

# Structure of the eukaryotic replicative CMG helicase suggests a pumpjack motion for translocation

Zuanning Yuan<sup>1,2</sup>, Lin Bai<sup>2</sup>, Jingchuan Sun<sup>2</sup>, Roxana Georgescu<sup>3,4</sup>, Jun Liu<sup>5</sup>, Michael E O'Donnell<sup>3,4</sup> & Huilin Li<sup>1,2</sup>

The CMG helicase is composed of Cdc45, Mcm2–7 and GINS. Here we report the structure of the *Saccharomyces cerevisiae* CMG, determined by cryo-EM at a resolution of 3.7–4.8 Å. The structure reveals that GINS and Cdc45 scaffold the N tier of the helicase while enabling motion of the AAA+ C tier. CMG exists in two alternating conformations, compact and extended, thus suggesting that the helicase moves like an inchworm. The N-terminal regions of Mcm2–7, braced by Cdc45–GINS, form a rigid platform upon which the AAA+ C domains make longitudinal motions, nodding up and down like an oil-rig pumpjack attached to a stable platform. The Mcm ring is remodeled in CMG relative to the inactive Mcm2–7 double hexamer. The Mcm5 winged-helix domain is inserted into the central channel, thus blocking entry of double-stranded DNA and supporting a steric-exclusion DNA-unwinding model.

In eukaryotes, two Mcm2–7 rings are loaded onto DNA at origins of replication during G1 phase<sup>1,2</sup>. The double hexamer is inactive as a helicase but is activated upon entry into S phase by the Dbf4-dependent Cdc7 kinase, cyclin-dependent kinase and several transient factors that activate origins; subsequently, Cdc45 and GINS assemble onto each Mcm2–7 complex, thereby forming two active CMG helicases that surround each strand of parental DNA<sup>3,4</sup>. CMG then translocates along single-stranded (ss) DNA in the 3'-to-5' direction, thus facilitating bidirectional replication<sup>5–7</sup>. Several archaeal MCM crystal structures are known<sup>8–12</sup>, and the atomic model of the yeast double hexamer of Mcm2–7 has been determined by cryo-EM<sup>13</sup>. These structures have revealed that the Mcm2–7 core forms a two-tiered ring structure: an N-terminal domain (NTD) tier composed of a helical subdomain, a zinc-binding motif and an OB motif, and a C-terminal domain (CTD) tier containing the AAA+ motors. The structure of the 11-protein CMG complex is known only at low resolution: single-particle EM of fruit fly and yeast CMG has indicated that GINS and Cdc45 bind to one side of the Mcm2–7 hexamer, thus forming a double ring-like structure<sup>14–16</sup>.

Most studies of replicative helicases have focused on homohexamers such as papilloma virus E1, SV40 T antigen, archaeal Mcm, *Escherichia coli* DnaB and bacteriophage T7 gene 4 protein<sup>17</sup>. In comparison, studies of the asymmetric 11-protein CMG are scarce, and many fundamental questions remain. Why is CMG an asymmetric ring? How does it translocate along DNA? Does it function differently from the rotary homohexameric helicases? Why are some ATP sites dispensable for CMG helicase activity, whereas others are essential? What is the location in CMG where double-stranded (ds) DNA is separated? There are two general models for the unwinding point:

(i) the steric-exclusion model, in which the dsDNA split point is on the C surface, and leading-strand ssDNA passes through the central channel while lagging-strand ssDNA is sterically excluded to the outside of the ring and (ii) the ploughshare or side channel–extrusion model, in which the duplex is split inside the hexamer channel, and one strand is extruded from a side channel between the CTD and NTD tiers<sup>18</sup>.

Translocation of a helicase along DNA is presumed to require at least two DNA-binding sites that differ in relative distance from each other during the ATP cycle, thus promoting movement along DNA<sup>17</sup>. Structural studies of papilloma virus E1 helicase have suggested the presence of six equivalent sites, the DNA-binding loops in the motor domains, that undergo changes in relative distance from one another in a rotary fashion, thereby escorting ssDNA through the ring in steps of one nucleotide per ATP<sup>10</sup>. The structure of bacterial DnaB has also revealed that the motor domain of each subunit binds ssDNA, and it has been proposed that entire protomers translocate along ssDNA in a spiral staircase-like manner, in steps of two nucleotides per ATP<sup>19</sup>. CMG lacks the six-fold symmetry of homohexamers, and whether CMG acts in a rotary fashion is unknown.

To address the conformational changes during CMG translocation, and CMG's mode of unwinding (steric exclusion or side-channel extrusion), we used cryo-EM to derive an atomic model of yeast CMG. We found that the Cdc45–GINS–Mcm architecture is sculpted to rigidify the Mcm2–7 NTD ring while still allowing movements within the Mcm2–7 CTD motor ring. The CTD ring adopts two different conformations, tilted and untilted, relative to a fixed NTD ring, thus giving rise to compact and extended forms of CMG with a maximum domain translocation of about 20 Å. Five of the Mcm2–7 subunits contain C-terminal winged-helix domains (WHDs) that

<sup>1</sup>Department of Biochemistry & Cell Biology, Stony Brook University, Stony Brook, New York, USA. <sup>2</sup>Biology Department, Brookhaven National Laboratory, Upton, New York, USA. <sup>3</sup>DNA Replication Laboratory, Rockefeller University, New York, New York, USA. <sup>4</sup>Howard Hughes Medical Institute, Rockefeller University, New York, New York, USA. <sup>5</sup>Department of Pathology and Laboratory Medicine, University of Texas Medical School at Houston, Houston, Texas, USA. Correspondence should be addressed to M.E.O'D. (odonnell@rockefeller.edu) or H.L. (hli@bnl.gov).

