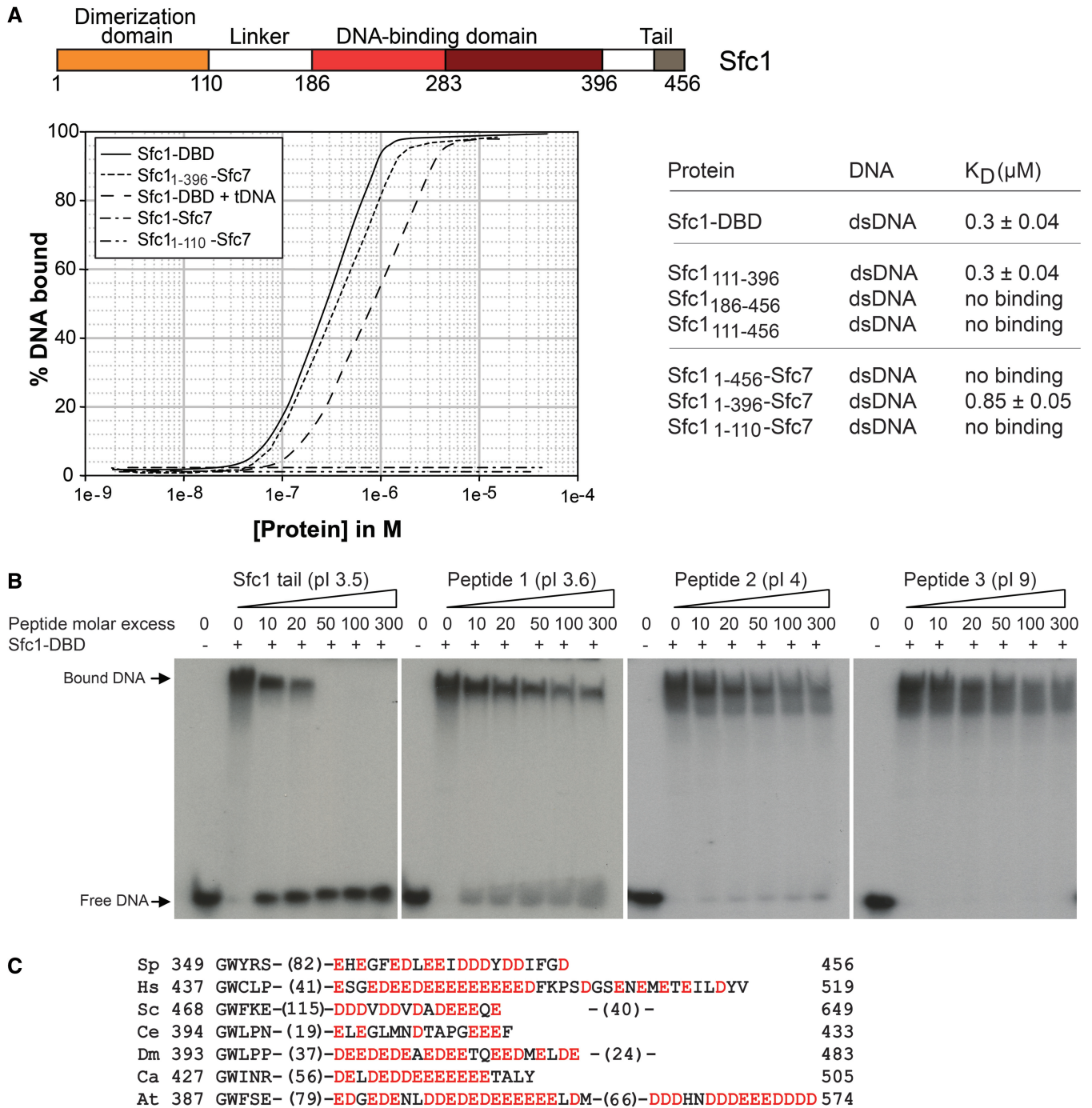
residues 435–456 are sufficient for the autoinhibition of the DNA-binding activity of Sfc1-DBD. The binding of the acidic tail to Sfc1-DBD appears to be specific because other peptides of similar lengths and pI values, but with

unrelated amino acid sequences, are less efficient in disrupting DNA binding (Figure 5B). Cryptic DBDs have been also described in TFIIIF Rap30 (32) and TFIIIB subunit Brf1 (33). Autoinhibitory regions generally





