





Structural studies of RFC^{Ctf18} reveal a novel chromatin recruitment role for Dcc1

Benjamin O Wade¹, Hon Wing Liu², Catarina P Samora², Frank Uhlmann² & Martin R Singleton^{1,*}



Replication factor C complexes load and unload processivity clamps from DNA and are involved in multiple DNA replication and repair pathways. The RFC^{Ctf18} variant complex is required for activation of the intra-S-phase checkpoint at stalled replication forks and aids the establishment of sister chromatid cohesion. Unlike other RFC complexes, RFC^{Ctf18} contains two non-Rfc subunits, Dcc1 and Ctf8. Here, we present the crystal structure of the Dcc1-Ctf8 heterodimer bound to the C-terminus of Ctf18. We find that the C-terminus of Dcc1 contains three-winged helix domains, which bind to both ssDNA and dsDNA. We further show that these domains are required for full recruitment of the complex to chromatin, and correct activation of the replication checkpoint. These findings provide the first structural data on a eukaryotic seven-subunit clamp loader and define a new biochemical activity for Dcc1.

Keywords Ctf18; Dcc1; sister chromatid cohesion; S-phase checkpoint; X-ray crystallography

Subject Categories Cell Cycle; DNA Replication, Repair & Recombination; Structural Biology

DOI 10.15252/embr.201642825 | Received 2 June 2016 | Revised 5 January 2017 | Accepted 9 January 2017 | Published online 10 February 2017

EMBO Reports (2017) 18: 558-568

Introduction

Accurate duplication and segregation of chromosomes is vital to maintain integrity of the genetic material. To facilitate this, cells have evolved checkpoint mechanisms to ensure that aberrant processes are detected and where possible, repaired. One such set of checkpoints occur during DNA replication. The so-called intra-Sphase checkpoints are required to maintain genome integrity at stalled replication forks [1–5]. These are characterised by a build-up of single-stranded DNA (ssDNA) that can arise from uncoupled helicase and polymerase activity, resulting from DNA lesions in the template. This ssDNA is recognised by the sensor kinase Mec1 (ATR in humans) which interacts with the ssDNA-binding protein replication protein A (RPA) [6]. Activated Mec1 in turn recruits and activates the diffusible checkpoint kinase Rad53 (Chk2) in

conjunction with Mrc1 (Claspin) [7–10]. The activation of Rad53 is associated with its hyper-phosphorylation and induces a range of downstream effects including the stabilisation of stalled forks, repression of late-firing origins [11,12], up-regulation of dNTP synthesis and transcription of DNA repair genes [13–15].

A number of studies in yeast have shown the requirement for another protein complex, RFC^{Ctf18} as being required for full activation of Rad53 [16-18]. RFCCtf18 is a variant replication factor C-like complex, the basic function of which is to load or unload a DNA sliding clamp in an ATP-dependent manner [19,20]. All RFC complexes contain four common subunits (Rfc2-5) and a single variable large subunit that defines the function of the complex [21]. The prototypical member is RFCRfc1, which loads PCNA onto DNA at replication forks and is essential for all replication and repair processes. The best understood variant is probably RFC^{Rad24}, which participates in the DNA damage checkpoint response by loading a specialised clamp formed by Rad17, Mec3 and Ddc1 [22-26]. The functions of other variants, RFC^{Elg1} and RFC^{Ctf18} are less clear. RFC^{Elg1} has been implicated in PCNA unloading [27-29] and maintenance of genomic stability [30-36], while RFCCtf18 was originally identified in screens for chromosome mis-segregation [18,37,38]. In addition to containing the Ctf18 variant of the Rfc1 subunit, ${\rm RFC}^{\rm Ctf18}$ also contains the proteins Dcc1 and Ctf8, making it the only RFC variant formed of seven subunits. RFCCtf18 can both unload and load PCNA in vitro [39-41]; however, studies in yeast have suggested that its in vivo role is as a loader, as cells lacking Ctf18 show a reduced level of PCNA at sites of DNA replication [42]. The finding that Ctf18 deletion also impinges upon sister chromatid cohesion (SCC) has led to it being classified as a non-essential SCC establishment factor, together with other replication-associated proteins such as Chl1, Ctf4, Csm3, Mrc1 and Tof1 [38,43-47].

Despite the characterisation of the phenotypic effects of RFC^{Ctf18} deletion and its implication in multiple cellular pathways, the biochemical and mechanistic properties of the complex are not fully understood. As the Dcc1-Ctf8 sub-complex is not required for the PCNA-loading or unloading reaction [39,40], we hypothesised that it might play a regulatory role. To investigate this, we have solved the crystal structure of the Dcc1-Ctf8 heterodimer bound to the C-terminus of Ctf18 (Ctf18^c). To our surprise, we found that the C-terminus of Dcc1 contains three-winged helix (WH) domains, which we demonstrate can bind to both ssDNA and double-stranded DNA (dsDNA). We further show that these domains aid recruitment

¹ Structural Biology of Chromosome Segregation Laboratory, The Francis Crick Institute, London, UK

² Chromosome Segregation Laboratory, The Francis Crick Institute, London, UK *Corresponding author. Tel: +44 203 796 2034; E-mail: martin.singleton@crick.ac.uk

559

of the complex to chromatin *in vivo* and are required for checkpoint activation

Results

Construct design and structure solution

RFC^{Ctf18} is formed of seven subunits: Ctf18, Rfc2, Rfc3, Rfc4, Rfc5, Ctf8 and Dcc1 (Fig 1A). The core complex, which is sufficient to load and unload PCNA from DNA [37], is formed of Ctf18 and the four small Rfc subunits. Each of these proteins contains an AAA+ATPase domain. Previous studies have shown that Dcc1 and Ctf8

are able to form a stable sub-complex *in vitro* and that the C-terminus of Ctf18 is required for the interaction with the Dcc1-Ctf8 sub-complex [39]. To further investigate the interaction between Ctf18 and the Dcc1-Ctf8 sub-complex and the function of Dcc1-Ctf8, we designed a range of constructs using the *S. cerevisiae* proteins for structural studies. We identified a construct that expressed well in *E. coli* that includes the last 76 residues of Ctf18 (residues 666–741) with full-length Dcc1 and Ctf8. Crystals of this complex (designated Ctf18^C-Dcc1-Ctf8) were grown that diffracted to 2.4 Å resolution. While working on this complex, we also obtained crystals of the C-terminal domain of Dcc1 (residues 90–380) that diffracted to 2 Å. Phase information for this crystal form was determined using multiple-wavelength anomalous diffraction

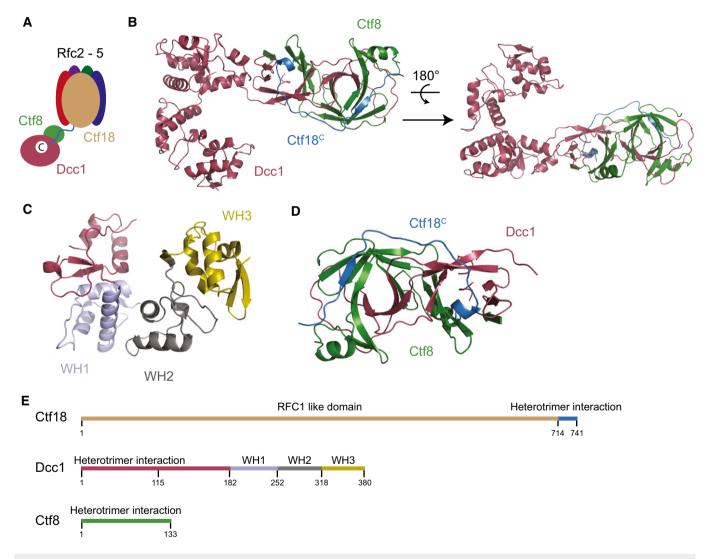


Figure 1. The overall structure of the heterotrimer.