Received 20 November 2015; accepted 6 January 2016; published online 8 February 2016; doi:10.1038/nsmb.3170

extend outward from the CTD and appear flexible and poorly defined in the double hexamer<sup>13</sup>. However, in CMG the WHDs of Mcm5 and Mcm6 are localized, and the WHD of Mcm5 is pushed into the interior axial channel of the CTD motor ring, where it restricts the large-diameter pore to a size too narrow to accommodate dsDNA. These results support a steric-exclusion mechanism for CMG, in which leading-strand ssDNA threads through the axial channel, and the lagging strand is excluded, as indicated by biochemical studies in the *Xenopus* system<sup>20</sup>. The high-resolution structures of the two CMG conformers observed here indicate that the CTD motor ring tilts up and down relative to a rigid NTD ring–Cdc45–GINS platform, in a manner akin to the nodding head of an oil-rig pumpjack. Alternation between these structures suggests that CMG might move on DNA in a linear fashion, ratcheting along ssDNA during unwinding without a need for rotary action.

## RESULTS

### 3D reconstruction of CMG helicase

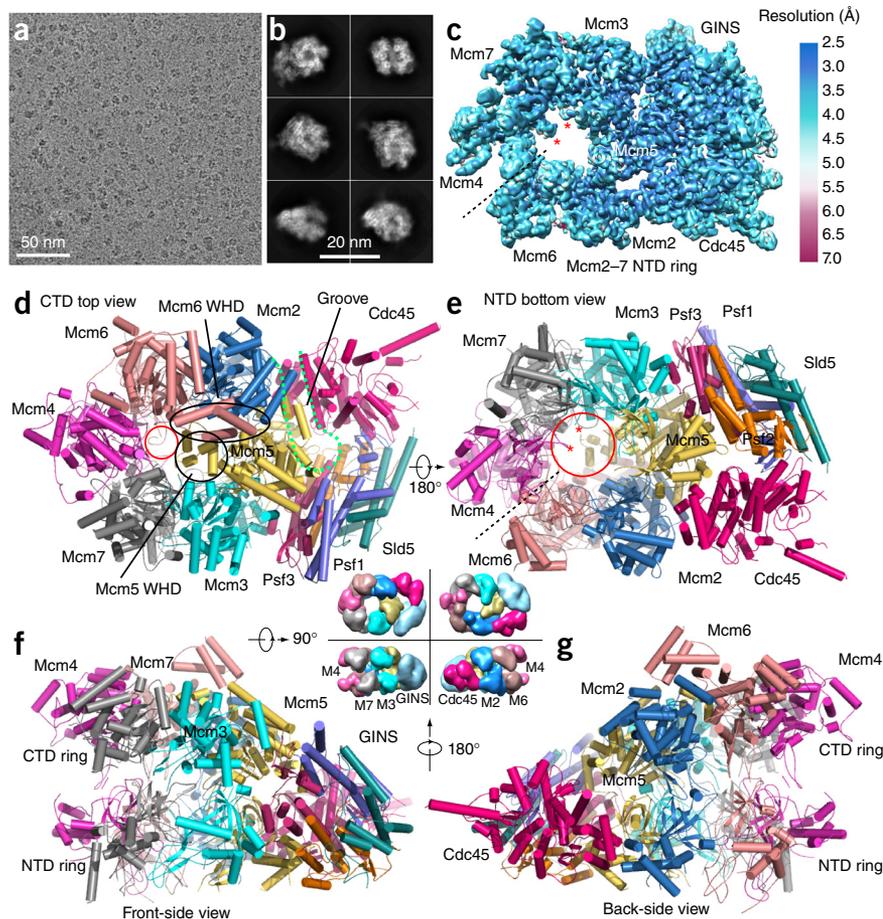
Recombinant *S. cerevisiae* CMG was expressed and purified from yeast, and cryo-EM images were recorded on a Titan Krios microscope with a K2 Summit direct electron detector (Fig. 1a,b). After two-dimensional (2D) and three-dimensional (3D) classification, we initially reconstructed a 3D density map of the full CMG complex at 3.8-Å resolution from ~470,000 particle images (Supplementary Table 1 and Supplementary Figs. 1 and 2). Interestingly, the Cdc45–GINS–Mcm2–7 NTD ring showed solid density, but the density of the Mcm2–7 CTD ring was much weaker, thus indicating partial flexibility. Excluding the flexible CTD motor region, the 3D density map had an estimated resolution of 3.7 Å (Fig. 1c and Supplementary Fig. 3). We subsequently carried out focused 3D classification in the CTD motor region and obtained two conformers for the CMG helicase, at 4.8 Å and 4.7 Å (Supplementary Figs. 4–6; discussed below). The lower resolution was probably a result of a decreased

number of particles belonging to each conformer but could also have been an indication that the helicase exists in a continuum of conformations<sup>21</sup>. This possibility was underscored by the gold-standard Fourier-shell-correlation curve, which started to drop as early as 6.5 Å. We first built the Cdc45, GINS and the Mcm2–7 NTD ring on the basis of the 3.7-Å-resolution density map as well as the previously reported inactive yeast Mcm2–7 structure and that of the human GINS homolog<sup>13,22–24</sup>. We then used the 3.7-Å structure and the 3.8-Å structure of the CTD ring of the inactive Mcm2–7 to model CMG conformers I and II.

