Structural and dynamic mechanisms of CBF3-guided centromeric nucleosome formation

Ruifang Guan¹,6, Tengfei Lian²,6, Bing-Rui Zhou¹, Emily He³, Carl Wu⁴, Martin Singleton⁵ & Yawen Bai¹

Accurate chromosome segregation relies on the specific centromeric nucleosome–kinetochore interface. In budding yeast, the centromere CBF3 complex guides the deposition of CENP-A, an H3 variant, to form the centromeric nucleosome in a DNA sequence-dependent manner. Here, we determine the structures of the centromeric nucleosome containing the native CEN3 DNA and the CBF3 core bound to the canonical nucleosome containing an engineered CEN3 DNA. The centromeric nucleosome core structure contains 115 base pair DNA including a CCG motif. The CBF3 core specifically recognizes the nucleosomal CCG motif through the Gal4 domain while allosterically altering the DNA conformation. Cryo-EM, modeling, and mutational studies reveal that the CBF3 core forms dynamic interactions with core histones H2B and CENP-A in the CEN3 nucleosome. Our results provide insights into the structure of the budding yeast centromeric nucleosome and the mechanism of its assembly, which have implications for analogous processes of human centromeric nucleosome formation.
Centromeres mediate the attachment of chromosomes to the mitotic spindle by the kinetochore complex that binds microtubules, which is responsible for accurate chromosome segregation during mitosis. Mis-segregation of chromosomes can lead to aneuploidy, a hallmark of cancer. Centromeres are marked by specific nucleosomes in which the canonical histone H3 is replaced by the CENP-A variant. Human centromeres are regional, including megabase DNA with repeats of the same DNA fragment. Studies using native chromatin immunoprecipitation followed by sequencing of CENP-A-containing particles reveal that the octamer is the major form of CENP-A nucleosomes in normal centromeres and naturally occurring neocentromeres. Structures of human CENP-A nucleosome core particle containing one half human palindromic DNA, Widom 601 DNA, or a native α-satellite DNA have been solved at near-atomic/atomic resolution using X-ray crystallography and cryo-EM.

The CEN3 nucleosome is different from those in the earlier two structural models. In the structure of the CEN3 CENP-ACse4 nucleosome, the CNT motif, located between the super-helical locations 3 and 4 of the DNA, is accessible for binding by the Gal4 domain (Fig. 1c, d and Supplementary Fig. 1a–f, and Table 1). The high quality of the density map allowed us to build a structural model of the nucleosome with a uniquely positioned DNA (Fig. 1c, d and Supplementary Fig. 1g–j).

In contrast, budding yeast chromosomes have point centromeres with single nucleosomes that are defined by conserved ~125 nucleotide segments, including three centromere determining elements (CDEs): CDEI (8 bp), CDEII (~80–90 bp, and AT-rich), and CDEIII (~25 bp). The CDE-IIEM structure of budding yeast CENP-A (Cse4 or CENP-A Cse4) nucleosome containing the Widom 601 nucleosome positioning DNA was determined at 2.7 Å resolution. However, the structure of the nucleosome containing the native α-satellite DNA is determined at 2.6 Å resolution using a single-chain antibody fragment (scFv)-assisted cryo-EM method in which each nucleotide can be resolved.

In the structure, we found that 115 bp DNA, including CDEII (83 bp), CDEIII (26 bp), and 6 bp of CDEIII (the region on the right side of CDEIII) (Fig. 1a), interacted with the core histones and formed a well-defined left-handed ~1.3-turn super-helical structure. The end regions of the CEN3 DNA (2–14 and 131–137) showed much weaker density, indicating that they have flexible conformations. In comparison, the recently reported structure of the CENP-A Cse4 nucleosome core containing the non-native Widom 601 DNA includes 119 bp structured DNA, even though the core histones in both of the CEN3 and 601 CENP-A Cse4 nucleosomes show similar structures with a root mean square deviation (RMSD) of 0.9 Å (Supplementary Fig. 1k). In contrast, the human CENP-A nucleosome core particle containing the native α-satellite DNA includes 145 bp. Amino acid sequence alignment of the human CENP-A and budding yeast CENP-A Cse4 shows that the lack of several positively charged residues in the α helix of CENP-A Cse4 is the likely cause for the flexible DNA ends in the CEN3 CENP-A Cse4 nucleosome (Fig. 1e). Also, the DNA position in our cryo-EM structure of the CEN3 nucleosome is different from those in the earlier two structural models.