- A Schematic showing the overall organisation of RFCCtf18.
- B The structure of the heterotrimeric complex Dcc1 (maroon), Ctf8 (green) and Ctf18^C (blue).
- C Detailed view of the C-terminal domain of Dcc1 with WH1 (blue), WH2 (grey) and WH3 (yellow) identified.
- D The interaction domain between Ctf18 (blue) and the sub-complex, Dcc1 (red) and Ctf8 (green).
- E Linear representation of the overall domain organisation of the heterotrimer.

Data information: See also Figs EV1 and EV2.

© 2017 The Francis Crick Institute EMBO reports Vol 18 | No 4 | 2017

(MAD) of selenomethionine-substituted protein. The refined structure was then used as a search model to determine the structure of the Ctf18^C-Dcc1-Ctf8 heterotrimer by molecular replacement (MR). The final structures were of high quality with full statistics of the crystallographic analyses given in Appendix Table S1.

Overall structure of the Ctf18^C-Dcc1-Ctf8 heterotrimer

The overall structure of Ctf18 C -Dcc1-Ctf8 is extended and looks like a "hook", as the C-terminus of Dcc1 folds back towards Ctf8 (Fig 1B). This C-terminus is primarily α -helical, while the N-terminus

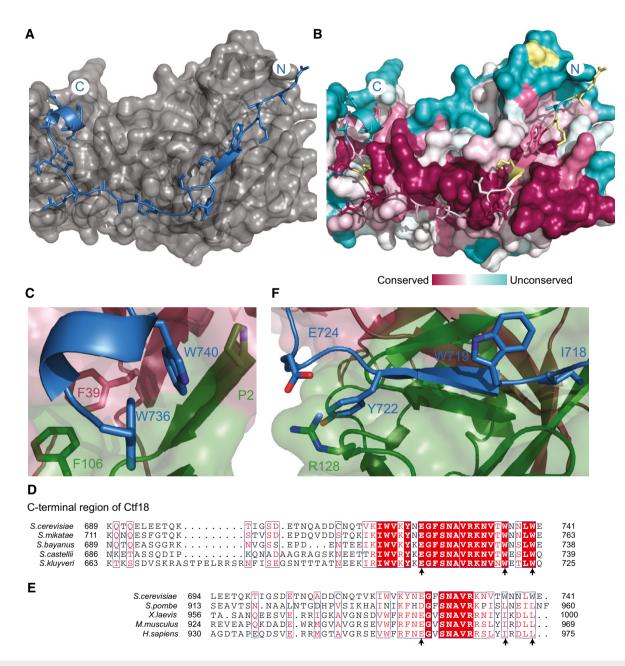


Figure 2. Ctf18 binding to the Dcc1-Ctf8 sub-complex.

- A The binding interface of Ctf18 (blue) residues 714–740 to the surface of Dcc1-Ctf8 (grey).
- B Surface conservation of the Ctf18-binding interface, shown in the same orientation as in (A). Yellow residues indicate that there was insufficient data from the alignment.
- C The C-terminal helix of Ctf18 (blue) with Ctf8 (green) and Dcc1 (maroon). The conserved tryptophans are shown in stick representation.
- D Sequence alignment of yeast species with arrows indicating the location of conserved interacting residues.
- E An expanded sequence alignment from yeast to human showing the functionally equivalent regions. Arrows indicate conserved interacting residues.
- F The interaction of the β -strand region of Ctf18 (blue) with Ctf8 (green) and Dcc1 (maroon).

560 EMBO reports Vol 18 | No 4 | 2017 (s) 2017 The Francis Crick Institute

of the protein and the majority of Ctf8 consist of β -strands. To identify whether any common domains were present in the heterotrimer, we performed a three-dimensional homology search using the Dali server [48]. The results indicated that Dcc1 contains three-winged helix (WH) domains from residues 182–252, 253–317 and 318–380, which we have called WH1, WH2 and WH3, respectively (Fig 1C). These WH domains show structural homology to the WH domains from PKZ [49], cullin-1 [50] and DsrD [51], with Z-scores of 5.7, 7.4 and 4.9, respectively (Fig EV1). WH domains are often

involved in DNA binding or mediating protein–protein interactions [52], suggesting possible functions for this structural feature. The N-terminus of Dcc1 from residues 1 to 115 forms a large heterodimer interface with Ctf8 (Fig 1D). This domain is comprised of 2 α -helices and 18 interconnecting β -strands from both Dcc1 and Ctf8. The overall structure of the heterodimer resembles the "triple" β -barrel fold which has previously been observed in RNase H2 [53], and the transcription factors Rap30/74 [54], Sfc1/7 [55] and A49/34.5 [56] (Fig EV2). This architecture appears to be used

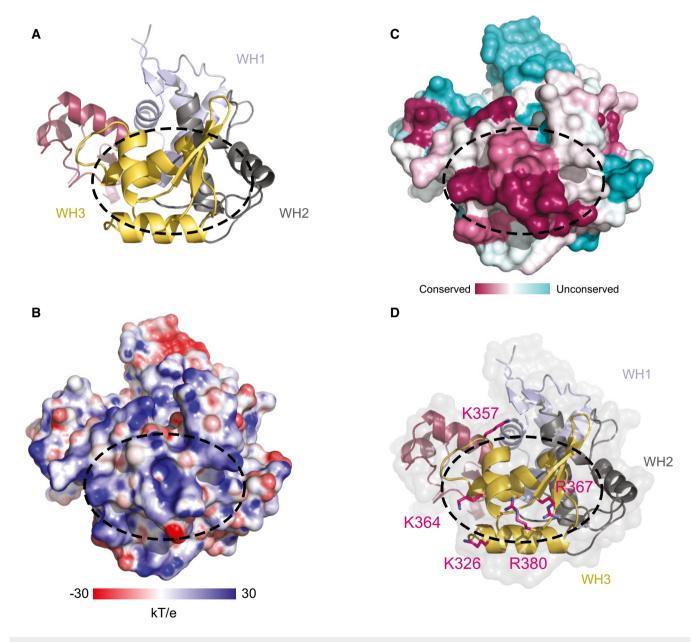


Figure 3. DNA-binding domain of Dcc1.

- A Ribbon diagram of the C-terminal domain of Dcc1 with the triple WH domains: WH1 (blue), WH2 (grey) and WH3 (yellow).
- B The surface charge of the C-terminal domain of Dcc1 with positive areas in blue and negative in red.
- C The surface conservation of Dcc1 with highly conserved regions in dark pink and the unconserved regions in pale blue.
- D Potential DNA-binding basic residues tested by mutational analysis depicted in stick representation.