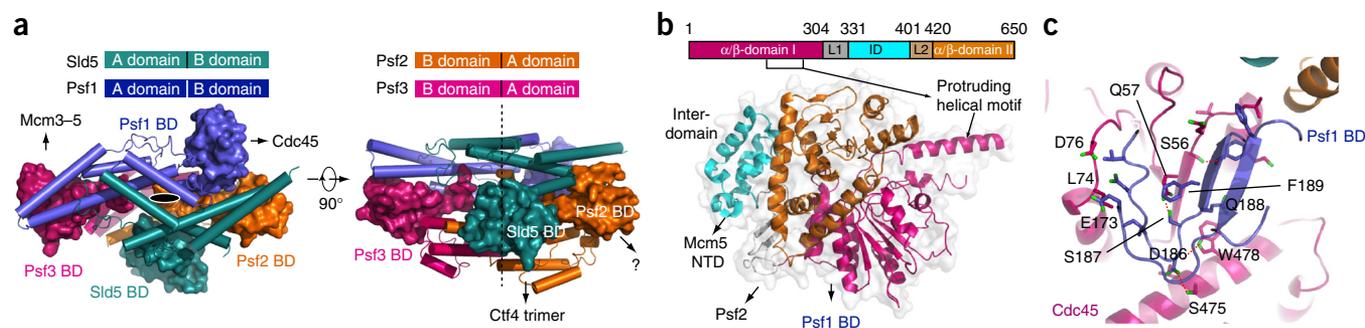
### Overall structure of CMG

CMG has a relatively large ~20-Å-diameter opening at the bottom NTD ring, where the leading strand presumably exits the Mcm2–7 axial channel (Fig. 1c–g). The axial channel in the NTD ring is partially obstructed by two protruding loops that connect the first two β-strands of the OB subdomains in Mcm4 and Mcm7 (Fig. 1c,e). These loops bind ssDNA in the archaeal MCM hexamer, and two conserved DNA-binding arginine residues in the yeast Mcm4 and Mcm7 OB domains are essential for function<sup>25</sup>. In CMG, these OB domains are positioned directly below the ssDNA pore of the CTD ring where DNA enters (Fig. 1d). Therefore, the leading strand coming from the CTD ring probably encounters and interacts with the two channel-lining OB folds in the NTD ring.

The Mcm2–7 NTD ring is highly stabilized by an extensive network of bracing interactions with Cdc45 and GINS. Interestingly, the only contacts between Cdc45 and Mcm2 or Mcm5 are with the NTD, and Cdc45 makes no contact with the CTDs, thus resulting in a remarkably wide groove between the Mcm2–7 CTD and Cdc45 (Fig. 1d–g).



**Figure 1** Cryo-EM and overall structure of the *S. cerevisiae* CMG complex. (a) A typical motion-corrected raw image from ~8,000 images of frozen CMG particles recorded on a direct detector. (b) Six selected 2D averages representing the particles in different views. (c) 3D cryo-EM map of CMG, color-coded according to local resolution. The overall resolution is 3.7 Å. This density map includes the Cdc45–GINS–Mcm2–7 NTD ring. The two red asterisks mark the hairpin loops in the NTD OB subdomains of Mcm4 and Mcm7 that protrude into the axial channel. (d–g) Cartoon views of the full-length atomic model of CMG conformer I as viewed from the CTD motor ring (d), NTD bottom (e), front side (f) and back side (g). The dashed black lines in c and e mark the smaller interface between the NTDs of Mcm4 and Mcm6. The dashed green curve in d marks the groove between Cdc45 and Mcm2–5. The small red circle in d marks the constricted opening (10 Å) at the CTD motor ring. The large red circle in e marks the 20-Å channel opening at the bottom. The black circle and the oval in d mark the WHDs of Mcm5 and Mcm6, respectively. The central four surface representations of CMG in d–g are an aid to understanding the high-resolution structures. They represent the lower-resolution EM map of CMG (EMD-6463)<sup>16</sup>.



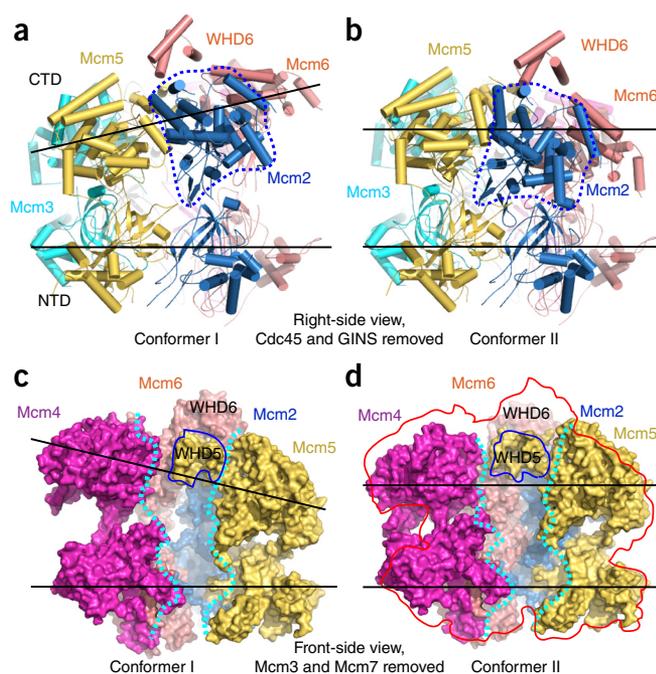
**Figure 2** Structure and interactions of yeast GINS and Cdc45. **(a)** The full-length GINS structure in top and side views. Domain A is shown in cartoon and domain B in surface. Top, schematic showing that all four subunits have a similar two-domain architecture, but domains A and B in Psf2 and Psf3 are inverted in comparison to those in Sld5 and Psf1. The black oval and dashed line indicate the pseudo two-fold-symmetric arrangement of the GINS complex. The question mark indicates that the Psf2 B domain may have an unknown interacting partner in the functional replisome. **(b)** The full-length structure of Cdc45 in cartoon representation superimposed on the transparent surface view. The two  $\alpha/\beta$ -domains are shown in magenta and brown, respectively. The middle all-helix interdomain (ID) is connected to the N-terminal and C-terminal  $\alpha/\beta$ -domains via two linkers (L1 and L2). Part of L1 and L2 further form a two stranded  $\beta$ -sheet. The ID primarily interacts with NTDs of Mcm2 and Mcm5. The L1 loop region interacts with Psf2. **(c)** Extensive interactions between GINS Psf1 C-terminal BD (blue) and Cdc45 (magenta). The two  $\beta$ -strands of Psf1 BD adjoin with the four  $\beta$ -strands in  $\alpha/\beta$ -domain I of Cdc45, forming a six-stranded  $\beta$ -sheet core across the two proteins. Four hydrogen bonds are drawn as dashed red lines.