In the structure of the CEN3 CENP-A Cse4 nucleosome, the CCG motif, located between the super-helical locations 3 and 4 of the DNA, is accessible for binding by the Gal4 domain (Fig. 1c), suggesting that the CBF3 core could bind to the CEN3 nucleosome. To test it, we conducted the nucleosome binding study using the CBF3 core purified from budding yeast cells after treating it with the phosphatase (bacteria lambda protein). Previous studies have shown that the purified CBF3 core is phosphorylated, and dephosphorylation is required for CBF3 core binding to CEN3 DNA. Indeed, we found that the CBF3 core but not the purified phosphorylated form could bind to the CEN3 CENP-A Cse4 nucleosome in an electrophoretic mobility shift.
We tried to use the cryo-EM method to determine the structure of the CENP-ACse4 CEN3 DNA nucleosome in complex with the CBF3core. However, the CEN3 DNA again dissociated from the nucleosome bound to CBF3core on the cryo-EM grid, and the antibody fragment could not bind to the CBF3core–nucleosome complex (Supplementary Fig. 2b).

To stabilize the nucleosome, we engineered a hybrid DNA, CEN3-601, by using the CDEIII and its neighboring regions to substitute the corresponding region in the 601 DNA (Fig.2a and Supplementary Fig. 2c). We found only the CBF3core was associated with the nucleosome when CBF3 was mixed with the nucleosome; Ndc10 dissociated from the complex (Fig. 2b).

CBF3core showed a similar affinity to the CENP-ACse4 nucleosome with either the native CEN3 or the CEN3-601 DNA (Supplementary Fig. 2d). CBF3 and CBF3core bound to the CEN3 nucleosome with similar affinity (Supplementary Fig. 2f). Also, only dephosphorylated CBF3core or the CBF3core with the L1 loop deletion mutant of Skp1 could bind to the CEN3-601 nucleosome (Supplementary Fig. 2g, h). We also found that CBF3core bound to the CENP-ACse4 nucleosome only slightly better than the H3 nucleosome and deletion of the L1 loop in CENP-ACse4, a major difference between H3 and CENP-ACse4 on the surface of the nucleosomes, showed little effect on binding affinity (Fig. 2c).

However, a single base pair shift in the incorporated DNA position led to a weaker binding of CBF3core to the nucleosome (Fig. 2d). We were able to observe intact particles of the CBF3core bound to the CEN3-601 H3 nucleosome and obtained the cryo-EM density map at 4.2 Å resolution. It allowed us to build the structural models that showed multiple conformations for the major core region of the CBF3core (Fig. 2e, f, Supplementary Fig. 3 and Table 1).

In the CBF3core–nucleosome complex, CBF3core bound to the nucleosome through specific recognition of the CCG motif by the Gal4 domain in one of the two Cep3 subunits (Figs. 2f and 3a). The CDEIII DNA region associated with H2B in the free CEN3 CENP-ACse4 nucleosome was detached, and the full CDEIIIR DNA was in the naked form (Fig. 3a). The CDEIIIR site that interacts with the L1 loop and the α1 helix of core histone H2B in the free nucleosome is ~15 bp away from the GCC motif site. This structural feature suggests that the Gal4 domain binding to the CCG motif could have an allosteric inhibitory effect on the binding of the DNA by H2B (Supplementary Fig. 4a). Notably, recent studies have also shown that the binding of the pioneer transcription factor Sox2/Sox11 HMG domain to the nucleosome can also lead to the detachment of DNA from the core histones through allosteric effects. In addition, in our case, it appears that the Gal4 domain binding also makes the CDEIIIR DNA region more rigid as observed by cryo-EM.