Data information: The dotted areas indicate the presumptive DNA-binding region. Orientation of the structures is the same as in (A).

© 2017 The Francis Crick Institute EMBO reports Vol 18 | No 4 | 2017 561

primarily as a protein interaction motif, linking separate functional domains of larger complexes. The similarity of Dcc1-Ctf8 to the transcription factors listed also extends to sharing a modular structure consisting of a hetero-dimerisation fold preceding tandem WH domains (Fig 1E).

Ctf18 binds to the Dcc1-Ctf8 heterodimer

Ctf18 binds as an extended peptide to a conserved groove present in the dimer interface of Dcc1-Ctf8 (Fig 2A and B). The extreme Cterminus of Ctf18 forms a short helix, which is partially buried in a hydrophobic pocket of the first "barrel" of Dcc1-Ctf8. The structure shows that W736 and W740 make crucial contacts with other aromatic residues of both Dcc1 and Ctf8 as well as the extreme Nterminal proline of Ctf8 (Fig 2C). These tryptophans are totally conserved in yeast and are always hydrophobic residues in higher eukaryotes, suggesting the overall binding mechanism is conserved (Fig 2D and E). A recent study supports this binding model as coimmunoprecipitation experiments in cells show that mutating both W736 and W740 in Ctf18 disrupt binding to Dcc1-Ctf8 [57]. Furthermore, mutating both residues causes a defect in the phosphorylation of Rad53 in hydroxyurea-treated cells, indicating impaired checkpoint response. The N-terminal section of Ctf18 in our structure forms a beta sheet with Ctf8, with E724 from Ctf18 forming a salt bridge with R128 from Ctf8 (Fig 2F). Residues 666-713 of Ctf18 appear to be unstructured, suggesting they form a flexible linker between the Dcc1-Ctf8 module and the rest of the Rfc1-like domain of Ctf18. The overall binding mechanism is similar to that by which RNase H2A binds to RNase H2B and H2C [53] (Fig EV2) where an extended peptide from H2A binds along the H2B-H2C dimer interface.

Dcc1 is able to bind DNA through the conserved WH domains

To further investigate the function of Dcc1, we studied both the surface charge and conservation of the WH domains (Fig 3A-D). A large conserved basic patch was present on the surface of WH3. This suggested the domain could interact with DNA as basic residues are able to bind to the negatively charged backbone of DNA. To test this, we carried out an electrophoretic mobility shift assay (EMSA) with fluorescently labelled ssDNA and dsDNA oligonucleotides. The Dcc1 C-terminal construct was titrated against a constant DNA concentration. A gel shift was observed with both ssDNA and dsDNA, suggesting Dcc1 could interact with both substrates (Fig 4A). To further dissect the contribution of the individual WH domains to DNA binding, we deleted WH3 (residue 319-380) from Dcc1 and repeated the EMSA. The results of the initial experiments indicate that WH1 and WH2 could not bind either ssDNA or dsDNA (Fig 4B). However, when we reduced the DNA concentration and overexposed the gel, a slight

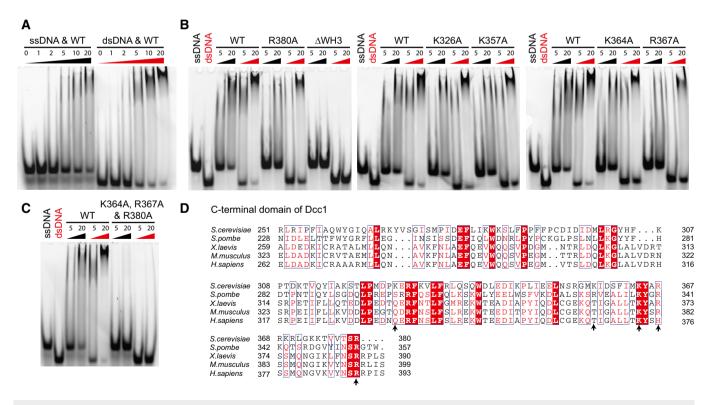
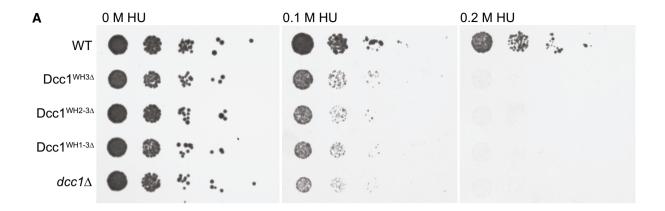


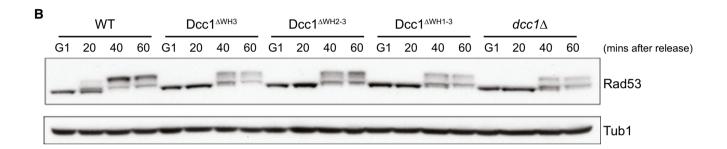
Figure 4. DNA binding analyses of Dcc1.

562

- A DNA binding analysis conducted via an EMSA. 22-mer dsDNA or 25-mer ssDNA was fluorescently labelled, and the C-terminus of Dcc1 (residues 90–380) was titrated in at concentrations indicated. Protein concentrations are given in μM.
- B DNA binding properties of Dcc1 point mutants and WH3 deletion. Protein concentrations are given in μM.
- C DNA binding by triple mutant. Protein concentrations are given in μM .
- D Sequence alignment from yeast to human of WH2 and WH3, with arrows indicating putative basic DNA-binding residues.

EMBO reports Vol 18 | No 4 | 2017 © 2017 The Francis Crick Institute





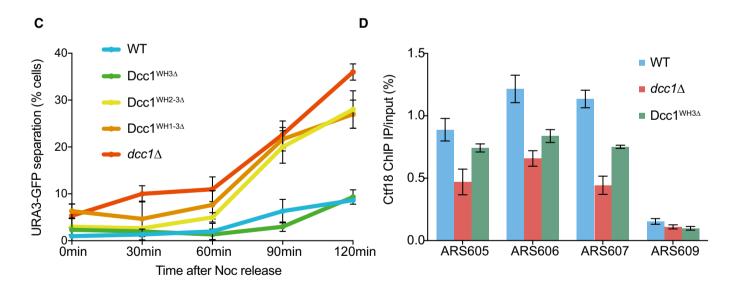


Figure 5. Functional analyses of WH domains.