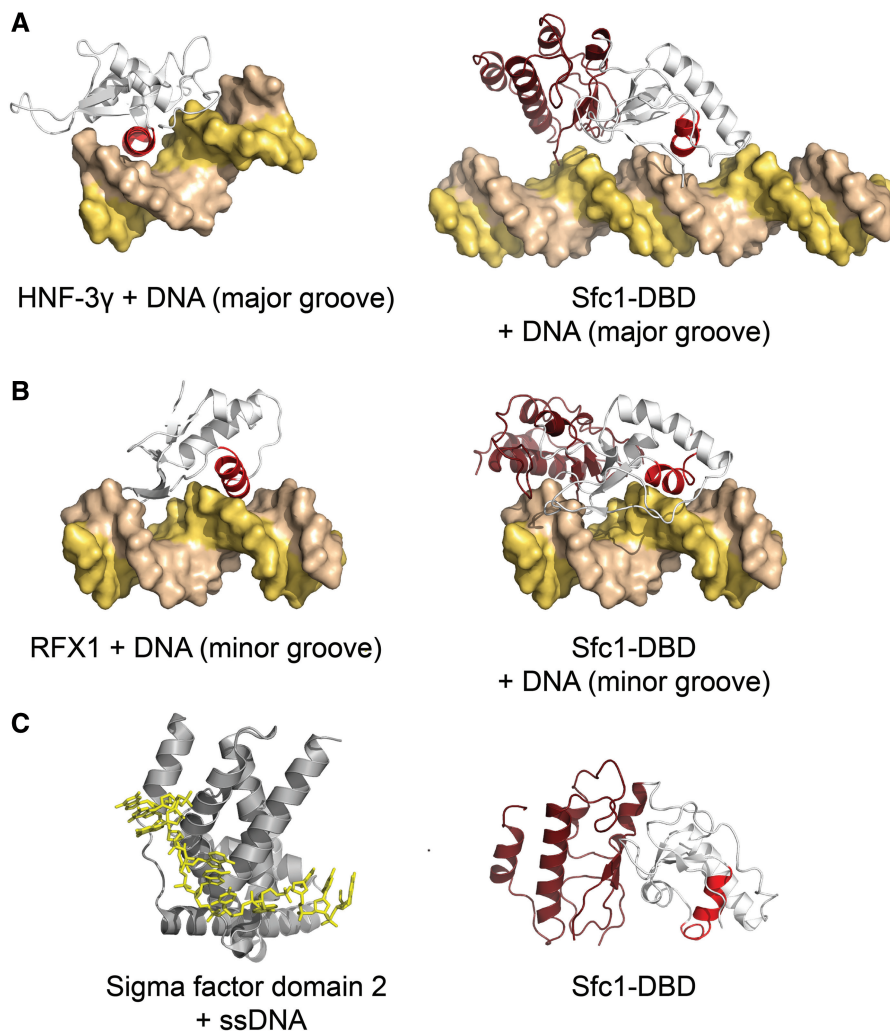
**Figure 4.** DNA binding of Sfc1-DBD to ssDNA and dsDNA. (A) EMSA of Sfc1-DBD and Sfc1-DBD mutant proteins bound to 25meric ssDNA (panel 1) and dsDNA (panel 2). Arrows indicate Sfc1-DBD-DNA complexes. Increasing amounts of non-radioactive (cold) dsDNA or ssDNA were added to compete for binding. The amounts of single-stranded or double-stranded competitor DNA are depicted above the lane in μM. The same experiments using mutant proteins Sfc1-DBD R237A/R238A and R269A/K294A that change the large and small basic patches, respectively, are depicted in panels 3/4 and 5/6. In panel 1, plus indicates dsDNA formed through strand-exchange with radioactive ssDNA. (B) Filter binding assays of Sfc1-DBD, Sfc1<sub>186-283</sub> lacking the WHI domain, Sfc1-DBD R237A/R238A and Sfc1-DBD R269A/K294A mutant proteins bound to dsDNA (ds), non-coding ssDNA (nc-ss) and coding ssDNA (c-ss). The resulting  $K_D$  values are listed in the right table. Asterisks in (B) and (C) indicate those experiments where the corresponding experimental data are depicted in the left panel. (C) Filter binding assays of Sfc1-DBD and Sfc1-DBD mutant proteins bound to a 25 bp, double-stranded A-box DNA site. As in (B), resulting  $K_D$  values are listed in the right table.



**Figure 5.** Auto-inhibition of the DNA binding activity of the Sfc1/Sfc7 heterodimer by its C-terminal tail. (A) Filter binding assays of different Sfc1 and Sfc1/Sfc7 heterodimer constructs to a 25 bp double-stranded A-box DNA site. The domain structure of Sfc1 is given as reference. (B) EMSA of Sfc1-DBD bound to radioactively labelled DNA with the 22mer acidic Sfc1 tail peptide (Sfc1<sub>435-456</sub>) added in increasing amounts to compete for DNA-binding (left panel). Control peptides 1–3 compete less efficiently for DNA binding (three right panels). The molar excess of tail peptide to protein is indicated. For each peptide pI value, length and sequence is given: Sfc1 tail, pI 3.5, 22mer: QEHEGFEDLEEIDDDYDDIFGD. Peptide 1, pI 3.6, 28mer: HLIQDEEYDDEDVPHDLQLSEDEYNSER. Peptide 2, pI 4, 31mer: NSKRLEEDNDEEQSHHKTKQAVSYDEPDED. Peptide 3, pI 9: 23mer: QLPPPKLSMVGFPFLYKYKANP. (C) Alignment of acidic tail residues from Sfc1 *S. pombe* (Sp) with orthologues from human (Hs), *S. cerevisiae* (Sc), *C. elegans* (Ce), *Drosophila melanogaster* (Dm), *Candida albicans* (Ca), *Arabidopsis thaliana* (At). Acidic residues are depicted in red.

increase specificity of DNA recognition, as they only allow DNA binding under well-defined conditions, thereby preventing non-specific binding. Although the acidic C-terminal tail is conserved between different species

(Figure 5C), its functional importance might vary between species. In *S. cerevisiae*, the corresponding C-terminal residues in subunit τ95 are not required for viability at least under normal growth conditions (34).