This architectural facet allows major changes to occur in the CTD ring between the two CMG conformers while the NTD ring is still stabilized and held rigid. Specifically, in conformer I, the CTD ring is tilted relative to the NTD ring through upward displacement of the motor domains of the Mcm2, Mcm6 and Mcm4 subunits, and the interface between the CTDs of Mcm2 and Mcm5 springs open. The presence of ATP sites at the CTD-subunit interfaces suggests that ATP hydrolysis at the Mcm2-5 interface leads to conformer I, in which this interface is disrupted. The other function that the Cdc45-GINS-Mcm NTD architecture may serve is to hold onto DNA when the Mcm2-5 interface opens, because DNA is known to pass through this interface during origin activation<sup>14,26,27</sup>. If DNA were to escape the CTD ring via the dynamic Mcm2-5 interface, the Cdc45-GINS subunits that span the Mcm2-5 interface would prevent the helicase from falling off the DNA, as indicated by previous CMG-DNA cross-linking studies<sup>28</sup>. Within the Mcm2-7 NTD ring, the interface with the least buried

surface area is the Mcm4-6 interface (**Fig. 1c,e**). If CMG needs to open for bypass of a leading-strand block, the Mcm4-6 interface may serve as a 'pressure-release valve' for NTD-ring opening.

In the active helicase compared to the inactive double-hexameric structure, Mcm2-7 is substantially remodeled. We observed defined WHD densities for Mcm5 and Mcm6, unlike those in the Mcm2-7 double hexamer<sup>13</sup>, whose WHDs are flexible and poorly defined on the CTD surface outside the central channel (**Fig. 1d-g** and **Supplementary Figs. 5 and 6**). Notably, the WHD of Mcm5 is stuffed inside the central pore, and the Mcm6 WHD stacks on top of it. The interior location of the Mcm5 WHD is consistent with that in our earlier MS analysis demonstrating that the only WHD that cross-links to residues lining the channel is the Mcm5 WHD, even though all other Mcm subunits except Mcm2 also contain a WHD<sup>16</sup> (**Supplementary Fig. 7a-c**). The location of the Mcm5 WHD constricts the diameter of the central channel to  $\sim 10$  Å, a diameter too small to accommodate dsDNA, which is 20 Å in diameter (**Fig. 1d**). Therefore, the Mcm5 WHD may provide a structural basis for the steric-exclusion mechanism of DNA unwinding. The most dramatic alteration between CMG and the double hexamer is the existence of at least two distinct conformers in the CTD AAA+ motor ring of CMG (described below).

The surface of CMG is predominantly of mixed charge (**Supplementary Fig. 7d-g**), whereas the lumen of the CTD pore is positively charged and suitable for entry of ssDNA. The axial channel



**Figure 3** Side-by-side comparison of conformer I and conformer II in the Mcm2-7 region of CMG helicase. **(a,b)** Comparison of the two conformations, shown in cartoon representation and viewed from the right side, from Cdc45 and GINS (which are both removed for clarity) with the CTD motor ring on top and the NTD ring at the bottom. The two black lines show that the CTD motor ring is tilted by  $\sim 10^\circ$  with respect to the NTD ring in **a** but is nearly parallel to the NTD ring in **b**. The dashed blue line encircles the Mcm2 CTD. This domain is rotated  $30^\circ$  clockwise in **b** compared to **a**. **(c,d)** Front-side surface view of conformer I (**c**) and conformer II (**d**), with the front Mcm3 and Mcm7 removed to show the Mcm2-7 axial channel, which is approximately demarcated by two dashed cyan curves. Cdc45 and GINS at the right side are also removed for clarity. In both conformers, the Mcm5 WHD is located inside the axial channel. The Mcm6 WHD sits on top of the Mcm5 WHD. The red shape in **d** is a contour of **c** superimposed on conformer II to show the changes in the CTD motor ring between the two conformers.

**Figure 4** Superposition of CMG conformers I and II. **(a,b)** Top view **(a)** and back-side view **(b)** showing the movement of the Mcm2–6–4 half motor ring. **(c)** Front-side view showing the smaller movements of the Mcm5–3–7 half ring. Conformer I is shown in dark gray and conformer II in color. The distances shown in **b** and **c** are measured for the AAA+ domains between the two conformers.

