Insights into Centromere DNA Bending Revealed by the Cryo-EM Structure of the Core Centromere Binding Factor 3 with Ndc10

Highlights

- Cryo-EM studies of CBF3 reveal the core complex has a deep asymmetric channel
- The size, conservation, and charge of the channel suggest that it binds centromere DNA
- Unique insertions in the Ctf13 F box provide the binding site for one Ndc10 monomer
- The Ndc10 DNA-binding site is in plane with and perpendicular to the CBF3 channel

Authors

Wenjuan Zhang, Natalya Lukoynova, Shomon Miah, Jonathan Lucas, Cara K. Vaughan

c.vaughan@mail.cryst.bbk.ac.uk

In Brief

Zhang et al. determine two cryo-EM structures of the budding yeast centromere binding factor CBF3. These structures show that CBF3 must bend the centromere DNA element CDEIII, and this may promote looping of the centromere to accommodate a centromeric nucleosome.

Data and Software Availability

6FE8
6GSA

Zhang et al., 2018, Cell Reports 24, 744–754
July 17, 2018 © 2018 The Authors.
https://doi.org/10.1016/j.celrep.2018.06.068
INSIGHTS INTO CENTROMERE DNA BENDING REVEALED BY THE CRYO-EM STRUCTURE OF THE CORE CENTROMERE BINDING FACTOR 3 WITH NDC10

Wenjuan Zhang,1,2 Natalya Lukoyanova,1 Shomon Miah,1 Jonathan Lucas,2 and Cara K. Vaughan1,3,*

1Institute of Structural and Molecular Biology, Birkbeck College, Malet Street, London WC1E 7HX, UK
2Institute of Structural and Molecular Biology, University College London, Gower Street, London WC1E 6BT, UK
3Lead Contact
*Correspondence: c.vaughan@mail.cryst.bbk.ac.uk
https://doi.org/10.1016/j.celrep.2018.06.068

SUMMARY

The centromere binding factor 3 (CBF3) complex binds the third centromere DNA element in organisms with point centromeres, such as S. cerevisiae. It is an essential complex for assembly of the kinetochores in these organisms, as it facilitates genetic centromere specification and allows association of all other kinetochores’ components. We determined high-resolution structures of the core complex of CBF3 alone and in association with a monomeric construct of Ndc10, using cryoelectron microscopy (cryo-EM). We identify the DNA-binding site of the complex and present a model in which CBF3 induces a tight bend in centromeric DNA, thus facilitating assembly of the centromeric nucleosome.

INTRODUCTION

The integrity of genetic information passed through generations relies on faithful segregation of chromosomes during mitosis. The kinetochore, a mega-Dalton protein assembly, enables this segregation by specifically associating with both the centromere (CEN) of sister chromatids and the microtubules of the mitotic spindle. Most eukaryotes have regional CENs with unique satellite repeats that vary in length but whose arrangement is largely conserved between chromosomes despite being unconserved in sequence. Regional CENs, which can be up to 5 Mb in length in humans, are specified epigenomically by the presence of an essential centromeric histone H3 variant, CENP-A, which is found at CEN DNA, where it is interspersed with canonical nucleosomes (Verdaasdonk and Bloom, 2011).

In contrast, budding yeasts, including S. cerevisiae, have evolved point CENs, comprising three CEN DNA elements (CDEs), typically of ~125 bp. CDEI (8 bp) and CDEII (~26 bp) are relatively conserved and border CDEII (78–86 bp), which, although largely unconserved at a base level, is characterized by generally greater than 90% A-T content (Fitzgerald-Hayes et al., 1982). Point CENs evolved from an ancestor with an epigenomically specified CEN (Malik and Henikoff, 2009), and this evolutionary transition introduced genetic specification of the CEN while retaining aspects of epigenomic specification, in particular the essential presence of the CENP-A homolog Cse4.

Unique to these organisms, the CEN binding factor 3 (CBF3) complex provides a physical link between the genetic and epigenomic mechanisms of CEN specification. It associates specifically with CDEII (Lechner and Carbon, 1991) and is responsible for deposition of Cse4, through a direct interaction with the centromeric histone chaperone Scm3/HJURP (Camahort et al., 2007). Its epigenomic role is emphasized by the observation that a synthetic kinetochore can be assembled in the absence of CDEs, provided functional CBF3 is available to stabilize Cse4 incorporation at the CEN (Ho et al., 2014).