- A Spot assay showing the effect of deleting WH domains. The resistance of dcc1\(\Delta\), Dcc1\(^{\Delta\)WH2-3}\) and Dcc1\(^{\Delta\)WH1-3}\) strains compared to wild-type cells in hydroxyurea (HU) was tested by spotting 10-fold serial dilutions on YPD agar containing either 0 M, 0.1 M or 0.2 M HU.
- B Checkpoint activation as assessed by Rad53 phosphorylation. Cells were arrested in G1 and released into medium containing 0.2 M HU for arrest in early S-phase. Samples were taken at the indicated time points and analysed for Rad53 phosphorylation by Western blotting. Tubulin was used as a loading control.
- C Sister chromatid cohesion assay using the same deletions as described above. The effect on sister chromatid cohesion was analysed by visualising tetracycline repressor-GFP fusion proteins bound to tetracycline operator arrays integrated at the *URA3* locus from cells arrested in a nocodazole-imposed mitotic arrest. The percentage of cells with two TetR-GFP dots were recorded. Three independent replicates were performed, and error bars indicate the standard deviation.
- D Analysis of Ctf18 levels at origins in HU-arrested cells. ChIP was conducted with an anti-HA antibody against HA-Ctf18. The enrichment was analysed by qPCR at three early firing origins (ARS605, ARS606 and ARS607) and a late-firing origin (ARS609) on the indicated yeast strains. The means and standard error of three independent experiments are plotted. Analysis of Ctf18 levels averaged over all three early origins by three-way ANOVA demonstrated statistically significant reductions in both the *dcc1*Δ (decrease = -0.555, *P* = 5.06897e-11) and Dcc1^{WH3Δ} (decrease = -0.3018, *P* = 1.1074164e-06) strains.

© 2017 The Francis Crick Institute EMBO reports Vol 18 | No. 4 | 2017 563

gel shift was observed with ssDNA but not dsDNA, indicating that WH1 and WH2 have a weak affinity for ssDNA, which is enhanced by the presence of WH3 (Fig EV3A). Dcc1 was also able to bind to dsDNA with a ssDNA extension at either the 3' or 5' end with an approximately equal affinity to single-stranded or duplex DNA (Fig EV3B). To identify individual residues involved in binding, we mutated basic residues in WH3 and compared them to wild-type (WT) protein. The residues K326, K357, K364, R367 and R380 (Fig 3D) were mutated to alanine and their effect on DNA binding assessed (Fig 4B). Mutation of K364, R367 or R380 reduced the affinity for both ssDNA and dsDNA, albeit weakly with K364, while the triple mutant K364A, R367A, R380A totally eliminated binding similar to the WH3 deletion (Fig 4C). These residues are widely conserved (Fig 4D) and suggest that the binding site for ssDNA and dsDNA overlaps to some extent.

Functional studies on the Dcc1 WH domains

Cells lacking Ctf18, Dcc1 or Ctf8 grown in the presence of hydroxyurea (HU) cannot properly activate the intra-S-phase checkpoint at stalled forks [16-18,58]. To further understand the function of the WH domains in Dcc1, we progressively deleted the WH domains and assessed checkpoint activation as assessed by Rad53 phosphorylation as well as viability when treated with hydroxyurea. We found that removal of any of the WH domains caused severe viability problems under HU treatment and showed delayed and reduced Rad53 phosphorylation, comparable to a complete Dcc1 deletion (Figs 5A and B, and EV4A). These defects are unlikely to be caused by impaired heterotrimer formation as Ctf18 and Ctf8 are still able to form a stable soluble complex with the WH3 or WH2-WH3 truncated Dcc1 proteins (Fig EV4B). The WH1-WH3 deletion is insoluble when recombinantly expressed but stable $Dcc1^{\Delta WH1-WH3}$ can be observed in cell extracts (Fig EV4A) suggesting it is also capable of binding Ctf8 and Ctf18, as free Dcc1 is quite unstable and rapidly degraded (data not shown).

We also find that deletion of the WH domains leads to sister chromatid cohesion defects (Fig 5C). Interestingly, however, while deletion of any of the WH domains had a severe checkpoint phenotype, deletion of the WH3 domain alone had a minimal impact on sister chromatid cohesion. We further analysed the function of the WH3 domain by chromatin immunoprecipitation (ChIP) and qPCR on Ctf18 at sites of replication origins in hydroxyurea-treated cells (Figs 5D and EV4C). Cells without the Dcc1 WH3 domain showed a statistically significant reduction of Ctf18 at replication origins, though not as much as complete Dcc1 deletions (Fig 5D). As Ctf18 deletion does not compromise entry into S-phase [16], and the truncated complex is still stable (Fig EV4B), this result suggests impaired recruitment of the RFC^{Ctf18} complex to stalled forks.

Discussion

RFC^{Ctf18} is a seven-subunit complex formed of Ctf18, Rfc2-5 and the stable Dcc1-Ctf8 sub-complex. The complex is required for activation of the intra-S-phase checkpoint and proper establishment of sister chromatid cohesion. To better understand the

function of RFC^{Ctf18}, we determined the crystal structure of the heterotrimeric Ctf18^C-Dcc1-Ctf8 complex and analysed *in vitro* and *in vivo* activities.

The structure of Ctf18^C-Dcc1-Ctf8 was completely unexpected and revealed that Dcc1 has three consecutive WH domains at its Cterminus, the third of which can bind to both ssDNA and dsDNA. To our knowledge, this is a completely novel organisation of WH domains. The N-terminus of Dcc1 and Ctf8 form a "triple" β-barrel fold, which can then bind Ctf18. This overall architecture has previously been found in a range of transcription factors and RNase H2. Like the transcription factors, RFCCtf18 uses this domain as a bridge between a larger complex and a variable number of WH domains able to bind DNA. This suggests a common ancestor of both a subset of transcription factors and RFCCtf18. RFC complexes are able to directly recognise and bind DNA structures such as template-primer junctions via multiple Rfc subunits [21,59]. The discovery that RFC^{Ctf18} contains additional binding sites for both dsDNA and ssDNA suggests that directed recruitment is an important part of its mechanism, as the intrinsic DNA binding activity of the core complex is sufficient for the basic loading and unloading activity [39,40].

Recent studies have also shown that RFCCtf18 is able to bind to DNA polymerase epsilon (Pol ε) via the catalytic subunit Pol2 [57,60,61]. This interaction requires the intact Ctf18^C-Ctf8-Dcc1 complex which we present in this study. Our structure shows that mutations in, or deletion of the Ctf18 C-terminus would dissociate the Dcc1-Ctf8 heterodimer from the complex, so disrupting the platform for Pol ϵ binding. The addition of two non-Rfc subunits to the RFC^{Ctf18} complex is reminiscent of the bacterial clamp loader, which employs two auxiliary subunits, χ and ψ to link the loader to single-stranded DNA-binding protein (SSB) upstream of the replication fork [62,63]. Dcc1 and Ctf8 bear no structural similarity to χ and ψ [64]. However, the structure of the bacterial clamp loader bound to a peptide derived from the ψ protein [65] as well as the location of the C-terminus of Rfc1 in the RFC-PCNA structure [66] suggests that the linker region of Ctf18 emerges adjacent to the "collar" domain of RFC. This would allow the Dcc1 subunit to interact with both ssDNA emerging from the clamp loader as well as Pol ε, situated on the same face of the complex (Fig 6).

Our structure shows that the Ctf8-Dcc1 interface acts as a binding site for the Ctf18 C-terminus, which together forms a platform capable of binding Pol 2. Given the proximity of the WH domains to this binding site, it seems likely that they are also capable of interacting with the Pol ϵ complex, and/or DNA within the replisome. The reduced occupancy of Ctf18 on chromatin observed following WH3 deletion supports this notion. Although all the WH domains are required for full checkpoint activation, we find that only WH1 and WH2 are required for correct sister chromatid cohesion. This separation of function might be explained if the cohesion pathway is more tolerant to reduced levels of chromatin bound Ctf18 relative to the checkpoint pathway. Alternatively, the WH3 domain might form a checkpoint-specific interaction with other proteins.