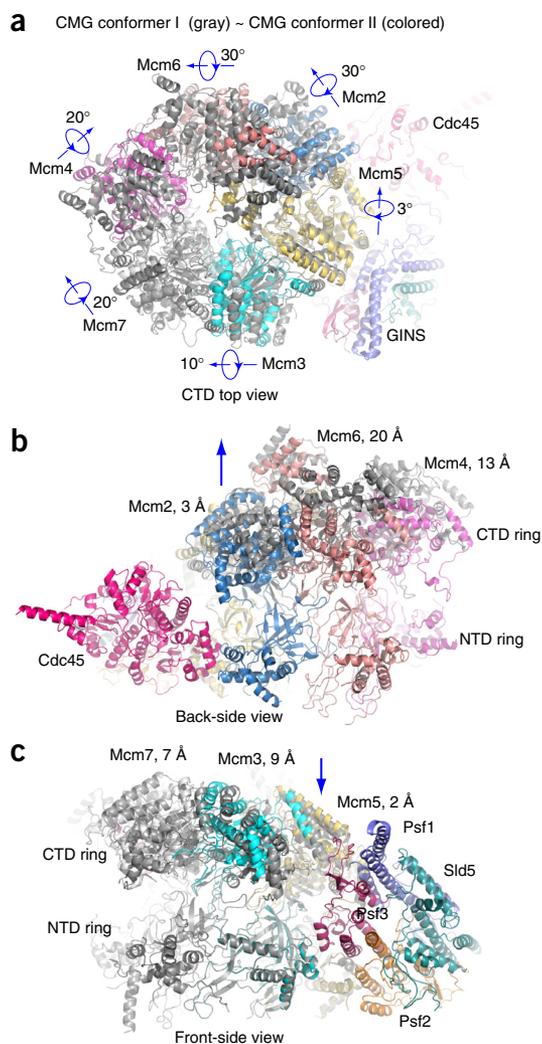
is positively charged in the NTD ring, largely because of the channeling loops of OB domains, which have previously been shown to interact with ssDNA<sup>25</sup>. Furthermore, there is a nearly continuous and positively charged path on the outer surface of Mcm3 that extends from the top CTD to the bottom of the CMG structure. Because the lagging strand interacts with the outer surface of an archaeal Mcm helicase<sup>29</sup>, it is possible that the lagging strand follows the positively charged route to reach the polymerase (Pol)  $\alpha$  polymerase-primase, which we have previously shown to be below the NTD ring of CMG<sup>16</sup>.

In the context of the CMG complex, the function of individual GINS subunits as scaffolding points for different proteins became clear. Each GINS subunit has a two-domain architecture: a four-helix-bundle region (A domain) and an  $\alpha/\beta$ -region (B domain (BD)). Similarly to the architecture of human GINS<sup>22–24</sup>, the A domains of the four subunits stack to form a pseudo two-fold-symmetric flat core onto which the four BDs are attached at the periphery (Fig. 2). The Psf1 BD is not visible in the crystal lattice of human GINS<sup>22–24</sup>. The Psf1 BD is not needed for GINS-complex formation, but it is essential for CMG formation and has been proposed to bind Cdc45 (ref. 24). In previous low-resolution EM maps of CMG, a density between GINS and Cdc45 was tentatively assigned to the Psf1 BD<sup>14,16</sup>. The CMG structure clearly showed that this density is the Psf1 BD, which folds into an  $\alpha/\beta$ -motif similarly to the corresponding domain in Sld5 (Fig. 2). The Psf1 BD interacts with Cdc45 by forming a  $\beta$ -sheet core crossing the two proteins. Interaction between the Psf1 BD and Cdc45 is of a mixed nature and is extensive, burying an interface of 820 Å<sup>2</sup>. The Psf3 BD clearly interacts with Mcm3–5, strengthening the Mcm3–5 interface (Fig. 1e,f). The Sld5 BD is located at the bottom and, together with a short N-terminal helix that was not resolved in the current structure, binds the Ctf4 trimer<sup>16,30</sup>. By analogy to the other B domains, the Psf2 BD may bind a protein that is not yet identified.

We found that Cdc45 displayed a prominent helical motif extending far beyond the previously determined low-resolution profile of CMG<sup>16</sup> (Figs. 1d,e and 2b). This helical protrusion is proximal to the catalytically active domain of Pol  $\epsilon$ , because a lysine at the tip of the helix cross-links to the N-terminal catalytic domain of Pol2 (ref. 16). Cdc45 has been predicted to have a RecJ-like fold<sup>31–33</sup>. The structure showed that yeast Cdc45 contains two RecJ-like  $\alpha/\beta$ -domains separated by a small helical interdomain (ID) (Fig. 2b). The helical ID contacts and stabilizes the NTDs of Mcm2 and Mcm5. The double  $\alpha/\beta$ -domain fold is also found in GT-B-type glycosyl-transferases<sup>34</sup>. Full-length RecJ forms an O-like structure with a pore in the middle surrounded by its four domains, and there is an ssDNA-binding groove between the N-terminal two  $\alpha/\beta$ -domains<sup>35,36</sup>. However, yeast Cdc45 did not form an analogous central pore or DNA-binding groove and supported the  $\alpha$ -helical protrusion instead.