The major core region of the CBF3core showing multiple conformations is connected to the Gal4 domain through a flexible linker and moves as a rigid body (relative to the Gal4-nucleosome...
Examination of the available structures of the CBF3core in the free and DNA-bound forms shows that the major core regions have the same structure but display different orientations relative to the Gal4 domain (Supplementary Fig. 4b, c).25,32,33 In some of the conformations, the Ctf3 subunit is close to the α2 helix of H2B and the L1 loop of H3 (Fig. 2e and Supplementary Fig. 4d). It explains why scFv could not bind to the CBF3core–nucleosome complex (Supplementary Figs. 2a and 4e).

To examine the interactions between the CBF3core and the CEN3 CENP-A Cse4 nucleosome, we built a structural model by substituting the H3 histone in the CEN3-601 nucleosome–CBF3core complex with CENP-A Cse4. We found that the Cbf13 subunit was close to the five charged residues of the α2 helix of H2B and the three residues in loop 1 of CENP-A Cse4 (Fig. 3b). The model suggests potential formation of a salt bridge between Ctf13 K282 and H2B E117 and hydrophobic interactions between Cft13 P297 and CENP-A Cse4 K1987 (Supplemental Fig. 4d). We mutated the residues in the α2 helix of H2B to Ala, deleted the three residues in loop 1 of CENP-A Cse4, and measured the binding affinities of CBF3core to the nucleosome–Gal4 region.

**Fig. 2** The cryo-EM structure of CBF3core bound to the CEN3-601 nucleosome. a Illustration of engineering of the CEN3-601 DNA sequence and single bp shifts of the incorporated DNA. The alignment of the CEN3 and 601 DNA sequences are based on the structural alignment of the CEN3 CENP-A Cse4 and human 601 CENP-A nucleosomes (PDB: 6BUZ). b EMSA assay of CBF3 and CBF3core binding to the CEN3-601 DNA nucleosome, showing that Ndc10 was competed out by the nucleosome and only CBF3core bound to the nucleosome. Three experiments were repeated independently with similar results. c EMSA assay of CBF3core binding to the CEN3-601 DNA nucleosome containing CENP-A Cse4, CENP-A Cse4 with deletion of the L1 loop (CENP-A Cse4 ΔL1), and H3 histones. Two experiments were repeated independently with similar results. d EMSA assay of the effects of single bp shift of the CCG location in the CEN3-601 DNA on the binding affinity between CBF3core and the nucleosomes (top). The quantified intensity ratio of the CBF3core–nucleosome complex (first band above the nucleosome) over the total nucleosome (bottom). Data were presented as mean values. Error bars represent standard deviation values from three (n = 3) independently performed experiments. e Density maps of the CBF3core–nucleosome complex showing two extreme conformations in which the main core regions of CBF3core show different conformational movement relative the nucleosome–Gal4 region. f Different views of the CBF3core–nucleosome structure with CBF3core in one conformation. Middle top shows the cartoon of the CBF3core–nucleosome complex.
nucleosomes containing the histone mutants (Fig. 3c and Supplementary Fig. 4f). The binding dissociation constant ($K_d$) for the nucleosome with wild type histones is $\sim0.32 \mu M$. Each of the two mutants showed little effects on the binding affinity, and when combined, they decreased the binding affinity by only less than a factor of two (Fig. 3c). These results showed that CBF3core only made weak and dynamic contacts with the core histones.

Discussion

In this study, we determined the structure of the CEN3 CENP-A$^{Cse4}$ nucleosome, which shows that the CEN3 DNA mainly uses CEDII and CDEIII to interact with the core histones. In the structure, the CDEII is in the linker DNA region, allowing for binding by CBF1 in a specific sequence-dependent manner. Our structure is consistent with the in vivo nucleosome mapping results, which show that no significant region beyond CEN3 CDEs are resistant to MNase digestion42,43. In contrast, in the recently proposed "proximity" model for CENP-A$^{Cse4}$ deposition by CBF3, the CDEIII is located in the linker DNA region45. Also, the dyad position in the CEN3 CENP-A$^{Cse4}$ nucleosome structural model that is proposed based on hydroxyl radical footprinting results and computational modeling is $\sim12$ nucleotide away from the dyad in our structure44, leading to the conclusion that the Mif2 AT-hook domain contacts only one side of the nucleosome dyad44. Using the dyad location in our cryo-EM structure (position at 74) (Fig. 1), Mif2 would contact both sides of the nucleosome dyad. These results show our understanding of the determinant of nucleosome positioning is limited and it is still
not possible to predict nucleosome positioning accurately. In particular, it is intriguing that the CEN3 nucleosome is uniquely positioned considering the fact that its sequence is highly AT-rich.