CBF3 comprises two homodimers, of Cep3 and Ndc10, and a Ctf13-Skp1 heterodimer. Cep3 provides sequence specificity through binuclear zinc-cluster (Zn2Cys6) domains homologous to those found in GAL4-like fungal transcription factors (Lechner, 1994; MacPherson et al., 2006). These domains bind a pseudo-TGT/CCG palindrome in CDEIII (Espelin et al., 1997). Ndc10 contributes both non-specific DNA binding (Cho and Harrison, 2011; Espelin et al., 1997) and association with Scm3/HJURP (Camahort et al., 2007). The Skp1-Ctf13 heterodimer interacts with Cep3, Ndc10, and CDEIII at a completely conserved G, centrally positioned between the TGT/CCG half-sites (Espelin et al., 1997).

The molecular mechanism of CBF3 association with CDEIII is unknown, despite a wealth of genetic, biochemical, and structural data (Bellizzi et al., 2007; Cho and Harrison, 2011; Leber et al., 2018; Perriches and Singleton, 2012; Purvis and Singleton, 2008), because of an absence of atomic-resolution structural information for the CBF3 complex as a whole.

Herein we present cryoelectron microscopy (cryo-EM) structures of the CBF3 core complex (CBF3CC4N) alone at atomic resolution and in complex with a monomeric construct of Ndc10 (CBF3CC4N + Ndc10 D1-2). The core complex forms a deep channel that is both highly charged and strongly conserved and is perfectly sized to accommodate DNA. The structure of CBF3CC4N + Ndc10 D1-2 identifies the Ndc10 D1-2 interface on the core complex and indicates a second DNA-binding interface that is in plane with, and perpendicular to, the channel of the core. Combining our in vitro experiments, which identify the channel as the likely site of DNA association with the core complex, and existing structural data for the DNA-binding site of Ndc10 D1-2, we propose a model that accounts for previously published genetic, biochemical, and biophysical experiments and implicates the core complex as well as Ndc10 in bending DNA to accommodate the centromeric nucleosome.
RESULTS

Cryo-EM Studies of CBF3

Dissection of the assembly and turnover of CBF3 in vivo indicated that Cep3 and Ctf13 associate early in the CBF3 assembly pathway, whereas association with Ndc10 occurred at much later time points (Rodrigo-Brenni et al., 2004). These results were in agreement with earlier in vitro reconstitution experiments (Russell et al., 1999); therefore we expressed and purified a CBF3CC comprising the Ctf13-Skp1 heterodimer and the Cep3 homodimer (Figure 1A). Full-length Cep3 rendered CBF3CC relatively unstable, but the complex with an N-terminally truncated Cep3, in which the Zn$_2$Cys$_6$ domains were missing, yielded a 220 kDa complex (CBF3CC$^{DN}$) that was purified to homogeneity and was suitable for cryo-EM studies (Figures S1A and S1B). Two-dimensional (2D) classification of motion-corrected cryo-EM images generated class averages with clear secondary structural details (Figure S1C) that enabled de novo reconstruction of a three-dimensional (3D) map (Figure S1D). After refinement, the final map was calculated from 187,606 particles, with a resolution of 3.7 Å judged using the Fourier shell correlation (FSC) gold standard method (Figures S1E–S1H; Table S1). The CBF3CC$^{DN}$ map has a horseshoe shape with a deep central channel (Figure 1B).

Ndc10 is the remaining and largest component of the CBF3 complex (110 kDa) which, like Cep3, is dimeric. Current data suggest that each of its five domains contributes different functions (Cho and Harrison, 2011): domain 1 associates with both the CBF3 core and Cbf1, a non-essential CDEI-binding protein; domain 2 provides the principal DNA-binding interface; domain 3 contributes dimerization; and domains 4 and 5 are responsible for recruitment of Scm3/HJURP.

In order to identify the binding site for Ndc10 on CBF3CC$^{DN}$ we assembled a complex of CBF3CC$^{DN}$ bound to a monomeric construct of Ndc10 comprising domains 1 and 2 (Ndc10 D1-2). After co-elution by size exclusion chromatography, this complex was suitable for cryo-EM studies. Two-dimensional classification revealed high-resolution classes in which secondary structural elements could be identified, and these data generated a 4.2 Å map (Figures 1C and S2A–S2E; Table S1).

Atomic Model of CBF3CC$^{DN}$

Secondary structure details and side chains were clearly visible throughout the CBF3CC$^{DN}$ map (Figures 2A–2D), enabling an atomic model for the core complex to be built and refined (Figure 2E; Table S2). The Cep3 homodimer was readily recognized, facilitated by its largely helical structure and the identification of an approximately two-fold axis, which could also be seen in the 2D classes (Figure S1C). The 2.5 Å crystal structure of the Cep3$^{DN}$ homodimer (Purvis and Singleton, 2008) was fit in the map as a rigid body, and this structure is essentially unchanged after refinement with a final C$\alpha$ RMSD of 1.1 Å. The consistently
high quality of the remaining density, assigned to the Skp1-Ctf13 heterodimer, has highlighted significant structural and chemical details that were not visible in the previously published polyalanine model of Ctf13 within CBF3CC (Leber et al., 2018).