It has been reported that RFC^{Ctf18} also acts in additional pathways, for example telomere positioning and stability [67–69] and opens up the intriguing possibility that Dcc1 has other, as yet undetermined activities that require robust DNA interactions. Determining the precise details of exactly how and where this multi-faceted protein complex contributes to DNA replication and chromosome segregation activities will provide an interesting ongoing challenge.

564 EMBO reports Vol 18 | No 4 | 2017 (© 2017 The Francis Crick Institute

Benjamin O Wade et al Structural studies of RFC^{Ctf18}

EMBO reports

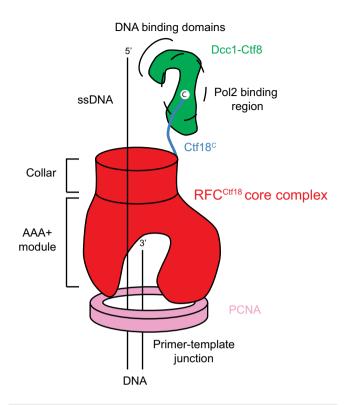


Figure 6. Organisation of the RFC^{Ctf18} complex. Schematic diagram showing the location of the Ctf8-Dcc1 heterodimer in relation to the clamp loader/clamp and likely location of the Pol ϵ binding site.

Materials and Methods

Protein expression and purification

The genes coding for the subunits Ctf18 (residues 666-741), Dcc1 (full length) and Ctf8 (full length) were amplified from genomic DNA and cloned into the vectors pET-28a, pCDFDuet and pETDuet, respectively. The Ctf18 subunit had the addition of an N-terminal His₆ tag and TEV protease site. These subunits were then sequenced to confirm that no mutations had been introduced. The vectors containing the subunits were transformed into BL21(DE3)-RIL competent cells (Agilent) using standard procedures. An overnight culture of freshly transformed cells was grown at 37°C in LB, which was used to inoculate a 1 l culture. These cells were then grown at 37°C until an OD of 0.6 was reached, upon which expression of the complex was induced by the addition of 0.25 mM IPTG at 16°C for 18 h. The cells were harvested and sonicated in 50 mM Tris pH 7.5, 300 mM NaCl, 10% glycerol and 0.5 mM TCEP and then centrifuged at 35,000 \times g, 4°C for 1 h. The supernatant was loaded onto a 5 ml His-Trap HP (GE Healthcare) and eluted using imidazole. The tag was then cleaved from the Ctf18 subunit by the addition of TEV protease at 0.04 mg/ml and left overnight at 4°C. The protein was loaded onto a MonoQ 10/300 GL (GE Healthcare) column equilibrated in 20 mM Tris pH 7.5, 100 mM NaCl and 0.5 mM TCEP. The protein was eluted in a high concentration of NaCl. The complex was concentrated and loaded on an equilibrated HiLoad 16/60 Superdex 75 column (GE Healthcare) in 10 mM Tris pH 7.5, 150 mM NaCl and 0.5 mM TCEP. Eluted protein-containing fractions were analysed for purity by SDS–PAGE. Selenomethionine incorporated protein was obtained by growing the cells in minimal media where selenomethionine was the only source of methionine and purification of the protein was identical to the native protein. The Dcc1 truncation (residues 90–380) was expressed and purified using the same protocol as for the full complex.

Crystallisation and structure solution

The selenomethionine incorporated C-terminus of Dcc1 was concentrated to 20 mg/ml and crystallised in 2 μ l sitting drops with equal volumes of protein to reservoir solution at 20°C. The reservoir contained 0.2 M potassium sodium tartrate and 18% PEG 3350. Crystals were harvested after 2 days in 20% ethylene glycol. A two wavelength MAD data set was collected at Diamond Light Source on beamline IO3. The AutoSHARP [70] package was used to solve the structure, and Buccaneer [71] was used to trace the initial model. Iterative rounds of rebuilding and refinement were carried out in Coot [72] and Refmac5 [73]. The final model was then fitted by molecular replacement to a higher resolution native data set obtained from a different crystal form grown in the same conditions. Ctf18^C-Dcc1-Ctf8 was concentrated to 35 mg/ml and was crystallised in 4 μ l sitting drops with equal volumes of protein to reservoir solution. The reservoir contained 0.1 M Bis-Tris propane pH 6.3, 0.2 M NaBr and 17% PEG 3350. Crystals were grown at 20°C for 10 days before harvesting in 15% ethylene glycol and stored in liquid nitrogen. Native data were collected at Diamond Light Source on beamline IO3. Molecular replacement was carried out in Phaser [74] using the C-terminus of Dcc1 as a start model. An initial model was traced using phenix.autobuild [75], and iterative rounds of rebuilding and refinement were carried out as before. Some spots fell beyond the edge of the detector accounting for reduced completeness in the outer shell. Full statistics for data collection and refinement are shown in Appendix Table S1.

DNA binding analysis

DNA binding analysis was conducted on both dsDNA and ssDNA. The dsDNA had the sequence TCTCCACAGGAAACGGAGGGGT and a fluorescent 6-FAM tag on the 5' sense strand. The DNA was annealed and purified using standard techniques. The ssDNA used was a 25-mer oligo d(T), with a 5' 6-FAM fluorescent tag. Both oligos were synthesised and purchased from Sigma-Aldrich. A 6% non-denaturing DNA retardation gel was pre-run at 4°C at 3 mA for 1 h in 0.5× TBE. DNA was at 1 μM unless otherwise stated and protein titrated at increasing concentrations. Samples were loaded and run for 75 min at 5 mA, 4°C. The gel was then analysed using a Typhoon 9400 imager excited at 488 nm, and emission was detected at 520 nm at 500 V.

Yeast strains and cultures

All strains used in this study were of W303 background and are listed in Table 1. Gene deletions and epitope tagging of endogenous genes were performed by gene targeting using polymerase chain reaction (PCR) products [76,77]. Cells were grown in rich YP medium supplemented with 2% glucose as the carbon source [78]. Synchronisation

© 2017 The Francis Crick Institute EMBO reports Vol 18 | No 4 | 2017 565

Table 1. Yeast strains used in this study.

Y141 W303 wild type

Y4950 MATa dcc1Δ::LEU2

Y4951 MATa dcc1(WH3Δ)-Pk₉::TRP1

Y4325 MATa ctf18Δ::TRP1

Y1235 MATa CTF18-HA₆::HIS3

Y4952 MATa dcc1∆::LEU CTF18-HA6::KAN^R

K7100 MATa URA3::tetOs HIS3::tetR-GFP

Y4954 MATa ura3::3xURA3 tetO112, his3::HIS3 tetR-GFP dcc1∆::LEU2

Y4955 MATa ura3::3xURA3 tetO112, his3::HIS3 tetR-GFP dcc1(WH3Δ)-Pk₉::TRP1

Y5202 MATa ura3::3xURA3 tetO112, his3::HIS3 tetR-GFP DCC1-Pkg::TRP1

Y5203 MATa ura3::3xURA3 tetO112, his3::HIS3 tetR-GFP $dcc1(WH2-3\Delta)-Pk_9::TRP1$

Y5204 MATa ura3::3xURA3 tetO112, his3::HIS3 tetR-GFP $dcc1(WH1-3\Delta)-Pk_9::TRP1$

Y5245 MATa CTF18-HA6::HIS3 DCC1-Pk9::TRP1

Y5246 MATa CTF18-HA₆::HIS3 dcc1(WH3Δ)-Pk₉::TRP1

Y5247 MATa CTF18-HA₆::HIS3 dcc1(WH2-3Δ)-Pk₉::TRP1

Y5248 MATa CTF18-HA₆::HIS3 dcc1(WH1-3Δ)-Pk₉::TRP1

in G1 was performed by addition of α -factor (0.4 μ g/ml) for 2 h. To arrest cells in early S-phase, G1 synchronised cultures were filtered, washed and re-suspended in fresh medium containing 0.2 M hydroxyurea for 60 min. Arrest in G2/M was achieved by release into medium containing 5 μ g/ml nocodazole for 120 min.