**Figure 6.** DNA recognition by Sfc1-DBD. (A) dsDNA recognition by HNF-3 $\gamma$  (PDB ID: 1VTN) with the recognition helix coloured in red contacting the major groove of the DNA (left panel). Sfc1-DBD is depicted in the same orientation contacting ideal B-form DNA in the major groove through its recognition helix coloured in red. (B) dsDNA recognition by RFX1 (PDB ID: 1DP7) with the recognition helix contacting the minor groove of the DNA (left panel). Sfc1-DBD is depicted in the same orientation contacting ideal B-form DNA in the minor groove. (C) ssDNA recognition by  $\sigma$  factor domain 2 (PDB ID: 3UGO). Sfc1-DBD is presented in the right panel as reference.

## DISCUSSION

### Role of the Sfc1-Sfc7 heterodimer within TFIIC

Here, we provide first insight into the structural organization of two subunits of the TFIIC  $\tau$ A subcomplex. Our results identified the *S. pombe* Sfc1-Sfc7 heterodimer as stable subcomplex within TFIIC that comprises a Sfc1-Sfc7 triple- $\beta$ -barrel dimerization module and a DBD in the C-terminal moiety of Sfc1. In *S. cerevisiae*, the corresponding  $\tau$ 95- $\tau$ 55 complex has been also described as an independent heterodimer (9). Additional interactions with other TFIIC subunits occur between the Sfc1 linker region that connects the Sfc1 dimerization domain and its C-terminal DBD and  $\tau$ A subunit Sfc4 (data not shown). At present, it is unknown whether the Sfc1-Sfc7 heterodimer also interacts with any of the  $\tau$ B subunits. However, *S. cerevisiae*  $\tau$ 95-E447K mutant (corresponding to Sfc1-DBD residue D328) considerably destabilizes the entire TFIIC-DNA complex (35). Because binding of  $\tau$ A

to the A-box contributes little to the overall DNA-binding affinity of TFIIC (see earlier in the text), this suggested that  $\tau$ 95 might also interact with  $\tau$ B subunits thereby indirectly influencing the binding to these subunits to the B-box (35). Sfc1-DBD interacts with ssDNA and dsDNA with little sequence specificity. Binding of ssDNA and dsDNA involves overlapping binding sites as all site-directed mutants similarly affect ssDNA and dsDNA binding. The binding affinities of Sfc1-DBD and the Sfc1-Sfc7 heterodimer to a 25 bp dsDNA target site and of the entire  $\tau$ A complex to a longer DNA target site (data not shown) are similar. This suggests that the Sfc1-DBD contains most of the important residues required for A-box binding and that Sfc1-DBD is the complete DBD of the  $\tau$ A subcomplex.

### DNA recognition by Sfc1-DBD

Our results show that Sfc1-DBD is able to recognize ssDNA and dsDNA, although the way the Sfc1-DBD



recognizes A-box DNA still requires further analysis. Superimposing Sfc1-DBD onto the crystal structure of a typical winged-helix domain-DNA complex (HNF-3 $\gamma$ -DNA) bound to dsDNA, where the recognition helix H3 protrudes into the major groove, shows that DNA recognition by Sfc1-DBD considerably differs from canonical B-form DNA-recognition by winged-helix domains (36). Although the Sfc1-DBD recognition helix approximately fits into the major groove of B-form DNA (Figure 6A), adjacent conserved residues immediately following strand  $\beta$ 4 of the WHI domain (residues FKLK) would clash with the DNA backbone. To avoid clashes and to maximize interaction between Sfc1-DBD and dsDNA, the DNA needs to bend in such a way that the recognition helix  $\alpha$ 3 could still bind into the major groove, whereas residues FKLK could reach into the minor groove of the DNA. A second alternative dsDNA-binding mode is also possible, where the recognition helix contacts the minor groove as observed in the RFX1-DNA complex (36), whereas residues of the WHI domain would contact the major groove (Figure 6B).