The most striking secondary structural feature of this model is an eight-stranded parallel β sheet. This is part of a larger solenoid structure, corresponding to the predicted leucine-rich repeat (LRR) fold of Ctf13, and comprises eight LRR motifs, rather than seven LRRs identified previously (Leber et al., 2018). The LRRs are flanked at each end by additional domains (Figures 2E and S3A). The N-terminal BTB/POZ domain of yeast Skp1 (Orlicky et al., 2003) could be fit as a rigid body into one flanking domain. As previously observed (Leber et al., 2018), the CBF3 Skp1-Fbox structure is atypical, with the Skp1 helices α7 and α8 reoriented by 86° and 60°, respectively (Figures S3B, S3C, S5A, and S5B). Our model also reveals a unique linker subdomain between the F box and the LRRs, comprising a three-stranded anti-parallel β sheet and a long α-helix (Figure S3A). The majority of the F box is well conserved between Saccharomycetaceae, so this atypical structure is likely to be found in all Skp1-Ctf13 homologs (Figure S4A). In addition, there is an ~50 amino acid insertion between α2 and α3 of the F box for which there is no electron density, despite the latter half having significant sequence conservation (see further results below).

The remaining flanking density is an α-β subdomain that decorates the C-terminal end of the Ctf13 LRR, formed by insertions within the last 3 LRRs of Ctf13. This subdomain contacts and stabilizes a typically disordered acidic loop of Skp1 such that the Skp1-Ctf13 heterodimer forms a toroidal structure in which both the N-terminal F box and the C-terminal α-β LRR-insertion domain of Ctf13 associate with either end of Skp1 (Figures 2E and S5C). Other LRR-containing F box proteins form toroids, notably both the TIR1 and COI1 plant hormone receptors (Sheard et al., 2010; Tan et al., 2007), but in these cases the toroid is formed by the LRR domain alone, with Skp1 and the LRR domains almost perpendicular to each other. In contrast, for Skp1-Ctf13, the unusual Skp1-F box interaction forces both Skp1 and Ctf13 into the same plane (Figures 2E and S3D).

The Ctf13-Skp1 heterodimer forms the left side of the horseshoe and makes extensive contacts with the base, formed by the “proximal” monomer of the Cep3 homodimer (Figures 2E and S5D–S5F). This interface includes density for part of a loop in Cep3, from residues 330–339, not previously visible in the crystal structures. The “distal” monomer of the Cep3 dimer forms the remaining side of the horseshoe (Figure 2E). This arrangement positions the truncated Zn2Cys6 domains of (Cep3-DN)2 at opposite ends of the channel created by the horseshoe. The position of Ctf13 relative to (Cep3-DN)2 results in the channel being considerably narrower next to the N terminus of the distal Cep3-DN monomer (Figure 2E).

CBF3CCΔN Channel Is the Putative Binding Site for CDEIII
Ctf13 and Cep3 line the channel with basic residues that are strongly conserved between Saccharomycetaceae (Figures 3A–3C). In Ctf13, a series of arginine and lysine residues extend like fingers from the inter-LRR turns of LRRs 1–6 into this channel
Figure 3. The CBF3CCΔN Channel Is the Putative Binding Site for CDEIII

Views of each side of the inner surface of the channel colored by protein (color as in D) (A), electrostatic potential (from −5 [red] to +5 [blue] TeV) (B), and conservation (purple, white, and green denote high, medium, and low conservation, respectively) (C). (D) View down the two-fold axis of the Cep3ΔN homodimer. R308Ctf13 lies directly along the two-fold axis (orange ball and stick). Conserved basic residues in Cep3 are highlighted in magenta ball and stick.

The charge and conservation within the channel, and the striking relative orientation of the arginine residues of neighboring LRRs in Ctf13, suggested that the channel may provide the binding site for CEN DNA, with these residues potentially contributing direct interactions with CDEIII. In order to test this model, we carried out electrophoretic mobility shift assays. Full-length Cep3 alone binds CDEIII tightly, whereas Cep3ΔN, with truncated Zn2Cys6 domains, does not bind CDEIII DNA, consistent with previous observations (Figure 4A, lanes 5–7 and 10) (Purvis and Singleton, 2008). In contrast, CBF3CCΔN shows a DNA gel shift (Figure 4A, lanes 2–4 and 5–7), indicating that association with the Skp1-Ctf13 heterodimer significantly enhances the affinity of Cep3ΔN for CDEIII DNA. This observed shift does not require dephosphorylation of Skp1 (Leber et al., 2018). Labeled CDEIII DNA could be competed with either unlabeled CDEIII DNA or a random DNA duplex of equal length (Figure 4A, lanes 8 and 9) indicating that the association is not sequence specific. Our in vitro data therefore provide conclusive evidence that the Zn2Cys6 domains are not required for the association of the CBF3CC complex with DNA. However, the Zn2Cys6 domains are the sole determinants of sequence specificity, as the remaining component of the CBF3, Ndc10, has also been shown to contribute affinity not but specificity to the CDEIII association (Cho and Harrison, 2011).