Yeast molecular biology techniques

Chromatin immunoprecipitation was performed as previously described [79], as was the quantitative analysis of chromatin immunoprecipitates [80]. Analysis of sister chromatid cohesion was done by visualising tetracycline repressor-GFP fusion proteins bound to tetracycline operator arrays integrated at the *URA3* locus on chromosome 5, as previously described [81]. Hydroxyurea (HU) resistance was tested by spotting 10-fold serial dilutions of cultured strains on YPD agar containing 0 M, 0.1 M or 0.2 M HU. Antibodies used for Western blotting or chromatin immunoprecipitation were anti-HA (clone F-7, Santa Cruz), anti-tubulin (clone YOL1/34, Serotec) and an antibody specific to Rad53 (ab104232, Abcam).

Data availability

Coordinates and structure factors of Ctf18^C-Ctf8-Dcc1 and Dcc1^{90–380} structures have been deposited in the Protein Data Bank under accession codes 5MSM and 5MSN, respectively.

Expanded View for this article is available online.

Acknowledgements

We wish to thank A. Purkiss for assistance with X-ray data collection, the Crick Proteomics platform for mass spectrometric analyses and G. Kelly for assistance with data analysis. This work was supported by the Francis Crick Institute, which receives its core funding from Cancer Research UK (FC001155, FC001198), the Medical Research Council (FC001155, FC001198) and the Wellcome Trust (FC001155, FC001198).

Author contributions

BOW cloned, expressed, purified and crystallised all constructs and carried out DNA binding experiments. BOW and MRS collected and analysed X-ray data. CPS and HWL carried out *in vivo* experiments. BOW, FU and MRS conceived and designed the study. All authors provided input into the manuscript.

Conflict of interest

The authors declare that they have no conflict of interest.

References

- Errico A, Costanzo V (2012) Mechanisms of replication fork protection: a safeguard for genome stability. Crit Rev Biochem Mol Biol 47: 222 – 235
- Friedel AM, Pike BL, Gasser SM (2009) ATR/Mec1: coordinating fork stability and repair. Curr Opin Cell Biol 21: 237 – 244
- Labib K, De Piccoli G (2011) Surviving chromosome replication: the many roles of the S-phase checkpoint pathway. Philos Trans R Soc Lond B Biol Sci 366: 3554 – 3561
- Zegerman P, Diffley JFX (2009) DNA replication as a target of the DNA damage checkpoint. DNA Repair 8: 1077 – 1088
- 5. Zou L (2013) Four pillars of the S-phase checkpoint. Genes Dev 27: 227 233
- Zou L, Liu D, Elledge SJ (2003) Replication protein A-mediated recruitment and activation of Rad17 complexes. Proc Natl Acad Sci USA 100: 13827 – 13832
- Katou Y, Kanoh Y, Bando M, Noguchi H, Tanaka H, Ashikari T, Sugimoto K, Shirahige K (2003) S-phase checkpoint proteins Tof1 and Mrc1 form a stable replication-pausing complex. Nature 424: 1078–1083
- Alcasabas AA, Osborn AJ, Bachant J, Hu F, Werler PJ, Bousset K, Furuya K, Diffley JF, Carr AM, Elledge SJ (2001) Mrc1 transduces signals of DNA replication stress to activate Rad53. Nat Cell Biol 3: 958 – 965
- Kumagai A, Dunphy WG (2000) Claspin, a novel protein required for the activation of Chk1 during a DNA replication checkpoint response in Xenopus egg extracts. Mol Cell 6: 839–849
- Osborn AJ, Elledge SJ (2003) Mrc1 is a replication fork component whose phosphorylation in response to DNA replication stress activates Rad53. Genes Dev 17: 1755–1767
- Santocanale C, Diffley JF (1998) A Mec1- and Rad53-dependent checkpoint controls late-firing origins of DNA replication. *Nature* 395: 615-618
- Shirahige K, Hori Y, Shiraishi K, Yamashita M, Takahashi K, Obuse C, Tsurimoto T, Yoshikawa H (1998) Regulation of DNA-replication origins during cell-cycle progression. *Nature* 395: 618 – 621
- Zhao X, Muller EG, Rothstein R (1998) A suppressor of two essential checkpoint genes identifies a novel protein that negatively affects dNTP pools. Mol Cell 2: 329–340
- 14. Zhou Z, Elledge SJ (1993) DUN1 encodes a protein kinase that controls the DNA damage response in yeast. *Cell* 75: 1119–1127
- Huang M, Zhou Z, Elledge SJ (1998) The DNA replication and damage checkpoint pathways induce transcription by inhibition of the Crt1 repressor. Cell 94: 595–605
- Crabbé L, Thomas A, Pantesco V, De Vos J, Pasero P, Lengronne A (2010)
 Analysis of replication profiles reveals key role of RFC-Ctf18 in yeast replication stress response. Nat Struct Mol Biol 17: 1391–1397
- Kubota T, Hiraga S-I, Yamada K, Lamond AI, Donaldson AD (2011) Quantitative proteomic analysis of chromatin reveals that Ctf18 acts in the DNA replication checkpoint. Mol Cell Proteomics 10: M110.005561