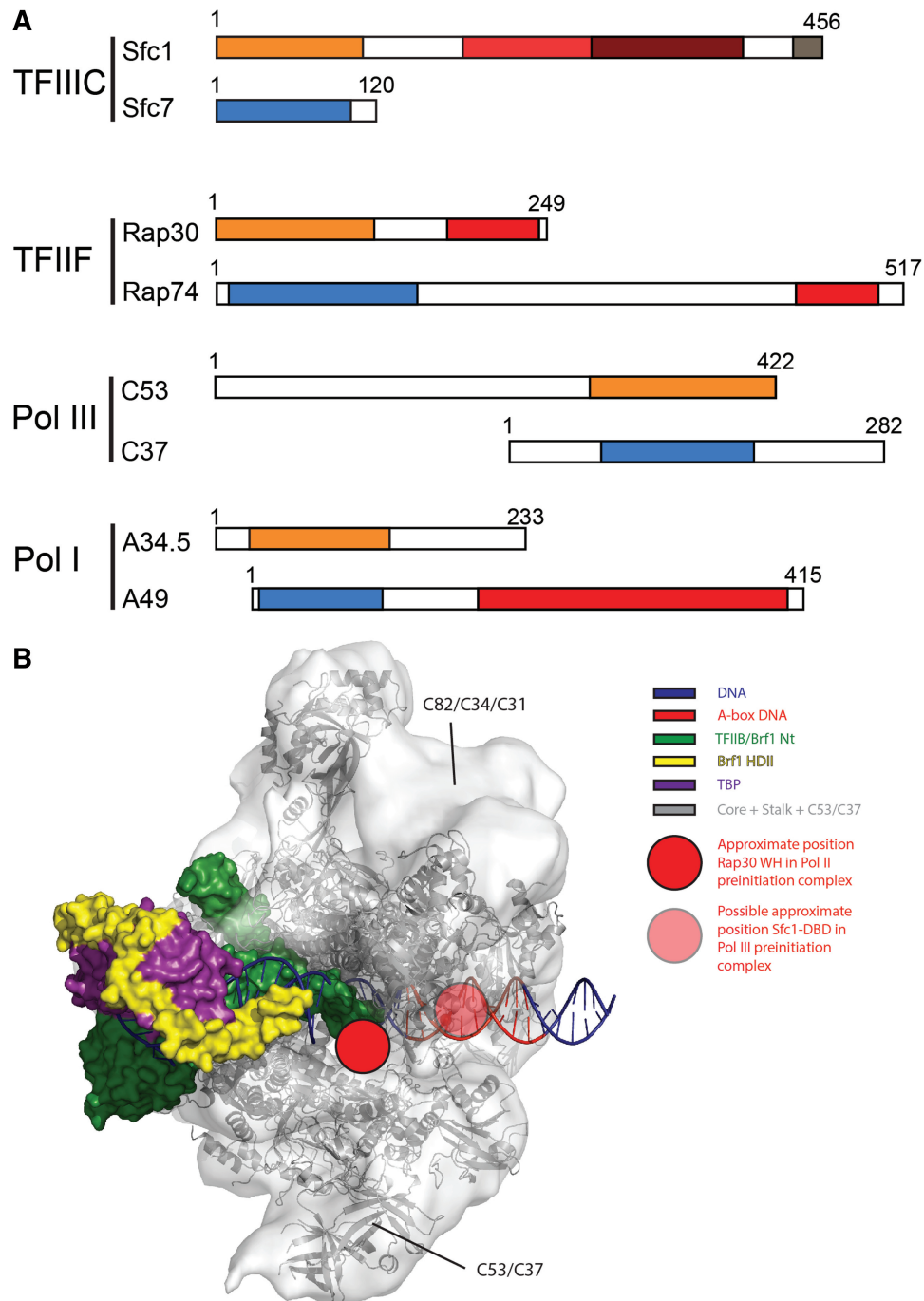
The crystal structure of the bacterial  $\sigma$  factor domain 2 bound to a single-stranded  $-10$  element sequence (37) suggests yet a third alternative DNA-binding mode for the binding of Sfc1-DBD to A-box DNA. In the  $\sigma$  factor domain 2/DNA complex, the dsDNA promoter is opened, the non-template strand is extruded, and DNA bases of the single strand are recognized by intricate hydrophobic and polar interactions resulting in the sequence-specific readout of the  $-10$  element sequence (Figure 6C). The binding characteristics of Sfc1-DBD and the Sfc1-Sfc7 heterodimer described here and of TFIIC purified from *S. cerevisiae* (31) are consistent with a similar DNA-binding mode. Stillman *et al.* observed that  $\tau$ A binding to the A-box contributes little to the TFIIC binding affinity, whereas  $\tau$ B binding to the B-box provides the major contribution. Furthermore,  $\tau$ A binding to the A-box appears to be temperature dependent, not sequence-specific and can be inhibited by ssDNA. These observations had led to the suggestion that the A-box undergoes a closed to open transition during  $\tau$ A binding similar as observed during promoter opening by bacterial DNA-dependent RNA polymerases (31). Interestingly, a similar role in promoter opening has been also proposed for TFIIF Rap30 that also shares some sequence similarity with bacterial  $\sigma$  factor domain 4 (32).

#### Similarities between TFIIC Sfc1/Sfc7 and TFIIF Rap30/Rap74

Unexpectedly, the TFIIC Sfc1/Sfc7 heterodimer resembles the Pol II-specific general transcription factor TFIIF. In mammals and *Drosophila*, TFIIF consists of subunits Rap30 and Rap74, whereas yeast TFIIF comprises subunits Tfg2 and Tfg1 and a non-essential third subunit (38). TFIIF is required for RNA polymerase II PIC assembly (39) and interacts directly with DNA and other PIC components (40–42). TFIIF is also required for open complex formation (43) and transcription start site selection (44,45) and has been shown to function *in vitro* during transcriptional elongation (43). The overall domain

architecture of TFIIF Rap30/Rap74 and TFIIC Sfc1/Sfc7 is similar (Figure 7A). Both heterodimers use N-terminal triple  $\beta$ -barrel domains as dimerization modules (28) and TFIIF Rap30/Rap74 and TFIIC Sfc1 (but not Sfc7) contain winged-helix domains as cryptic DBDs (29,46). The N-terminal region of Sfc1 involved in forming the triple  $\beta$ -barrel domain is most similar to Rap30, whereas the Sfc1 winged-helix domain is most similar to the winged-helix domain of Rap30 (among all winged-helix domains present in the PDB), suggesting a possible common origin of Sfc1 and TFIIF Rap30.