We also find that CBF3 can bind to the CBF3core–nucleosome while causing dissociation of Ndc10 from the CBF3core–nucleosome complex. Alignment of the structures of the CBF3core–nucleosome and (CBF3)$_2$–DNA (PDB ID: 6GYS) complexes through alignment of the CCGs region and Gal4 domains in the two structures. CBF3core binds to DNA and the nucleosome in different orientations and the two Ndc10 subunits (cyan) have steric clash with the nucleosome (magenta) as indicated by circles (black).

The dyad in the proximity model is at nucleotide sequence number 34 (Fig. 1a). CDEI could be available for binding of CBF1 in both the CBF3core–CEN3 nucleosome complex and free CEN3 nucleosome. The CBF3core–CEN3 nucleosome complex can only recruit one copy of CCAN on the CDEI side. The free CEN3 nucleosome can recruit two copies of CCAN.

Based on the above results, we speculate a structural and dynamic mechanism for a two-step process for specifying the budding yeast centromeric nucleosome (Fig. 4c, d)\(^\text{50}\). After DNA replication, CBF3, in association with the Scm3–Cse4–H4, could target centromeric DNA by specifically recognizing the CCG motif in CDEIII\(^\text{25}\). Upon the CENP-ACse4 deposition to the CEN DNA by Scm3, the CENP-ACse4 nucleosome could partially form. Meanwhile, the CBF3core is still attached to the CDEIII DNA. During this process, intermediate steps involving various partially formed nucleosomes might occur\(^\text{29}\) and Scm3–Ndc10 dissociates.
from the CBF3core but could still interact with the linker or nucleosomal DNA in a sequence-independent manner25,36. In our model, Ndc10 would serve as a dynamic adapter between CBF3core and Scm3, instead of playing an architectural role in the previous “looping” and “proximity” models25,36. The dynamic interactions of CBF3core with the core histones could help prevent the core histones from dissociating from the DNA, compensating the weak affinity between the core histones and the AT-rich DNA. The AT-rich sequence in CEN3 could also play a role in enhancing the binding specificity of the CCG motif by the G4 domain. The position of the CEN3 in our cryo-EM structure allows CBF3core to bind to the nucleosome while Ndc10 would clash with the nucleosomal DNA. Notably, in the “proximity” model, Ndc10 can coexist with the nucleosome. However, the nucleosome with the hypothetical dyad at nucleotide position 34 is not at an intrinsically favored position. It might only exist as a transient intermediate. The N-terminal domain of CEN10 (NTD, 1-344) and CBF3core containing Cep3 domain, Skp1 or Skp1-ΔL (37–64 deleted), and Ctf19 were purified as expressed in E. coli (Supplementary Table 2)26, to obtain the dephosphorylated CBF3core, the phosphorylated CBF3core (PMP), and Mcn1 were added into the sample, then followed by lambda protein phosphatase. The mixture was incubated at 30 °C for 30 min. The dephosphorylated CBF3core was purified by size-exclusion chromatography (Superose 6, 10/100, GE Healthcare). The final buffer was 5 mM NaCl and 1 mM DTT. The purified proteins were collected and stored at −80 °C. To obtain the dephosphorylated CBF3 core, purified dephosphorylated CBF3core and Ndc10 NTD were mixed at 1:1 molar ratio and incubated at 4 °C for 2 h. The mixture was loaded onto Superose 6 Increase 10/100 (GE Healthcare). The peak fraction was assessed by SDS–PAGE and Coomassie stain. The dephosphorylated CBF3 was stored at −80 °C.