### CMG fluctuates between planar and tilted forms

The cryo-EM images were recorded for pure CMG with no added nucleotides or DNA, so we expected the structure to be in the apo form. Indeed, we observed no nucleotide or DNA density in any of our 3D reconstructions. Among the ~470,000 CMG particles used for 3D reconstruction, ~40% of the particles existed in the extended and



tilted form termed conformer I, and ~20% of the particles were in the untilted planar form termed conformer II (Fig. 3 and Supplementary Figs. 4 and 5). In conformer I, when viewed from the Cdc45–GINS side of the structure, the CTD ring containing the AAA+ motors of Mcm2–7 is tilted by ~10° with respect to the NTD ring, leading to an approximate spiral arrangement of domains composing the CTD ring (Fig. 3a). In conformer II, the CTD motor ring is approximately parallel to the NTD ring, and CMG is more compact than in conformer I (Fig. 3b). The Mcm5 WHD is inside the axial channel in both conformers (Fig. 3c,d). By superimposing the two conformers, we found that each of the six Mcm2–7 CTD AAA+ motor domains undergoes both rotation and translation, to varying degrees (Fig. 4a–c and Supplementary Videos 1 and 2). The Mcm6 AAA+ motor domain in conformer I, compared to conformer II, is translated upward 20 Å, while the AAA+ motor domains of the neighboring Mcm2 and Mcm4 are shifted up by 3 Å and 13 Å, respectively (Fig. 4b). When viewed from the right side (Cdc45 and GINS), the Mcm2 AAA+ motor domain in conformer I, compared to conformer II, is rotated 30° and has moved away from the Mcm5 AAA+ domain, thus leaving a small gap between Mcm2 and Mcm5 (Fig. 3a,b). In overview, conformer I is an expanded form of CMG, whereas conformer II is more compact. It is possible that the CTD motor ring may have additional conformations that were not captured in the current focused 3D classification, because the two conformers accounted for ~60% of the CMG particles. Earlier low-resolution EM studies of CMG have indicated

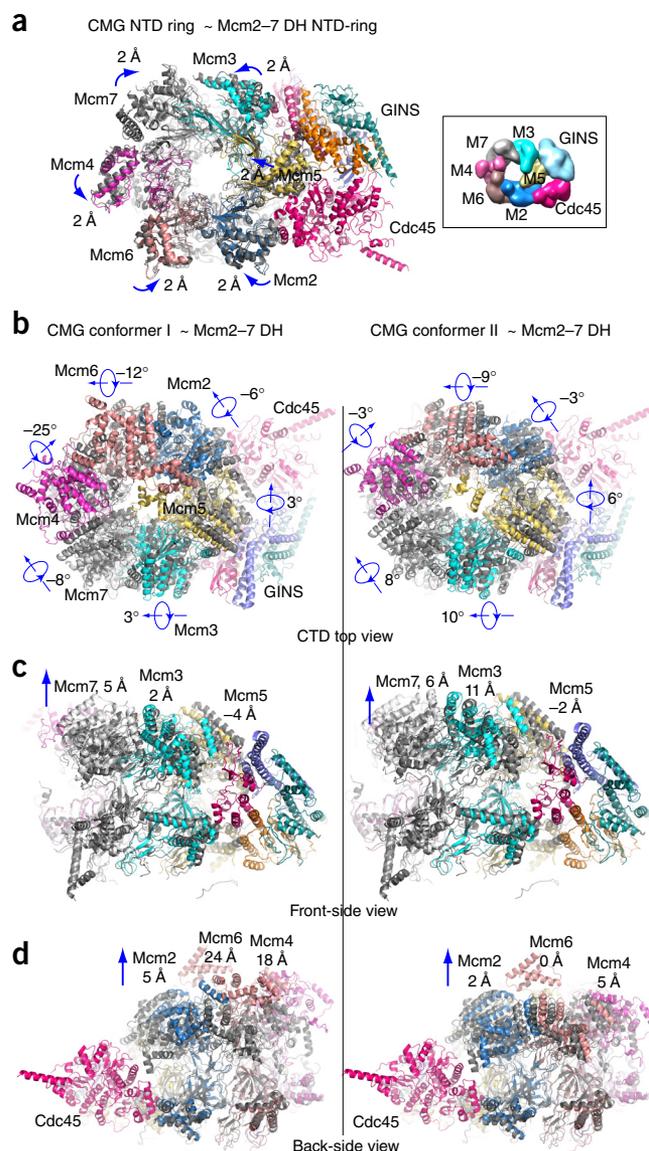
**Figure 5** Remodeling changes between Mcm2–7 in the double hexamer (DH) and in the active CMG conformers. **(a)** Superimposition of the Mcm2–7 NTD-tier ring of CMG (colored cartoon representation) with that of the inactive double hexamer (gray cartoon representation). The double-hexamer Mcm2–7 structure is (EMD-6338)<sup>13</sup>. The direction and distance of rigid-body domain movements are noted for each subunit. Inset, surface representation of the NTD view of the previously reported 3D density map of CMG (EMD-6463)<sup>16</sup>. **(b–d)** Comparisons of the top view **(b)**, the front-side view **(c)** and back-side view **(d)** of the CTD tier of Mcm2–7 of the double hexamer with the Mcm2–7 CTD tier of CMG conformer I (left) and CMG conformer II (right). Each CTD AAA+ domain undergoes a combination of rotation and translation. The degree of rotation is shown in **b** and the vertical translation in **c** for Mcm3, Mcm5 and Mcm7, and in **d** for Mcm2, Mcm4 and Mcm6. Another major change is the insertion of the Mcm5 WHD into the interior channel in the active helicase, but this is not illustrated here because the deposited Protein Data Bank coordinates of the inactive Mcm2–7 hexamer do not contain the Mcm5 WHD. The Mcm5 WHD in DH is visualized as weak density outside the central channel<sup>13</sup>.

two forms: an apo form with an opening in the AAA+ tier between Mcm2–5 and an ATP form with both tiers closed<sup>14,15</sup>. These forms probably correspond to conformers I and II, respectively.