Distant homologies have been also identified between Pol II-specific general transcription factor TFIIF Rap30/Rap74, Pol I subunits A34.5/A49 and Pol III subunits C53/C37, but also between the Pol II-specific general transcription factor TFIIE and Pol III subunits C34 and C82 [reviewed in (47)]. This has led to the suggestion that Pol I and Pol III permanently recruited general transcription factors while they are dissociable in the Pol II enzyme (48). Pol I subunits A34.5/A49 and Pol III subunits C53/C37 are similarly positioned on the RNA polymerase core and are thought to be the functional equivalent to TFIIF Rap30/Rap74 (49–52). Although detailed structural information of the entire C37/C53 heterodimer or parts of it are lacking, comparison of the overall domain organization of TFIIC Sfc1/Sfc7 and Pol III C53/C37 heterodimers suggests that the Sfc1/Sfc7 heterodimer even closer resembles TFIIF Rap30/Rap74 than Pol III C53/C37. Sfc1 and TFIIF Rap30 share a similar overall domain organization where the dimerization domain is followed by a winged-helix domain. In contrast, in C53 the dimerization domain is located at the C-terminus (Figure 7A). Furthermore, we could not find any predicted winged-helix domains for *S. cerevisiae* or *S. pombe* C53 and C37, nor for human C53, using homology modelling (53). In contrast, the human homologue of C37 (which has been described to be more homologous to Rap74 and is considerably larger than its yeast homologues) potentially contains one or two winged-helix domains. Recent cryo-electron microscopy reconstructions located the C-terminal Rap30 winged-helix domain in the human Pol II transcription initiation complex (54). The corresponding position in Pol III is indicated and lies in close proximity to the A-box DNA site recognized by Sfc1-DBD in a model of an initial open Pol III complex (Figure 7B). The C53/C37 subcomplex has been initially implicated in transcription termination (55), but it is also involved in transcription initiation and promoter opening (56). The C53/C37 dimerization module (similar as the corresponding modules in TFIIF Rap30/Rap74 and in A34.5/A49) occupies a position at one side of the Pol III DNA-binding cleft (Figure 7B). Photo-cross-linking experiments demonstrated that the N-terminal extension of C53 reaches across the DNA-binding cleft and contacts Pol III-specific subunits C34 and C82, the Pol III stalk and subunit TFIIC  $\tau$ 131, whereas the C-terminal extension of subunit C37 contacts the active center of Pol III and presumably also the DNA transcription bubble (57). Given their close spatial proximity, Sfc1/Sfc7 could assist C53/C37 during Pol III transcription. Accordingly, the C53/C37 heterodimer would initiate promoter opening,



**Figure 7.** TFIIF-like role of the Sfc1/Sfc7 heterodimer in transcription. **(A)** Overall domain organization of TFIIC Sfc1/Sfc7, TFIIF Rap30/Rap74, Pol III C53/C37 and Pol I A34.5/A49. Dimerization domains related to TFIIF Rap30 and Rap74 are depicted in orange and blue, respectively. The winged-helix domains in TFIIF Rap30, Rap74, Sfc1-DBD and the tandem winged-helix domain in A49 are depicted in bright red, the Sfc1-WHI domain in Sfc1 dark red, and the Sfc1 acidic tail in dark grey. For TFIIC, the *S. pombe* orthologues, for TFIIF, the human orthologues and for Pol I and III, the *S. cerevisiae* orthologues are depicted. **(B)** Model of the Pol III open preinitiation complex. The open PIC complex was prepared by fitting the elongation complex of the Pol II structure (PDB ID: 1I6H) and the homology model of the C53/C37 dimerization domains into the Pol III EM density map (EMDB code 1802). Yeast TFIIB residues 22–211 were obtained from the initially transcribing Pol II complex in complex with TFIIB (PDB ID: 4BBS) after superimposing the Pol II Rpb2 subunits of this structure and the elongation complex. Crystal structures of Pol II-TFIIB (PDB ID: 3K7A), TFIIB-TBP-DNA (PDB ID: 1VOL) and the Brf1 homology region II in complex with TBP and DNA (PDB ID: 1NGM) were used to complete the open PIC. DNA was extended using B-form DNA with the position of the A-box (basepair +8 to +19) indicated in red. Red and transparent red dots indicate the positions of Rap30 winged-helix domain (54) and of Sfc1-DBD, respectively. The colour code is given in the legend at the right.

whereas the Sfc1/Sfc7 heterodimer would bind to the internal A-box promoter DNA, subsequently Sfc1-DBD would extrude and stabilize the non-coding DNA strand similar as observed in the  $\sigma$  factor domain 2/ssDNA complex. Open complex stabilization by the Sfc1/Sfc7 heterodimer would allow Pol III to transcribe across the internal A-box promoter site of tRNA genes, while maintaining contacts with TFIIC, as it has been postulated as requirement for 'facilitated re-initiation' during Pol III transcription. Indeed, the high rate of transcription at Pol III genes, together with the short length of these genes, suggests the possibility that a large part of Pol III tRNA genes could be single-stranded during active transcription. It might be energetically more favourable to keep the gene-internal promoter open rather than to repeatedly open and close the DNA. The fact that the ssDNA-binding factor Sub1 plays an important role in Pol III transcription further supports this hypothesis (58,59).