We tested the contribution of the highly conserved arginines 307, 308, and 330 from Ctf13 and lysines 265 and 364 from Cep3 to this interaction. The triple mutation to alanine in Ctf13 reduces association of CBF3CCΔN with DNA (Figure 4B). This effect is amplified by further mutation of Lys265 and Lys364 to alanine in Cep3 (Figure S6A). However, we note that although the mutant complex purifies in a manner consistent with a fully folded complex, it is destabilized relative to the wild-type (WT) and Ctf13-triple-alanine mutant (Figure S6B). Considered together, these data are consistent with a model in which basic residues lining the channel contribute affinity to DNA binding.

Ndc10 Associates with the Atypical F Box of Ctf13

Our map of the complex between CBF3CCΔN and Ndc10 D1-2 has excellent density for the core complex and lower resolution for Ndc10, suggesting flexibility between the two. Focused refinement of Ndc10 D1-2 yielded a second map with improved density for Ndc10 (Figures S2E and S2F; Table S1). Superposition of the common parts of these two maps yielded a composite map, which combined the high-resolution core from the original map and the improved Ndc10 density from the second map. The atomic model of CBF3CCΔN fits readily into our composite map, and the remaining density accommodates the crystal structure of S. cerevisiae Ndc10 D1-2 (Perriches and Singleton, 2012) (Figure 5A). Closer inspection reveals that the majority of the interface is formed from structural elements that are disordered in the structures of the isolated components. Ndc10 binds to the F box of Ctf13 on the outside of the channel, and two unique structural elements of this atypical F box are responsible for the interaction. Specifically, within Ctf13, the long insertion between a2 and a3 (approximately amino acids 30–90), which is predicted to be helical and contains stretches of conservation particularly through residues 70–90 (Figure S4), and a turn between β2 and β3 of the antiparallel β sheet, locate to density at the interface with Ndc10 D1-2. The interface on Ndc10 is mediated primarily by residues at its N terminus (before Gln 44) which are also predicted to be helical, with minor contributions from the C-terminal end of a3 (Val138) and a disordered loop between a1 and a2 (Pro65–Ser72) (Figures 5B and 5C). The C-terminal residues of Skp1 (Glu189–Arg194) are also located at the interface. Together these structural elements
form a hinge between the core complex and Ndc10 leading to lower resolution for this region of the cryo-EM map.

The resulting complex associates tightly with DNA relative to the binding by CBF3CCΔN alone (Figure S7B).

**DISCUSSION**

**Structural Homology of Ctf13 Suggests an Evolutionary Link to Epigenomic Modifications of the Point CEN**

The consistently high resolution of our map for CBF3CCΔN allowed an atomic model of Ctf13 to be built. A search of PDBFold and Dali identifies the histone demethylase KDM2B as the closest known F box-containing structural homolog of Ctf13 (Figures S3F–S3G).

KDM2B is a human lysine demethylase from the JHDM1 family containing the Jumonji (JmjC) domain. In general, the members of this family are responsible for inducing transcriptional silencing through demethylation of H3K36. In lower eukaryotes, family members, such as Jhd1 in *S. cerevisiae*, comprise only the histone lysine demethylase domain, while in higher eukaryotes they have several additional domains that detect or alter epigenomic states (Klose et al., 2006). In humans, these include a zinc finger domain that recognizes methylated DNA and an F box and LRR domain that can recruit SCF E3 ubiquitin ligase activity (Han et al., 2016; Wong et al., 2016).

CENs have been subject to rapid recent evolutionary change, accounting for the wide diversity in CEN sequences among species. Evidence suggests that the point CENs of Saccharomyces cerevisiae evolved from an ancestor with a regional CEN, and the CBF3 complex co-evolved to meet the requirements of genetic specification (Malik and Henikoff, 2009). The structural homology between Ctf13 and KDM2B may indicate a molecular link between Ctf13 and epigenomic specification of point CENs. Such a link has already been shown to exist for Ndc10, as it associates with the histone chaperone Scm3/HJURP, thereby recruiting the centromeric histone Cse4 (CenH3) to the point CEN. Ctf13 may therefore also contribute a currently uncharacterized function in epigenomic specification of the budding yeast point CEN.