Preparation of DNA. The 136 bp CEN3 DNA was prepared as described27. The 147 bp CEN3-601 DNA, which contains 48 bp DNA of CBF3 binding site underlined, were prepared by PCR amplification followed by ethanol precipitation and purified using the POROS column.

The forward and reverse template DNA sequence are: ATCGGAATTCCGGGCTGAGGCGGCCTCAGATTGCGTACGCCTC-TAGCGCCCTTTAAGGCGATAGCAGTGTTATCGCGGATGATTTCTTACTATTTC, respectively (Supplementary Table 2). The PCR products were pelleted using 70% ethanol containing 0.3 M NaAc 5.2. The PCR products were resuspended by TE buffer. The sample was loaded to POROS column. Eight milliliters of 2 M Tris was added to the supernatant. The proteins were purified by size exclusion chromatography (Superose 6, 10/100, GE Healthcare) in the column. The column was washed with buffer E (20 mM HEPES, pH 7.3, 1 mM NaCl, and 1 mM DTT). The mixtures were dialyzed against refolding buffer (10 mM Tris-Cl at pH 7.4, 1 mM EDTA, 5 mM β-mercaptoethanol, 200 mM NaCl, 25 mM β-mercaptoethanol) for three times. The supernatant was loaded to Hitrap S column chromatography, and the DNA was eluted by a linear gradient of 0 to 1 M NaCl in buffer E. The UV spectrum was monitored using a Nanodrop 2000 spectrophotometer. All yeast CEN3 nucleosomes were reconstituted following the published protocol28. Brieﬂy, purified histone octamers and 136 bp DNA that contained hybrid histone octamers and 136 bp CEN3 DNA were mixed with the same stoichiometric ratio were loaded to Superdex 200 gel filtration column.

Reconstitution of nucleosomes. Core histone octamers were reconstituted first as described22. Purified recombinant histones in equal stoichiometric ratio were dissolved in unfolding buffer (7 M guanidine–HCl, 20 mM Tris–HCl at pH 7.4, 10 mM NaCl) and dialyzed for 2 h against refolding buffer (20 mM HEPES, pH 7.4, 1 mM EDTA, 5 mM β-mercaptoethanol, 200 mM NaCl, 25 mM β-mercaptoethanol) for one day at 4 °C twice. The mixture was centrifuged at 3500 g to remove any insoluble material. Soluble octamers were purified by size fractionation on a Superdex 200 gel filtration column.

All yeast CEN3 nucleosomes were reconstituted following the published protocol29. Brieﬂy, purified histone octamers and 136 bp CEN3 DNA were mixed with a 1:1.3 ratio of DNAoctamer in high-salt buffer (2 mM NaCl, 10 mM K+/Na+ Phosphate at pH 7.4, 1 mM EDTA, 0.02% NP-40, 5 mM β-mercaptoethanol). The 1 ml mixture in a dialysis bag was placed in 600 ml of the high-salt buffer and dialyzed for 60 min followed by salt gradient dialysis. Four liters of a low-salt buffer (100 mM NaCl, 10 mM K+/Na+ Phosphate at pH 7.4, 1 mM EDTA, 0.02% NP-40, 2 mM β-mercaptoethanol) were gradually pumped into dialysis buffer with a flow rate of 2 ml/min for 30 h. The dialysis bag was then dialyzed against low-salt buffer for 60 min. The dialysis was done at room temperature. The sample was then concentrated with the icelandic method and was used to test any insoluble material. The soluble nucleosomes were stored at 4 °C for less than 1 week.