566 EMBO reports Vol 18 | No 4 | 2017 (© 2017 The Francis Crick Institute

- Naiki T, Kondo T, Nakada D, Matsumoto K, Sugimoto K (2001) Chl12 (Ctf18) forms a novel replication factor C-related complex and functions redundantly with Rad24 in the DNA replication checkpoint pathway. Mol Cell Biol 21: 5838 – 5845
- Kelch BA, Makino DL, O'Donnell M, Kuriyan J (2012) Clamp loader
 ATPases and the evolution of DNA replication machinery. BMC Biol 10: 34
- O'Donnell M, Kuriyan J (2006) Clamp loaders and replication initiation.
 Curr Opin Struct Biol 16: 35–41
- 21. Majka J, Burgers PMJ (2004) The PCNA-RFC families of DNA clamps and clamp loaders. *Prog Nucleic Acid Res Mol Biol* 78: 227–260
- Green CM, Erdjument-Bromage H, Tempst P, Lowndes NF (2000) A novel Rad24 checkpoint protein complex closely related to replication factor C. Curr Biol 10: 39 – 42
- Shimomura T, Ando S, Matsumoto K, Sugimoto K (1998) Functional and physical interaction between Rad24 and Rfc5 in the yeast checkpoint pathways. Mol Cell Biol 18: 5485–5491
- Ia Torre-Ruiz de MA, Green CM, Lowndes NF (1998) RAD9 and RAD24 define two additive, interacting branches of the DNA damage checkpoint pathway in budding yeast normally required for Rad53 modification and activation. EMBO J 17: 2687 – 2698
- Naiki T, Shimomura T, Kondo T, Matsumoto K, Sugimoto K (2000) Rfc5, in cooperation with rad24, controls DNA damage checkpoints throughout the cell cycle in Saccharomyces cerevisiae. Mol Cell Biol 20: 5888 – 5896
- Bao S, Shen X, Shen K, Liu Y, Wang XF (1998) The mammalian Rad24 homologous to yeast Saccharomyces cerevisiae Rad24 and Schizosaccharomyces pombe Rad17 is involved in DNA damage checkpoint. Cell Growth Differ 9: 961–967
- Kubota T, Nishimura K, Kanemaki MT, Donaldson AD (2013) The Elg1 replication factor C-like complex functions in PCNA unloading during DNA replication. Mol Cell 50: 273 – 280
- Kubota T, Katou Y, Nakato R, Shirahige K, Donaldson AD (2015) Replication-coupled PCNA unloading by the Elg1 complex occurs genome-wide and requires Okazaki fragment ligation. Cell Rep 12: 774 787
- Lee K-Y, Fu H, Aladjem MI, Myung K (2013) ATAD5 regulates the lifespan of DNA replication factories by modulating PCNA level on the chromatin. J Cell Biol 200: 31 – 44
- Smolikov S, Mazor Y, Krauskopf A (2004) ELG1, a regulator of genome stability, has a role in telomere length regulation and in silencing. Proc Natl Acad Sci USA 101: 1656–1661
- Ben-Aroya S, Koren A, Liefshitz B, Steinlauf R, Kupiec M (2003) ELG1, a yeast gene required for genome stability, forms a complex related to replication factor C. Proc Natl Acad Sci USA 100: 9906–9911
- Kanellis P, Agyei R, Durocher D (2003) Elg1 forms an alternative PCNAinteracting RFC complex required to maintain genome stability. Curr Biol 13: 1583–1595
- Bellaoui M, Chang M, Ou J, Xu H, Boone C, Brown GW (2003) Elg1 forms an alternative RFC complex important for DNA replication and genome integrity. EMBO J 22: 4304 – 4313
- Banerjee S, Myung K (2004) Increased genome instability and telomere length in the elg1-deficient Saccharomyces cerevisiae mutant are regulated by S-phase checkpoints. Eukaryot Cell 3: 1557–1566
- 35. Maradeo ME, Skibbens RV (2009) The Elg1-RFC clamp-loading complex performs a role in sister chromatid cohesion. *PLoS ONE* 4: e4707
- Parnas O, Zipin-Roitman A, Mazor Y, Liefshitz B, Ben-Aroya S, Kupiec M (2009) The ELG1 clamp loader plays a role in sister chromatid cohesion. PLoS ONE 4: e5497

- Mayer ML, Gygi SP, Aebersold R, Hieter P (2001) Identification of RFC(Ctf18p, Ctf8p, Dcc1p): an alternative RFC complex required for sister chromatid cohesion in Saccharomyces cerevisiae. Mol Cell 7: 959 – 970
- Hanna JS, Kroll ES, Lundblad V, Spencer FA (2001) Saccharomyces cerevisiae CTF18 and CTF4 are required for sister chromatid cohesion. Mol Cell Biol 21: 3144 – 3158
- Bermudez VP, Maniwa Y, Tappin I, Ozato K, Yokomori K, Hurwitz J
 (2003) The alternative Ctf18-Dcc1-Ctf8-replication factor C complex required for sister chromatid cohesion loads proliferating cell nuclear antigen onto DNA. Proc Natl Acad Sci USA 100: 10237 – 10242
- Bylund GO, Burgers PMJ (2005) Replication protein A-directed unloading of PCNA by the Ctf18 cohesion establishment complex. Mol Cell Biol 25: 5445-5455
- Shiomi Y, Shinozaki A, Sugimoto K, Usukura J, Obuse C, Tsurimoto T (2004) The reconstituted human Chl12-RFC complex functions as a second PCNA loader. Genes Cells 9: 279 – 290
- Lengronne A, McIntyre J, Katou Y, Kanoh Y, Hopfner K-P, Shirahige K, Uhlmann F (2006) Establishment of sister chromatid cohesion at the Saccharomyces cerevisiae replication fork. Mol Cell 23: 787 – 799
- Borges V, Smith DJ, Whitehouse I, Uhlmann F (2013) An Eco1-independent sister chromatid cohesion establishment pathway in Saccharomyces cerevisiae. Chromosoma 122: 121–134
- Mayer ML, Pot I, Chang M, Xu H, Aneliunas V, Kwok T, Newitt R, Aebersold R, Boone C, Brown GW et al (2004) Identification of protein complexes required for efficient sister chromatid cohesion. Mol Biol Cell 15: 1736–1745
- 45. Petronczki M, Chwalla B, Siomos MF, Yokobayashi S, Helmhart W, Deutschbauer AM, Davis RW, Watanabe Y, Nasmyth K (2004) Sister-chromatid cohesion mediated by the alternative RF-CCtf18/Dcc1/Ctf8, the helicase Chl1 and the polymerase-α-associated protein Ctf4 is essential for chromatid disjunction during meiosis II. J Cell Sci 117: 3547 3559
- 46. Skibbens RV (2004) Chl1p, a DNA helicase-like protein in budding yeast, functions in sister-chromatid cohesion. *Genetics* 166: 33–42
- Warren CD, Eckley DM, Lee MS, Hanna JS, Hughes A, Peyser B, Jie C, Irizarry R, Spencer FA (2004) S-phase checkpoint genes safeguard highfidelity sister chromatid cohesion. Mol Biol Cell 15: 1724–1735
- Holm L, Sander C (1993) Protein structure comparison by alignment of distance matrices. J Mol Biol 233: 123–138
- 49. Kim D, Hur J, Park K, Bae S, Shin D, Ha SC, Hwang H-Y, Hohng S, Lee J-H, Lee S *et al* (2014) Distinct Z-DNA binding mode of a PKR-like protein kinase containing a Z-DNA binding domain (PKZ). *Nucleic Acids Res* 42: 5937 5948
- Scott DC, Monda JK, Bennett EJ, Harper JW, Schulman BA (2011)
 N-terminal acetylation acts as an avidity enhancer within an interconnected multiprotein complex. Science 334: 674–678
- Chatake T, Mizuno N, Voordouw G, Higuchi Y, Arai S, Tanaka I, Niimura N (2003) Crystallization and preliminary neutron analysis of the dissimilatory sulfite reductase D (DsrD) protein from the sulfate-reducing bacterium *Desulfovibrio vulgaris*. Acta Crystallogr D Biol Crystallogr 59: 2306 2309
- Harami GM, Gyimesi M, Kovács M (2013) From keys to bulldozers: expanding roles for winged helix domains in nucleic-acid-binding proteins. Trends Biochem Sci 38: 364–371
- Shaban NM, Harvey S, Perrino FW, Hollis T (2010) The structure of the mammalian RNase H2 complex provides insight into RNA.NA hybrid processing to prevent immune dysfunction. J Biol Chem 285: 3617–3624