**Model for CBF3 Association with CDEIII**

Our electrophoretic mobility shift assay (EMSA) results support a model in which CDEIII DNA binds in the CBF3CCΔN channel, and consistent with this the diameter of the channel accommodates a modeled DNA duplex (Figure S8A). Superposition of the crystal structure of DNA-bound Ndc10 D1-2 from the budding yeast *Kluyveromyces lactis* (Cho and Harrison, 2011) onto our model indicates that the DNA-binding site of Ndc10 is coplanar with, and approximately perpendicular to, the CBF3CCΔN channel (Figure 6A). Therefore, to satisfy both the core and Ndc10 DNA-binding sites, the DNA must bend significantly.

Previous crosslinking data identified crosslinks between Ctf13 and the completely conserved cytosine at the pseudo-dyad axis, and its neighboring 3′ thymine, on the bottom strand of CEN3 CDEIII (Espelin et al., 1997) (Figure 6B). Alignment of the pseudo-dyad axis of modeled CEN3 CDEIII DNA with the two-fold axis of Cep3 places the central cytosine in line with the conserved Arg308 of Ctf13. If the DNA is then oriented such...
that the most conserved surface of Ctf13 aligns with the conserved “right-hand” end of CDEIII, the CCG half-site is placed at the narrow point of the channel (Figure S8A). There is strong sequence conservation between the Zn$_2$Cys$_6$ domains of Cep3 and fungal transcription factors, including between residues that contribute to half-site recognition (Figure S8D). When the Zn$_2$Cys$_6$ of Cep3 is modeled using the prototypical GAL4 crystal structure (Marmorstein et al., 1992), the narrow gap perfectly accommodates the Zn$_2$Cys$_6$ domain and orients its C-terminal end toward the N-terminal end of Cep3 (Figure S8B). This superposition packs the Zn$_2$Cys$_6$ domain against a strongly conserved loop between LRR motifs 3 and 4 (Figure S4).

The equivalent superposition at the other half-site is sterically incompatible with linear DNA (Figure S8C). However, the CEN DNA of budding yeast, comprising CDEI–CDEIII, forms a centromeric nucleosome (Furuyama and Biggins, 2007) with atypical significant loss of association with CBF3 compared with the TGT half-site (Espelin et al., 1997), and genetic results identify the CCG triplet as the only bases within the CEN whose substitution cannot be supported in S. cerevisiae (Cumberledge and Carbon, 1987; Gaudet and Fitzgerald-Hayes, 1987; Jehn et al., 1991; McGrew et al., 1986; Ng and Carbon, 1987; Niedenthal et al., 1991). Leber et al. (2018) identified a binding site for the TGT-proximal Zn$_2$Cys$_6$ domain between Ctf13 and Cep3 in a manner that would not allow binding to the TGT half-site. Because this location is not compatible with the observation of crosslinks between the TGT half-site and Cep3 (Espelin et al., 1997), it may be possible that this is an inactive conformation and that the presence of DNA and/or dephosphorylation may unlock it from this docked position.

Linear extension of the modeled DNA beyond CDEIII to incorporate bases to the “right” of the CEN DNA places bases approximately 21–30 bp downstream of the CCG half-site at
the Ndc10 DNA-binding site (Figures 6B and 6C). This is in agreement with Ndc10 crosslinking results indicating that Ndc10 forms sequence-specific interactions with a series of thymines to the “right” side of CDEIII (Espelin et al., 1997). As a whole, the complex encompasses a stretch of approximately 50 bp of CDEIII, remarkably consistent with the established 56 bp DNaseI footprint of CBF3 (Lechner and Carbon, 1991).

The centromeric nucleosome of S. cerevisiae protects ∼121 bp DNA (Cole et al., 2011; Krassovsky et al., 2012), the majority of which is AT-rich CDEII. Independent experimental observations suggest that the CEN DNA is wound in a right-handed manner and that this is dependent on the sequences within CDEII and CDEIII (Díaz-Ingelmo et al., 2015; Furuyama and Henikoff, 2009), contrary to the left-handed DNA supercoiling in canonical nucleosomes. Similar observations in higher eukaryotes suggest that this is not unique to organisms with point CENs (Furuyama and Henikoff, 2009).

Our Ndc10 construct is monomeric, so the position of the second Ndc10 protomer is unknown. Nonetheless, the C terminus of our construct is oriented toward the CDEII-proximal side of the modeled DNA (Figure 6C), suggesting that second Ndc10 protomer may also orient toward CDEII. Such a location is consistent with gel shift data indicating that Ndc10 associates specifically with CDEII (Espelin et al., 2003) and biophysical and biochemical experiments showing the Ndc10 dimer binds two separate segments of DNA (Cho and Harrison, 2011). In this scenario, the curvature of CDEIII induced by the core complex and the CDEIII-bound Ndc10 protomer is further enhanced at CDEII by the second Ndc10 protomer, in agreement with topology analysis of plasmids containing CEN DNA in budding yeast (Díaz-Ingelmo et al., 2015) and thus accounting for the observed looping of CEN DNA (Pieprasanta et al., 1999).