### ACCESSION NUMBERS

Coordinates and structure factors have been deposited in the Protein Data Bank with accession numbers 4BJJ and 4BJI for Sfc1<sub>1-110(K48A,Q49A)</sub>/Sfc7<sub>1-101</sub> and Sfc1-DBD, respectively.

### SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

### ACKNOWLEDGEMENTS

The authors thank I. Kuciński and F. Rodriguez Calviño for assistance during the project. They acknowledge G. Male for providing *S. cerevisiae*  $\tau$ A and J. Kosinski for modelling Sfc1-DBD DNA complexes. They are also grateful to J. Acker, O. Lefebvre, C. Fernández-Tornero, G. Male and S. Glatt for comments on the manuscript. The authors acknowledge support by the EMBL Protein Expression and Purification and the Proteomics Core Facilities, the EMBL Heidelberg Crystallization Platform and the EMBL/ESRF Joint Structural Biology Group for access and support at ESRF beamlines.

### FUNDING

Funding for open access charge: European Molecular Biology Laboratory (EMBL), Heidelberg.

### REFERENCES

- Schramm, L. and Hernandez, N. (2002) Recruitment of RNA polymerase III to its target promoters. *Genes Dev.*, **16**, 2593–2620.
- Marck, C., Kachouri-Lafond, R., Lafontaine, I., Westhof, E., Dujon, B. and Grosjean, H. (2006) The RNA polymerase III-dependent family of genes in hemiascomycetes: comparative RNomics, decoding strategies, transcription and evolutionary implications. *Nucleic Acids Res.*, **34**, 1816–1835.
- Stillman, D.J. and Geiduschek, E.P. (1984) Differential binding of a *S. cerevisiae* RNA polymerase III transcription factor to two promoter segments of a tRNA gene. *EMBO J.*, **3**, 847–853.
- Iwasaki, O. and Noma, K.I. (2012) Global genome organization mediated by RNA polymerase III-transcribed genes in fission yeast. *Gene*, **493**, 195–200.
- Schultz, P., Marzouki, N., Marck, C., Ruet, A., Oudet, P. and Sentenac, A. (1989) The two DNA-binding domains of yeast transcription factor tau as observed by scanning transmission electron microscopy. *EMBO J.*, **8**, 3815–3824.
- Huang, Y. and Maraia, R.J. (2001) Comparison of the RNA polymerase III transcription machinery in *Schizosaccharomyces pombe*, *Saccharomyces cerevisiae* and human. *Nucleic Acids Res.*, **29**, 2675–2690.
- Ducrot, C., Lefebvre, O., Landrieux, E., Guirouilh-Barbat, J., Sentenac, A. and Acker, J. (2006) Reconstitution of the yeast RNA polymerase III transcription system with all recombinant factors. *J. Biol. Chem.*, **281**, 11685–11692.
- Bartholomew, B., Kassavetis, G.A., Braun, B.R. and Geiduschek, E.P. (1990) The subunit structure of *Saccharomyces cerevisiae* transcription factor IIIC probed with a novel photocrosslinking reagent. *EMBO J.*, **9**, 2197–2205.
- Manaud, N., Arrebola, R., Buffin-Meyer, B., Lefebvre, O., Voss, H., Riva, M., Conesa, C. and Sentenac, A. (1998) A chimeric subunit of yeast transcription factor IIIC forms a subcomplex with tau95. *Mol. Cell. Biol.*, **18**, 3191–3200.
- Dumay-Odelot, H., Marck, C., Durrieu-Gaillard, S., Lefebvre, O., Jourdain, S., Prochazkova, M., Pflieger, A. and Teichmann, M. (2007) Identification, molecular cloning, and characterization of the sixth subunit of human transcription factor TFIIC. *J. Biol. Chem.*, **282**, 17179–17189.
- Taylor, N.M., Glatt, S., Hennrich, M.L., von Scheven, G., Grottsch, H., Fernandez-Tornero, C., Rybin, V., Gavin, A.C., Kolb, P. and Muller, C.W. (2013) Structural and functional characterization of a phosphatase domain within yeast general transcription factor IIIC. *J. Biol. Chem.*, **288**, 15110–15120.
- Mylona, A., Fernández-Tornero, C., Legrand, P., Haupt, M., Sentenac, A., Acker, J. and Müller, C.W. (2006) Structure of the tau60/Delta tau91 subcomplex of yeast transcription factor IIIC: insights into preinitiation complex assembly. *Mol. Cell*, **24**, 221–232.
- Hubbard, S.J. (1998) The structural aspects of limited proteolysis of native proteins. *Biochim. Biophys. Acta*, **1382**, 191–206.
- Kabsch, W. (2010) XDS. *Acta Crystallogr. D*, **66**, 125–132.
- Evans, P. (2006) Scaling and assessment of data quality. *Acta Crystallogr. D*, **62**, 72–82.
- Sheldrick, G.M. (2008) A short history of SHELX. *Acta Crystallogr. A*, **64**, 112–122.
- Pape, T. and Schneider, T. (2004) HKL2MAP: a graphical user interface for macromolecular phasing with SHELX programs. *J. Appl. Crystallogr.*, **37**, 843–844.
- Vonrhein, C., Blanc, E., Roversi, P. and Bricogne, G. (2007) Automated structure solution with autoSHARP. *Methods Mol. Biol.*, **364**, 215–230.
- Abrahams, J.P. and Leslie, A.G.W. (1996) Methods used in the structure determination of bovine mitochondrial F1 ATPase. *Acta Crystallogr. D*, **52**, 30–42.
- Perrakis, A., Harkiolaki, M., Wilson, K.S. and Lamzin, V.S. (2001) ARP/wARP and molecular replacement. *Acta Crystallogr. D*, **57**, 1445–1450.
- Emsley, P., Lohkamp, B., Scott, W.G. and Cowtan, K. (2010) Features and development of Coot. *Acta Crystallogr. D*, **66**, 486–501.
- Adams, P.D., Afonine, P.V., Bunkóczi, G., Chen, V.B., Davis, I.W., Echols, N., Headd, J.J., Hung, L.W., Kapral, G.J., Grosse-Kunstleve, R.W. et al. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D*, **66**, 213–221.
- Chen, V.B., Arendall, W.B., Headd, J.J., Keedy, D.A., Immormino, R.M., Kapral, G.J., Murray, L.W., Richardson, J.S. and Richardson, D.C. (2010) MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D*, **66**, 12–21.