567

© 2017 The Francis Crick Institute EMBO reports Vol 18 | No 4 | 2017

- Gaiser F, Tan S, Richmond TJ (2000) Novel dimerization fold of RAP30/RAP74 in human TFIIF at 1.7 A resolution. J Mol Biol 302: 1119–1127
- Taylor NMI, Baudin F, von Scheven G, Müller CW (2013) RNA polymerase
 III-specific general transcription factor IIIC contains a heterodimer
 resembling TFIIF Rap30/Rap74. Nucleic Acids Res 41: 9183–9196
- Geiger SR, Lorenzen K, Schreieck A, Hanecker P, Kostrewa D, Heck AJR, Cramer P (2010) RNA polymerase I contains a TFIIF-related DNA-binding subcomplex. Mol Cell 39: 583 – 594
- 57. García-Rodríguez LJ, De Piccoli G, Marchesi V, Jones RC, Edmondson RD, Labib K (2015) A conserved Polε binding module in Ctf18-RFC is required for S-phase checkpoint activation downstream of Mec1. *Nucleic Acids Res* 43: 8830 8838
- 58. Pan X, Ye P, Yuan DS, Wang X, Bader JS, Boeke JD (2006) A DNA integrity network in the yeast *Saccharomyces cerevisiae*. *Cell* 124: 1069–1081
- Gomes XV, Burgers PM (2001) ATP utilization by yeast replication factor
 I. ATP-mediated interaction with DNA and with proliferating cell nuclear antigen. *J Biol Chem* 276: 34768 – 34775
- 60. Murakami T, Takano R, Takeo S, Taniguchi R, Ogawa K, Ohashi E, Tsurimoto T (2010) Stable interaction between the human proliferating cell nuclear antigen loader complex Ctf18-replication factor C (RFC) and DNA polymerase is mediated by the cohesion-specific subunits, Ctf18, Dcc1, and Ctf8. J Biol Chem 285: 34608 – 34615
- 61. Okimoto H, Tanaka S, Araki H, Ohashi E, Tsurimoto T (2016) Conserved interaction of Ctf18-RFC with DNA polymerase ϵ is critical for maintenance of genome stability in *Saccharomyces cerevisiae*. *Genes Cells* 21: 482 491
- 62. Glover BP, McHenry CS (1998) The chi psi subunits of DNA polymerase III holoenzyme bind to single-stranded DNA-binding protein (SSB) and facilitate replication of an SSB-coated template. *J Biol Chem* 273: 23476–23484
- Kelman Z, Yuzhakov A, Andjelkovic J, O'Donnell M (1998) Devoted to the lagging strand-the subunit of DNA polymerase III holoenzyme contacts SSB to promote processive elongation and sliding clamp assembly. EMBO J 17: 2436 – 2449
- 64. Gulbis JM, Kazmirski SL, Finkelstein J, Kelman Z, O'Donnell M, Kuriyan J (2004) Crystal structure of the chi:psi subassembly of the *Escherichia coli* DNA polymerase clamp-loader complex. *Eur J Biochem* 271: 439 449
- Simonetta KR, Kazmirski SL, Goedken ER, Cantor AJ, Kelch BA, McNally R, Seyedin SN, Makino DL, O'Donnell M, Kuriyan J (2009) The mechanism of ATP-dependent primer-template recognition by a clamp loader complex. Cell 137: 659–671
- Bowman GD, O'Donnell M, Kuriyan J (2004) Structural analysis of a eukaryotic sliding DNA clamp-clamp loader complex. *Nature* 429: 724 – 730
- 67. Hiraga S-I, Robertson ED, Donaldson AD (2006) The Ctf18 RFC-like complex positions yeast telomeres but does not specify their replication time. *EMBO J* 25: 1505–1514

- Pike BL, Heierhorst J (2007) Mdt1 facilitates efficient repair of blocked DNA double-strand breaks and recombinational maintenance of telomeres. Mol Cell Biol 27: 6532 – 6545
- Gao H, Moss DL, Parke C, Tatum D, Lustig AJ (2014) The Ctf18RFC clamp loader is essential for telomere stability in telomerase-negative and mre11 mutant alleles. PLoS ONE 9: e88633
- Bricogne G, Vonrhein C, Flensburg C, Schiltz M, Paciorek W (2003)
 Generation, representation and flow of phase information in structure determination: recent developments in and around SHARP 2.0. Acta Crystallogr D Biol Crystallogr 59: 2023 2030
- Cowtan K (2006) The Buccaneer software for automated model building.
 Tracing protein chains. Acta Crystallogr D Biol Crystallogr 62: 1002–1011
- 72. Emsley P, Cowtan K (2004) Coot: model-building tools for molecular graphics. *Acta Crystallogr D Biol Crystallogr* 60: 2126 2132
- Murshudov GN, Skubak P, Lebedev AA, Pannu NS, Steiner RA, Nicholls RA, Winn MD, Long F, Vagin AA (2011) REFMAC5 for the refinement of macromolecular crystal structures. Acta Cryst 67: 355 – 367
- McCoy AJ, Grosse-Kunstleve RW, Adams PD, Winn MD, Storoni LC, Read RJ (2007) Phaser crystallographic software. J Appl Crystallogr 40: 658 – 674
- Adams PD, Afonine PV, Bunkóczi G, Chen VB, Davis IW, Echols N, Headd JJ, Hung L-W, Kapral GJ, Grosse-Kunstleve RW et al (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr D Biol Crystallogr 66: 213 – 221
- Knop M, Siegers K, Pereira G, Zachariae W, Winsor B, Nasmyth K, Schiebel E (1999) Epitope tagging of yeast genes using a PCR-based strategy: more tags and improved practical routines. Yeast 15: 963-972
- Wach A, Brachat A, Pöhlmann R, Philippsen P (1994) New heterologous modules for classical or PCR-based gene disruptions in Saccharomyces cerevisiae. Yeast 10: 1793–1808
- Amberg DC, Burke D, Strathern JN, Laboratory CSH (2005) Methods in yeast genetics. Cold Spring Harbor, NY: Cold Spring Harbor Laboratory Press
- Lengronne A, Katou Y, Mori S, Yokobayashi S, Kelly GP, Itoh T, Watanabe Y, Shirahige K, Uhlmann F (2004) Cohesin relocation from sites of chromosomal loading to places of convergent transcription. *Nature* 430: 573 – 578
- Lopez-Serra L, Kelly G, Patel H, Stewart A, Uhlmann F (2014) The Scc2-Scc4 complex acts in sister chromatid cohesion and transcriptional regulation by maintaining nucleosome-free regions. Nat Genet 46: 1147 – 1151
- Michaelis C, Ciosk R, Nasmyth K (1997) Cohesins: chromosomal proteins that prevent premature separation of sister chromatids. Cell 91: 35 – 45



License: This is an open access article under the terms of the Creative Commons Attribution 4.0 License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

568 EMBO reports Vol 18 | No 4 | 2017 © 2017 The Francis Crick Institute