In our model, both a right-handed and a left-handed DNA supercoil can be accommodated (Figure 6C). A right-handed superhelical path would position CDEI close to Ndc10 D1, consistent with data indicating that the CDEI-binding protein Cbf1 associates directly with Ndc10 D1 (Cho and Harrison, 2011). Further structural data that definitively locate the second Ndc10 protomer will be required to resolve this ambiguity.

Our structures of the CBF3CCN alone and in complex with Ndc10 D1-2 and the resulting model of CDEIII binding explain a wealth of data published over several decades, including genetic analyses (Cumberledge and Carbon, 1987; Gaudet and...
Fitzgerald-Hayes, 1987; Hegemann et al., 1988; Jehn et al., 1991; McGrew et al., 1986; Ng and Carbon, 1987; Niedenthal et al., 1991), crosslinking (Espelin et al., 1997, 2003), and footprinting data for CBF3 (Lechner and Carbon, 1991). Our structures indicate the path of DNA is strongly curved at CDEIII and probably looped at CDEII by the second protomer of Ndc10, consistent with atomic force microscopy (AFM) experiments that show shortening and bending of CEN DNA when bound to CBF3 (Pietrasanta et al., 1999).

**EXPERIMENTAL PROCEDURES**

**Protein Expression and Purification**

Our experiments showed that the co-expression of Sgt1 helps the formation and stability of CBF3CCN; therefore we cloned and co-expressed Sgt1 with CBF3CCN components. The Sgt1 gene with a C-terminal Strep tag was cloned into a modified pRS424 vector containing the CTF13 gene with a C-terminal CBP tag. The SKP1 gene was cloned into the modified pRS426 vector containing CEP3 (residues 47–608) with an N-terminal His tag. Mutants were generated by site-directed mutagenesis PCR. Both plasmids were co-transformed into S. cerevisiae yeast strain BCY123 (MATα pep4::HIS3 prb1::LEU2 bar1::HIS6 lys2::GAL1/10GAL4 can1 ade2 trp1 his3 leu23,112) by cotransformation with cOmplete EDTA-free Protease Inhibitor Cocktail (Roche) and frozen as pellets in liquid N2. The pellets were lyzed using a freezer miler (SPEX Sample Prep). The complex was purified by celfulminulin resin and eluted by buffer B (10 mM Tris [pH 8.0], 500 mM NaCl, 1 mM Mg acetate, 2 mM imidazole, 4 mM CaCl2, and 0.2% 2-mercaptoethanol) supplemented with cComplete EDTA-free Protease Inhibitor Cocktail (Roche) and frozen as pellets in liquid N2. The pellets were lyzed using a freezer miler (SPEX Sample Prep). The complex was purified by celfulminulin resin and eluted by buffer B (10 mM Tris [pH 8.0], 500 mM NaCl, 1 mM Mg acetate, 2 mM imidazole, 4 mM CaCl2, and 0.2% 2-mercaptoethanol). The eluted protein was loaded onto a Mono Q column (GE Healthcare) pre-equilibrated with buffer C (20 mM Tris, 500 mM NaCl, 20 mM imidazole [pH 8.0], and 10 mM β-mercaptoethanol). The complex fractions were pooled and concentrated before loading onto a Superdex 200/25/150 GL (GE Healthcare) pre-equilibrated with S200 buffer (15 mM Tris [pH 8.0], 200 mM NaCl, and 2 mM DTT). The peak fraction was used to make EM grids.

Cep3FL with an N-terminal His-tag was expressed in S. cerevisiae using the same system as above and was purified using Ni-affinity chromatography, using the His-Trap, ion exchange, and size exclusion steps described above.

Domains 1 and 2 (residues 1–554) of NDC10 were cloned from yeast genomic DNA into the pRS426 vector with a C-terminal Strep tag. The (Ndc10 D1-2) pRS426 was co-transformed with an empty pRS424 plasmid to express the Ndc10 D1-2 protein. The protein was purified by Streptrap HP column (GE Healthcare) pre-equilibrated with Strept binding buffer (20 mM Tris-HCl [pH 8.0], 500 mM NaCl, 1 mM EDTA, and 3 mM DTT) and then eluted in Strept elution buffer (binding buffer + 3 mM d-desthiobiotin). Fractions containing the desired protein were pooled and purified further by Mono Q column and Superdex 200/25/150 GL. Purified CBF3CCN and Ndc10 D1-2 were mixed at 3:1 molar ratio and incubated at 4°C for 1 hr. The mixture was loaded onto Superose 6 Increase 3.2/300 (GE Healthcare). The peak fraction was used to make EM grids.