24. Doublet, S. (2007) Production of selenomethionyl proteins in prokaryotic and eukaryotic expression systems. *Methods Mol. Biol.*, **363**, 91–108.
25. Baudin, F., Bach, C., Cusack, S. and Ruigrok, R.W. (1994) Structure of influenza virus RNPI. Influenza virus nucleoprotein melts secondary structure in panhandle RNA and exposes the bases to the solvent. *EMBO J.*, **13**, 3158–3165.
26. Goldschmidt, L., Cooper, D.R., Derewenda, Z.S. and Eisenberg, D. (2007) Toward rational protein crystallization: A Web server for the design of crystallizable protein variants. *Protein Sci.*, **16**, 1569–1576.
27. Holm, L. and Rosenström, P. (2010) Dali server: conservation mapping in 3D. *Nucleic Acids Res.*, **38**, W545–549.
28. Gaiser, F., Tan, S. and Richmond, T.J. (2000) Novel dimerization fold of RAP30/RAP74 in human TFIIF at 1.7 Å resolution. *J. Mol. Biol.*, **302**, 1119–1127.
29. Groft, C.M., Uljon, S.N., Wang, R. and Werner, M.H. (1998) Structural homology between the Rap30 DNA-binding domain and linker histone H5: implications for preinitiation complex assembly. *Proc. Natl Acad. Sci. USA*, **95**, 9117–9122.
30. Baker, R.E., Gabrielsen, O. and Hall, B.D. (1986) Effects of tRNA<sup>Tyr</sup> point mutations on the binding of yeast RNA polymerase III transcription factor C. *J. Biol. Chem.*, **261**, 5275–5282.
31. Stillman, D.J., Caspers, P. and Geiduschek, E.P. (1985) Effects of temperature and single-stranded DNA on the interaction of an RNA polymerase III transcription factor with a tRNA gene. *Cell*, **40**, 311–317.
32. Tan, S., Garrett, K.P., Conaway, R.C. and Conaway, J.W. (1994) Cryptic DNA-binding domain in the C terminus of RNA polymerase II general transcription factor RAP30. *Proc. Natl Acad. Sci. USA*, **91**, 9808–9812.
33. Huet, J., Conesa, C., Carles, C. and Sentenac, A. (1997) A cryptic DNA binding domain at the COOH terminus of TFIIB70 affects formation, stability, and function of preinitiation complexes. *J. Biol. Chem.*, **272**, 18341–18349.
34. Aye, M., Dildine, S.L., Claypool, J.A., Jourdain, S. and Sandmeyer, S.B. (2001) A truncation mutant of the 95-kilodalton subunit of transcription factor IIIC reveals asymmetry in Ty3 integration. *Mol. Cell. Biol.*, **21**, 7839–7851.
35. Jourdain, S., Acker, J., Ducrot, C., Sentenac, A. and Lefebvre, O. (2003) The tau95 subunit of yeast TFIIC influences upstream and downstream functions of TFIIC-DNA complexes. *J. Biol. Chem.*, **278**, 10450–10457.
36. Gajiwala, K.S. and Burley, S.K. (2000) Winged helix proteins. *Curr. Opin. Struct. Biol.*, **10**, 110–116.
37. Feklistov, A. and Darst, S.A. (2011) Structural basis for promoter-10 element recognition by the bacterial RNA polymerase sigma subunit. *Cell*, **147**, 1257–1269.
38. Henry, N.L., Campbell, A.M., Feaver, W.J., Poon, D., Weil, P.A. and Kornberg, R.D. (1994) TFIIF-TAF-RNA polymerase II connection. *Genes Dev.*, **8**, 2868–2878.
39. Tan, S., Conaway, R.C. and Conaway, J.W. (1995) Dissection of transcription factor TFIIF functional domains required for initiation and elongation. *Proc. Natl Acad. Sci. USA*, **92**, 6042–6046.
40. Cabart, P., Ujvari, A., Pal, M. and Luse, D.S. (2012) Transcription factor TFIIF is not required for initiation by RNA polymerase II, but it is essential to stabilize transcription factor TFIIB in early elongation complexes. *Proc. Natl Acad. Sci. USA*, **108**, 15786–15791.
41. Fishburn, J. and Hahn, S. (2012) Architecture of the yeast RNA polymerase II open complex and regulation of activity by TFIIF. *Mol. Cell. Biol.*, **32**, 12–25.
42. Wang, B.Q. and Burton, Z.F. (1995) Functional domains of human RAP74 including a masked polymerase binding domain. *J. Biol. Chem.*, **270**, 27035–27044.