Methods
Expression and purification of histones. Recombinant histones H3, H4, CENP-A, and CENP-C were expressed using the in vitro transcription and translation system, and purified as described previously30. The histone octamers were subjected to size-exclusion chromatography (Superdex 200, 10/100, GE Healthcare) in the column. The column was washed with buffer E (20 mM HEPES, pH 7.4, 1 mM EDTA, 5 mM β-mercaptoethanol, 200 mM NaCl, 25 mM β-mercaptoethanol) for one day at 4 °C twice. The mixture was centrifuged at 3500 g to remove any insoluble material. Soluble octamers were purified by size fractionation on a Superdex 200 gel filtration column.
All yeast CEN3-601, CEN3-601St- and CEN3-601St- nucleosomes were reconstituted using the same protocol, except that the samples were incubated at 37°C for 1 h. Nucleosomes were further purified by sucrose density chromatography (TSKgel DEAE, TOSOH Bioscience, Japan) to remove free DNA and histones. The purified nucleosomes were dialyzed against TE buffer containing 10 mM Tris 7.4, 1 mM EDTA, and 2 mM DTT twice.

Preparation of the complex of the CEN3 CENP-A CBF3core nucleosome bound to scFv. We found that the immobility of the CEN3 CENP-A CBF3core nucleosome was sensitive to the temperature. The soluble nucleosomes were precipitated by incubation at 4°C for over 12 h. The sample was centrifuged at 12,000 g for 1 min to remove the supernatants. The pellet was resuspended in TE buffer. scFv was mixed with the CEN3 CENP-A CBF3core with a 3:1 ratio of scFv:nucleosome at room temperature for 1 h. Then the samples were concentrated for electron microscopy analyses.

Preparation of CBF3core complex with the CEN3-601 nucleosome. The CBF3core and CEN3-601 nucleosome complex was reconstituted by mixing purified dephosphorylated CBF3core with the CEN3-601 nucleosome at 4°C for 1 h. The mixed sample was dialyzed for overnight in a buffer of 10 mM HEPEs at pH 7.3, 50 mM NaCl, 1 mM EDTA, and 1 mM DTT at 4°C. The complex was purified and stabilized using the Graphf method as described. Briefly, the top solution contained 10 mM HEPEs at pH 7.3, 50 mM NaCl, and 10% glycerol (Sigma). The bottom solution contained 10 mM HEPEs a pH 7.3, 50 mM NaCl, 0.15% glutaraldehyde (Polysciences), and 30% glycerol (Sigma). After ultracentrifugation at 190,000 g for 15 h, the fraction was collected and dialyzed against the buffer containing 10 mM Tris·HCl at pH 7.4, 50 mM NaCl, and 2 mM DTT. Then the samples were concentrated for electron microscopy analyses.

Cryo-EM sample preparation and data collection. Three milliliters of nucleosome–scFv sample was loaded onto a glow-discharged holey carbon grid (Quantifoil 300 mesh Cu R1.2/1.3), and 3 ml of nucleosome–CBF3core sample was loaded onto a glow-discharged Lacey grids. The grids were blotted for 3 s at 20°C and images with dose weighting were used for particle picking and extraction. Particles were picked by Gautomatch (https://www.mrc-lmb.cam.ac.uk/kzhang/Gautomatch/) and images with dose weighting were used for defocus determination using CTFFIND4.86

Image processing. The data processing procedures were shown in supplementary Fig. 1 for nucleosome–scFv dataset and supplementary Fig. 3 for nucleosome–CBF3core dataset. The nucleosome–scFv dataset was processed using RELION/3.0-beta2 and the nucleosome–CBF3core dataset was processed using RELION/3.0-beta2 following the standard procedures in RELION3.96 The beam-induced image drift was corrected using MotionCor2.97 The averaged images without dose weighting were used for defocus determination using CTFIND4 and images with dose weighting were used for particle picking and extraction. Particles were picked by Gautomatch (https://www.mrc-lmb.cam.ac.uk/kzhang/Gautomatch/) using templates generated from datasets collected on a 200-kV microscope Tecnai F20.

For the scFv–nucleosome dataset, 342,749 particles were picked. Bad particles were removed by 2D classification. Then 192,020 particles were selected from 3D classification with two classes with good structural features. After re-centering, the best 143,164 particles were selected for consensus refinement. After Bayesian polishing, 3D auto-refine and post processing, a 3.1 Å map was generated for model building.