**Cryo-EM**

The samples (0.12 and 0.24 mg/ml for the first and second datasets of CBF3CCN + 0.15 mg/ml for CBF3CCN + Ndc10 D1-2) were applied to glow-discharged UltrAuFoil 1.2/1.3 300 mesh grids (Quantifoil Micro Tools). Cryo-EM data were acquired at the Electron Bio-Imaging Centre (eBIC) on a FEI Titan Krios at 300 keV (Thermo Fisher Scientific), equipped with a post Quantum energy filter K2 Summit direct detector (Gatan). Data collection was automatically carried out using EPU software (Thermo Fisher Scientific) at a magnification of 47,170 (1.06 Å2 pixel−1) to record 1,236 videos with a defocus range of −1.6 to −3.6 μm for the first CBF3CCN dataset and 1,101 videos with a defocus range of −1.0 to −4.0 μm for the second CBF3CCN dataset. The total exposure time of 10 s fractionated into 25 frames with a total dose of 46 e− Å−2 for first dataset and the total exposure time of 15 s fractionated into 40 frames with a total dose of 60.9 e− Å−2 for second dataset. Videos were aligned using MotionCor2 (Cheng et al., 2017).

An initial dataset of the complex of CBF3CCN + Ndc10 D1-2 was collected in-house on a Tecnai Polara operating at 300 keV, recording at 1.39 Å2 pixel−1 on a post Quantum energy filter K2 Summit direct detector operating in counting mode. The total dose was 54 e− Å−2 for 18 s and 60 frames with a defocus range of −1.5 to −4.0 μm. The final dataset for this complex, comprising 2,003 movies, was collected at eBIC on a Titan Krios microscope using the same parameters as CBF3CCN, except the total dose was 50 e− Å−2 and the defocus range was −1.5 to −3.5 μm.

**Image Processing**

Contrast transfer function (CTF) parameters were estimated using CTFFIND4 (Rohou and Grigorieff, 2015), and CTF correction and following image processing were performed using RELION 2.0 (Kimmann et al., 2016), unless otherwise noted. Resolution is reported using the gold-standard FSC (0.143 criterion) as described (Rosenthal and Henderson, 2003; Scheres and Chen, 2012), and temperature factors were determined and applied automatically in RELION 2.0. For CBF3CCN, a subset of the initial dataset was picked using an automatically generated Gaussian reference by Gautomatch (Urvanauitis et al., 2015), extracted using a 2002 pixel box, and then subjected to reference-free 2D classification. Some of resulting 2D class averages representing different views were selected to be low-pass-filtered to 25 Å and used as references for further automatic particle picking of the initial dataset. The automatically picked particles were screened manually, followed by reference-free 2D classification, which yielded 69,392 particles for subsequent processing. An ovoid generated with SPIDER (Shaih et al., 2008) was used as an initial model for 3D classification. The best 3D class was used to perform a 3D auto-refinement against all the good particles, resulting in a 4.9 Å map. After substitution of the particles contributing to this map by re-extraction from dose-weighted images calculated by MotionCor2, a further 3D auto-refinement provided a reconstruction at 4.7 Å overall resolution. A total of 212,724 particles from the second dataset were picked from dose-weighted images and selected for further processing after reference-free 2D classification. After the two datasets, 282,116 particles were input to 3D classification using an initial 3D reference obtained by low-pass filtering (50 Å) the 4.7 Å map. After 3D classification, the best class, with 187,606 particles, was subjected to 3D auto-refinement. This resulted in a reconstruction of 4.1 Å. The map was post-processed by RELION and sharpened by a negative B factor using an automated procedure, resulting in a 3.7 Å reconstruction.

For CBF3CCN + Ndc10 D1-2, particles were picked automatically from the initial Polara dataset by Gautomatch (Urvanauitis et al., 2015), extracted using a 1522 pixel box, and subjected to reference-free 2D classification, yielding 9,024 particles for subsequent processing. The particles were subjected to 3D classification using an initial 3D reference obtained by low-pass filtering (50 Å) the 4.7 Å map. After 3D classification, the best class, with 187,606 particles, was subjected to 3D auto-refinement. This resulted in a reconstruction of 4.1 Å. The map was post-processed by RELION and sharpened by a negative B factor using an automated procedure, resulting in a 4.0 Å reconstruction.