The two distinct conformers of CMG indicate a major difference relative to the single conformation of the inactive Mcm2–7 double hexamer. Remodeling of the Mcm2–7 ring transiting from the double hexamer to the two CMG conformers requires numerous domain translations of up to 20 Å and rotations of up to 30° (Fig. 5). The previous study of the CMG–Pol ε structure has determined numerous intrapeptide cross-links through an MS approach<sup>16</sup>. The precise structures of GINS–Cdc45 within CMG permitted mapping of cross-links between CMG and the Pol2 subunit of Pol ε, and revealed a Pol ε footprint on the helicase that further supports the conclusion that Pol ε rides ahead<sup>16</sup> rather than trailing behind the helicase (Fig. 6). Two cross-links between the active N half of Pol2 and Cdc45 locate to the tip of the protruding helix of Cdc45, thus suggesting a possible role of Cdc45 in communication between helicase and polymerase activity (Fig. 6).

### Nodding-pumpjack unwinding model

Constriction of the ~25-Å axial channel in Mcm2–7 to ~10 Å by the Mcm5 WHD restricts the CTD channel to allow entry of only ssDNA (Figs. 3c,d and 4a), consistently with a steric-exclusion mechanism of unwinding and results from studies in the *Xenopus* system<sup>20</sup>. Translocation of homohexameric helicases has been proposed to occur in a rotary fashion in which an opening between motor domains circularly permutes around the ring with ATP hydrolysis<sup>10,17,19</sup>. The heterohexameric Mcm2–7 within CMG may function in a rotary fashion, but there are compelling data that may suggest otherwise. First, in a rotational model, the opening and vertical displacement observed between CMG conformers I and II upon opening of the Mcm2–5 interface would need to occur in a sequentially permuted fashion at each Mcm-subunit interface around the ring. However, the GINS interacts with the AAA+ domain of Mcm5 and possibly Mcm3, and this interaction may block the rotary propagation of large AAA+ motions. Second, only two ATP sites are necessary for CMG helicase activity (Mcm5–2 and Mcm3–5), whereas two ATP sites (Mcm 4–7 and Mcm 6–4) can be mutated without observable effect; mutations in the remaining two sites reduce helicase activity only by half<sup>6</sup>. The two required ATP sites are at the hinge point between the two conformers, and the nonessential sites are the most distant from the hinge. Rotary models generally require all or most of the ATP sites. Finally, the compact and extended shapes of the two CMG conformers immediately suggest another type of mechanism. We propose a



speculative new ‘pumpjack’ model that is linear instead of rotational. We freely concede that there are few data to back this proposal, and further structural advances (for example, CMG–DNA structures), biochemical experiments and single-molecule studies are needed to test its applicability to CMG and its ability to be generalized to other helicases. Our hope is that the model may foster new ideas and future experimentation to test the hypothesis.

In alternation between the two conformers, a ratcheting or inchworm-like expansion-contraction process is apparent (Supplementary Videos 1 and 2). A more accurate analogy of the movements is that of an old-fashioned oil-rig pumpjack (which resembles a horse’s head nodding up and down) sitting on top of an immobile platform. In the case of CMG, the Mcm2–7 NTD ring and the Cdc45–GINS side brace appear rigid, like a platform, and the Mcm2–6–4 act like the nodding horse’s head. Because the Mcm5 WHD stays inside the channel in both conformers, it may function as a ‘pin’ that facilitates strand separation as ssDNA is pulled through the motor domains. On the basis of these structural features, we propose a nodding-pumpjack mechanism of DNA unwinding that proceeds in a simple linear fashion along DNA (Supplementary Video 3). In the nodding-pumpjack model, the binding of ATP brings the interface between

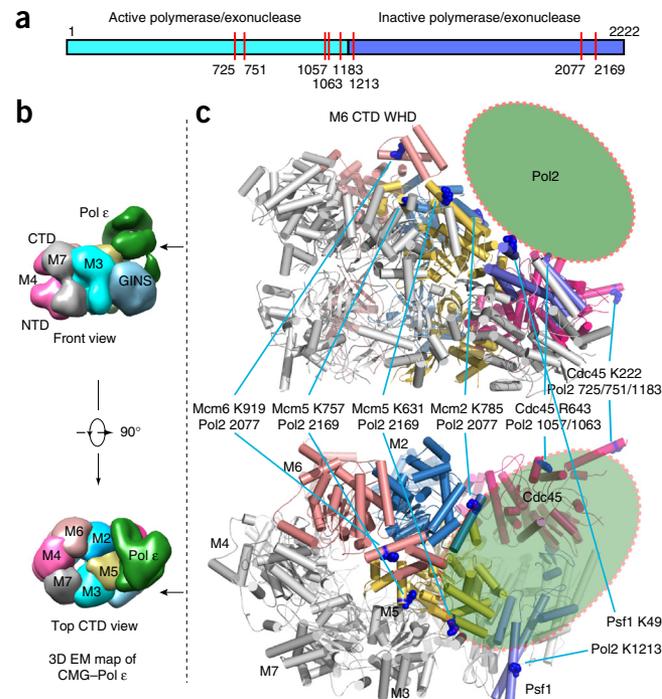
**Figure 6** Pol2 footprint on the atomic model of CMG helicase. (a) The two-domain architecture of Pol2, the catalytic subunit of the Pol  $\epsilon$  complex. The N-terminal half contains the polymerase and exonuclease activities. The C-terminal half is homologous to a B-family polymerase but lacks enzymatic activity. The red vertical bars indicate the locations of the residues in Pol2 that cross-link to subunits of CMG<sup>16</sup>. (b) Surface-rendered front and top views of the previously reported 3D density map of the CMG–Pol  $\epsilon$  complex (EMD-6465)<sup>16</sup>. (c) The top and bottom panels are the corresponding front and top views of the atomic model of the CMG helicase. The helicase residues that cross-link to Pol2 are shown as blue spheres. The green oval represents the approximate Pol2 position on CMG, as predicted from the cross-link footprint. Slashes denote ‘and’.