43. Yan, Q., Moreland, R.J., Conaway, J.W. and Conaway, R.C. (1999) Dual roles for transcription factor IIF in promoter escape by RNA polymerase II. *J. Biol. Chem.*, **274**, 35668–35675.
44. Chen, H.T., Warfield, L. and Hahn, S. (2007) The positions of TFIIF and TFIIE in the RNA polymerase II transcription preinitiation complex. *Nat. Struct. Mol. Biol.*, **14**, 696–703.
45. Freire-Picos, M.A., Krishnamurthy, S., Sun, Z.W. and Hampsey, M. (2005) Evidence that the Tfg1/Tfg2 dimer interface of TFIIF lies near the active center of the RNA polymerase II initiation complex. *Nucleic Acids Res.*, **33**, 5045–5052.
46. Kamada, K., De Angelis, J., Roeder, R.G. and Burley, S.K. (2001) Crystal structure of the C-terminal domain of the RAP74 subunit of human transcription factor IIF. *Proc. Natl Acad. Sci. USA*, **98**, 3115–3120.
47. Vannini, A. and Cramer, P. (2012) Conservation between the RNA polymerase I, II, and III transcription initiation machineries. *Mol. Cell*, **45**, 439–446.
48. Carter, R. and Drouin, G. (2010) The increase in the number of subunits in eukaryotic RNA polymerase III relative to RNA polymerase II is due to the permanent recruitment of general transcription factors. *Mol. Biol. Evol.*, **27**, 1035–1043.
49. Chen, Z.A., Jawhari, A., Fischer, L., Buchen, C., Tahir, S., Kamenski, T., Rasmussen, M., Lariviere, L., Bukowski-Wills, J.C., Nilges, M. *et al.* (2010) Architecture of the RNA polymerase II-TFIIF complex revealed by cross-linking and mass spectrometry. *EMBO J.*, **29**, 717–726.
50. Eichner, J., Chen, H.T., Warfield, L. and Hahn, S. (2010) Position of the general transcription factor TFIIF within the RNA polymerase II transcription preinitiation complex. *EMBO J.*, **29**, 706–716.
51. Fernández-Tornero, C., Böttcher, B., Rashid, U.J., Steuerwald, U., Flörchinger, B., Devos, D.P., Lindner, D. and Müller, C.W. (2010) Conformational flexibility of RNA polymerase III during transcriptional elongation. *EMBO J.*, **29**, 3762–3772.
52. Kuhn, C., Geiger, S.R., Baumli, S., Gartmann, M., Gerber, J., Jennebach, S., Mielke, T., Tschochner, H., Beckmann, R. and Cramer, P. (2007) Functional architecture of RNA polymerase I. *Cell*, **131**, 1260–1272.
53. Söding, J., Biegert, A. and Lupas, A.N. (2005) The HHpred interactive server for protein homology detection and structure prediction. *Nucleic Acids Res.*, **33**, W244–W248.
54. He, Y., Fang, J., Taatjes, D.J. and Nogales, E. (2013) Structural visualization of key steps in human transcription initiation. *Nature*, **495**, 481–486.
55. Landrieux, E., Alic, N., Ducrot, C., Acker, J., Riva, M. and Carles, C. (2006) A subcomplex of RNA polymerase III subunits involved in transcription termination and reinitiation. *EMBO J.*, **25**, 118–128.
56. Kassavetis, G.A., Prakash, P. and Shim, E. (2010) The C53/C37 subcomplex of RNA polymerase III lies near the active site and participates in promoter opening. *J. Biol. Chem.*, **285**, 2695–2706.
57. Wu, C.C., Lin, Y.C. and Chen, H.T. (2011) The TFIIF-like Rpc37/53 dimer lies at the center of a protein network to connect TFIIC, Bdp1, and the RNA polymerase III active center. *Mol. Cell. Biol.*, **31**, 2715–2728.
58. Rosonina, E., Willis, I.M. and Manley, J.L. (2009) Sub1 functions in osmoregulation and in transcription by both RNA polymerases II and III. *Mol. Cell. Biol.*, **29**, 2308–2321.
59. Tavenet, A., Suleau, A., Dubreuil, G., Ferrari, R., Ducrot, C., Michaut, M., Aude, J.C., Dieci, G., Lefebvre, O., Conesa, C. *et al.* (2009) Genome-wide location analysis reveals a role for Sub1 in RNA polymerase III transcription. *Proc. Natl Acad. Sci. USA*, **106**, 14265–14270.