For CBF3CCN + Ndc10 D1-2, particles were picked automatically from the initial Polara dataset by Gautomatch (Urvanauitis et al., 2015), extracted using a 1522 pixel box, and subjected to reference-free 2D classification, yielding 9,024 particles for subsequent processing. The particles were subjected to 3D classification using an initial 3D reference obtained by low-pass filtering (50 Å) the 4.7 Å map. After 3D classification, the best class, with 187,606 particles, was subjected to 3D auto-refinement. This resulted in a reconstruction of 4.1 Å. The map was post-processed by RELION and sharpened by a negative B factor using an automated procedure, resulting in a 4.0 Å reconstruction.
Model Building

Cep3 (Purvis and Singleton, 2008) and the BTB/POZ domain of Skp1 (Orlicky et al., 2003) were placed in the CBF3CCN map using Chimera (Pettersen et al., 2004). Ctf13 was built de novo using Coot (Emsley et al., 2010) in an early 4 Å map, using secondary structure predictions from both Psipred (Jones, 1999) and Phyre2 (Kelley et al., 2015), sequence conservation between common ascomycetes, and with reference to structural preferences for LR3 domains (Bella et al., 2008). Tracking of the main chain was assisted using a 6 Å-filtered map. Broken density frequently indicated flexible loops, some of which could be built. Sequence alignments were generated using ClustalW (Larkin et al., 2007) and annotated using ESPript (Gouet et al., 2003). A single round of Real Space Refinement in Phenix was used to refine geometry (Afonine et al., 2010). This model was then rigid body-fit into the final map at 3.7 Å, followed by a final polish in Coot and several rounds of refinement in Phenix and Refmac. Secondary structure restraints were initially generated using the Cep3 and Skp1 crystal structures and, for Ctf13, from within Phenix, with manual editing where deviations from the crystal structures were evident. In order to validate the model, FSCfree and FSCwork were calculated. Using Phenix, the atomic coordinates of the final model were randomly shifted by 0.5 Å and subsequently real space-refined against one unmasked half map (the “working” map). The resulting model was converted to a map and an FSC was calculated between it and both the working map (FSCwork) and the free half map (FSCfree).

The atomic model of CBF3CCN and the crystal structure of S. cerevisiae Ndc10 (PDB: 4ACO) were placed into the CBF3CCN plus Ndc10 D1-2 map using Chimera. DNA was modeled using 3D-DART (van Dijk and Bonvin, 2009).

Gel EMSAs

Protein-DNA interactions were evaluated using EMSA. Twenty-four picomoles singly or doubly labeled 39 bp CDEII dsDNA (ATTATATGGTATTTTAGTTC GAAAATGTTAA) was mixed with different amounts of protein with the indicated ratio of DNA to protein in the reaction buffer (25 mM HEPES [pH 8.0], 200 mM KCl, 2 mM DTT, 10% glycerol, 0.02% NP-40, 10 mM MgCl2 and 10 μM ZnCl2). For competition EMSAs, the unlabeled competitor DNA was 50 times more concentrated than the labeled DNA. The mixtures were incubated at room temperature for 45 min and resolved on a 3%–12% Bis-Tris native polyacrylamide gel at a constant voltage of 150 V at 4°C in 1× native PAGE running buffer for 110 min. After electrophoresis, the gel was scanned using an FLA-3000 fluorescent image analyzer (Fujifilm) excited with a 473 nm laser. EMSAs are representative of at least two repeats, with the exception of those in the Supplemental Information, which were carried out once.

Protein Stability Assay

The stability of the WT and mutant CBF3CCN was assessed using a fluorescence thermal shift assay. The melting temperature of 5 μM protein was measured in a 25 μL volume containing SYPRO orange (Molecular Probes) in a 96-well format using a real-time PCR detection system (Mylon et al., 2014) after buffer exchange into 15 mM HEPES (pH 8.0), 200 mM NaCl, and 2 mM DTT. The melting temperature was determined by non-linear regression analysis of the raw data in GraphPad Prism.

DATA AND SOFTWARE AVAILABILITY

The accession numbers of the EM maps and atomic coordinates reported in this paper are EMDB: EMD-4241, EMD-0051, EMD-0052 and PDB: 6FE8 and 6GSA.

SUPPLEMENTAL INFORMATION

Supplemental Information includes eight figures and two tables and can be found with this article online at https://doi.org/10.1016/j.celrep.2018.06.068.

ACKNOWLEDGMENTS

We are grateful to Martin Singleton for the Ctf13 and Cep3 genes; Kiyoshi Nagai for sharing his yeast expression system and helping us test expression of our constructs; Dan Clare, Corey Hecksel, and Alistair Siebert for data collection at eBIC; and Takanori Nakane for suggestions during data processing. W.Z. and C.K.V. were funded by the Biotechnology and Biological Sciences Research Council (BBSRC) (BB/J007595/1).

AUTHOR CONTRIBUTIONS


DECLARATION OF INTERESTS

The authors declare no competing interests.

Received: December 20, 2017
Revised: May 6, 2018
Accepted: June 15, 2018
Published: July 17, 2018

REFERENCES