For nucleosome–CBF3core dataset, 827,118 particles were picked. Bad particles and free nucleosome particles were removed in 2D classification. 194,679 particles were applied for 3D classification, and free nucleosome particles were further discarded. After re-centering, the 115,666 particles were selected for 3D refinement. The blurry density in 2D class averages and 3D reconstruction suggest the flexibility between nucleosome and CBF3core. We divided the particles into nucleosome part and CBF3core part using density subtraction. 3D auto-refine and post processing generated the final 4.0 Å map of density subtracted scFv-CBF3core and 4.2 Å map of density subtracted nucleosome. To elucidate the flexibility of the complex, multibody refinement was applied to study the relative motion of CBF3core to nucleosome (Supplementary Movie 1).

Model building and structure analysis. For the scFv–CEN3 nucleosome complex, an initial model of the CENP-A nucleosome histone octamer and scFv was generated using the free nucleosome structure reconstituted with human histone proteins (PDB: 6O1D)48. The model was fitted into the cryo-EM density map of scFv–CEN3 nucleosome complex. The CEN3 DNA sequence was built into the map from scratch in COOT and the histone octamer and scFv were optimized by manual rebuilding. The whole complex was refined using real space refinement in PHENIX.

For the nucleosome–CBF3core complex, initial model of nucleosome was generated with a rigid body fit into the density using previously built scFv–CEN3 nucleosome structure. DNA sequence was changed to CEN3-Widom 601 sequence based on CENP-A nucleosome with a Widom 601 DNA structure (PDB: 6UZU). Initial map was generated by rigid body fitting using scFv–CEN3 complex structure (PDB: 6GSY). Nucleosome and CBF3core structures were optimized by manually rebuilding in COOT followed by further refinement using real space refinement in PHENIX. Figures were made using UCSF Chimera and PyMOL (Version 1.8, Schrodinger, LLC. DeLano Scientific).

Electrophoretic mobility shift assay. Typical binding reactions of complex formation between CBF3/CBF3core and CEN3-601 nucleosomes were carried out for 60 min on ice in 10 mM Tris at pH 7.4, 75 mM NaCl, and 1 mM DTT. Reactions contained 530 nM nucleosome, and either 106, 212, 318, 424, 530, 636, 795, 1060, and 1166 nM CBFScore, or CBFScore. Ten microliters of the binding reactions were analyzed on 4% acrylamide gels in 0.2 x TBE 100 V for 90 min at 4°C. No ethidium bromide (EtBr) was added at this point to prevent potential disruption of DNA structure by EtBr. After electrophoresis, gels were stained with EtBr and the band intensity was quantified using Image J. The first band above the free nucleosome was taken as the CBF3core–nucleosome complex assuming 1:1 ratio between CBF3core and the nucleosomes. The fraction bound is calculated as the ratio of the intensity between this band and that of total free nucleosome. Binding data were fitted with the Hill equation and analyzed in Prism (Graphpad). Binding of CBFScore does not affect EtBr staining of the nucleosome.

For the binding reactions between CBF3/CBF3core and CEN3 nucleosomes, the mixture was incubated at room temperature for 60 min. Ten microliters of the binding reactions were analyzed by electrophoresis at 100 V for 20–30 min on native agarose gels (Seakem ME and Lonza LE) in 0.2 x TBE. After electrophoresis, gels were stained with SYBR Green I (Invitrogen) and visualized with a Fujifilm LAS-3000 camera. Images were exported into TIFF files for quantification using Image Quanti software (Amersham Biosciences).

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability. Three-dimensional cryo-EM density maps have been deposited in the Electron Microscopy Data Bank under accession numbers EMDB-22696 (CEN3 CENP-A CBF3core Nucleosome–scFv), EMDB-22698 (CEN3-601 H3 Nucleosome in Nucleosome–CBF3core), and EMDB-22697 (CBF3core in Nucleosome–CBF3core). The coordinates of atomic models have been deposited in the Protein Data Bank under accession numbers 7K78 (CEN3 CENP-A CBF3core Nucleosome–scFv), 7K7G (CEN3-601 H3 Nucleosome in Nucleosome–CBF3core), and 7K7N (CBF3core in Nucleosome–CBF3core). Previously published structures used in this study: 6O1D, 6UZU, 6GYS, 6UPH, 6FE8, 6FO7, and 6Q1D are available in PDB databank. All other relevant data supporting the key findings of this study are available within the article and its Supplementary Information files or from the corresponding authors upon request. Source data are provided with this paper.

Received: 14 June 2020; Accepted: 22 February 2021; Published online: 19 March 2021

References

1. Verdaasdonk, J. S. & Bloom, K. Centromeres: unique chromatin structures that drive chromosome segregation, Nat. Rev. Mol. Cell Biol. 12, 320–332 (2011).
2. McKinley, K. L. & Cheeseman, I. M. The molecular basis for centromere identity and function. Nat. Rev. Mol. Cell Biol. 17, 16–29 (2016).
3. Kops, G. J., Weaver, B. A. & Cleveland, D. W. On the road to cancer: aneuploidy and the mitotic checkpoint. Nat. Rev. Cancer 5, 773–785 (2005).
4. Earnshaw, W., Bordwell, B., Marino, C. & Rothfield, N. Three human chromosomal autoantigens are recognized by sera from patients with anti-centromere antibodies. J. Clin. Invest. 77, 426–430 (1986).
5. Meluh, P. B., Yang, P., Glowczewski, L., Koshland, D. & Smith, M. Cse4p is a component of the core centromere of Saccharomyces cerevisiae. Cell 94, 607–613 (1998).
6. Waye, J. S. & Willard, H. F. Nucleotide sequence heterogeneity of alpha satellite repetitive DNA: a survey of alphoid sequences from different human chromosomes. Nucleic Acids Res. 15, 7549–7569 (1987).
7. Manuelidis, L. & Wu, J. C. Homology between human and simian repeated DNA. Nature 276, 92–94 (1978).

NATURE COMMUNICATIONS | https://doi.org/10.1038/s41467-021-21985-9 | www.nature.com/naturecommunications
71. Pettersen, E. F. et al. UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).

**Acknowledgements**

We are grateful to Dr. Jiansen Jiang for the cryo-EM resource and guidance. We thank Dr. Hua Xiao for plasmids and suggestions on reconstitution of the CEN3 nucleosome, and Drs. Alex Kelly, Munira Basrai, Prashant Mishra, and Kentaro Ohkuni for helpful discussion. EM datasets were collected on the microscope in the NIH Multi-Institute Cryo-EM Facility (MICEF) and EM data processing was carried out using the NIH HPC Biowulf cluster (http://hpc.nih.gov). We thank Dr. Huaibin Wang for technical support on electron microscope. Our work was supported by the Francis Crick Institute, which receives its core funding from Cancer Research UK (FC001155), the Medical Research Council (FC001155) and the Wellcome Trust (FC001155) (M.R.S.), and by the intramural research program of Center for Cancer Research, National Cancer Institute, National Institutes of Health (Y.B.).

**Author contributions**

Y.B. and C.W. initiated the study on the CEN3 nucleosome. R.G., Y.B., and M.S. initiated the study on the CBF3–nucleosome complex. M.S. provided the CBF3 protein. B.-R.Z. provided the scFv protein. R.G. prepared the samples and conducted the biochemical experiments. E.H. assisted with mutation studies. R.G. and T.L. collected and processed the cryo-EM data. R.G. built the structural models. R.G. and Y.B. analyzed the structures and wrote the paper with input from all other authors.

**Funding**

Open Access funding provided by the National Institutes of Health (NIH).

**Competing interests**

The authors declare no competing interests.

**Additional information**

**Supplementary information** The online version contains supplementary material available at https://doi.org/10.1038/s41467-021-21985-9.

**Correspondence** and requests for materials should be addressed to Y.B.

**Peer review information** *Nature Communications* thanks the anonymous reviewers for their contribution to the peer review of this work. Peer reviewer reports are available.

**Reprints and permission information** is available at http://www.nature.com/reprints

**Publisher’s note** Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.