Mcm2 and Mcm5 into a closed position (conformer II), thus causing the Mcm2–6–4 horse’s head to nod and making CMG more compact (Fig. 7a,b). ATP hydrolysis enables CMG to spring open at the Mcm2–5 interfacial ATP site (conformer I), causing Mcm2 to rotate by 30° and the Mcm2–6–4 horse’s head to subsequently move upward (Fig. 7b,c), thereby readying the helicase for the next nodding cycle. In this scheme, the Cdc45–GINS–Mcm2–7 NTD ring acts as a platform upon which the CTD motor ring strokes up and down, powered by ATP binding and hydrolysis. The pumpjack model is essentially an inchworm-like ratchet that has extensive precedents in nonreplicative monomeric helicases<sup>37–43</sup>.

## DISCUSSION

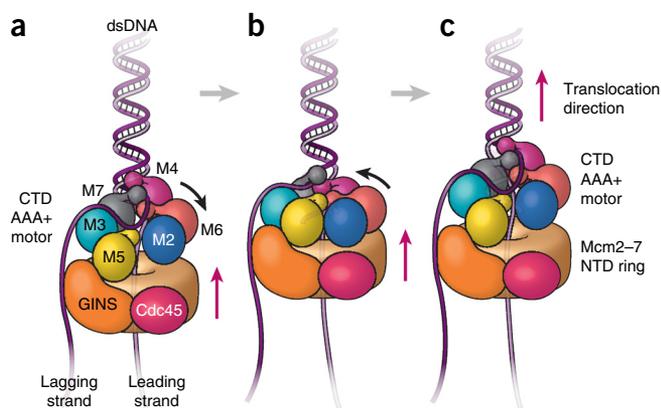
Enzymes are currently viewed as existing as a population of conformations, in which individual proteins rapidly sample many conformations, each with a role in catalysis<sup>44–46</sup>. Therefore, we presume that the two conformers of apo CMG are relevant to function and that ATP binding and hydrolysis act to direct them, as shown previously<sup>14,47</sup>. The two CMG conformers display a large change in distance between the CTD and NTD, and thus the axial channel for DNA becomes shorter or longer by alternating between the two forms. The simplest interpretation is that ATP drives successive cycles between the compact and extended forms, thus resulting in linear translocation (Supplementary Video 3).

Directional translocation along DNA can be produced by a protein that binds DNA at two sites and changes the intersite distance in two conformers that alternate with ATP hydrolysis. This would enable a hand-over-hand action similar to that of a person climbing a rope. The DNA-binding sites in papillomavirus E1 are six equivalent interior loops in the motor domains that are spirally arranged, and the two conformers in the E1 crystal suggest that they move in a rotary fashion, thereby escorting ssDNA through the central channel<sup>10,11</sup>. The bacterial DnaB helicase is shaped like a spiral, and each subunit binds DNA via interior loops in the six motor domains;



it has been proposed that entire protomers move along DNA in a spiral staircase-like manner<sup>19</sup>. During the rotary motion in E1 and DnaB, there is a gap between motor domains at the top and bottom of the spiral, and with each ATP-hydrolysis event the DNA-binding element at the bottom of the spiral ‘leapfrogs’ over the sites in the other subunits, thereby assuming a new position at the top of the spiral<sup>10,19</sup>. This leapfrog motion permutes the gap sequentially around the ring with ATP hydrolysis. The Rho, T7gp4, SV40 large T antigen and 19S proteasomal cap are hexameric ATPases that have also been discussed in terms of rotational processes<sup>17</sup>.

It is reasonable to propose that a gap between two motor domains of a circular homohexamer rotationally permutes around the ring. But CMG is asymmetric, and although it may still function through a rotational model, there are compelling reasons to propose a linear model instead (discussed earlier). Furthermore, both E1 and DnaB have a gap between two motor domains that has been proposed to remain open as it permutes around the ring. In contrast, one CMG conformer is a closed ring with no gap, and the other has an Mcm2–5 gap. It is possible that the open conformer is used only for origin activation, but because the Mcm2–5 gap is sealed in CMG



**Figure 7** Nodding-pumpjack model of CMG translocation. The leading strand (light purple) goes through the Mcm2–7 ring while the lagging strand (dark purple) is excluded to the outside. Mcm2–7 is composed of the CTD AAA+ motor ring (subunits labeled and colored) and the NTD ring (solid tan cylinder). The Mcm NTD ring, Cdc45 and GINS form a stable platform upon which the Mcm2–7 CTD AAA+ motor ring moves up and down. The small knobs attached to the Mcm AAA+ domains are the WHDs. The Mcm5 WHD plugs the central channel, thus preventing dsDNA entry and reinforcing the steric-exclusion model. (a) CMG conformer I is the expanded form. (b) ATP binding results in nodding of the CTD ring, thus forming the compact conformer II and bringing the NTD–GINS–Cdc45 platform upward. (c) ATP hydrolysis leads back to the expanded conformer I, which translocates the CTD motor domain ring upward. Although all six AAA+ domains move during the DNA-translocation cycle, the major movement is by Mcm2–6–4, as revealed by the two conformers captured by cryo-EM. Repeated cycles of ATP hydrolysis result in continued CMG translocation along the leading strand.

by the GINS–Cdc45–NTD ring, strand passage cannot occur, thus suggesting that the conformer is an intermediate in helicase action. It also seems reasonable that other interfaces of CMG may undergo open-closed transitions.

Translocation by a linear pumpjack ratchet requires only two DNA-binding sites. Therefore, we propose that there are only two DNA sites in CMG, although additional sites may be involved. One DNA site in CMG is probably the same as that established in homohexameric helicases and is probably formed by the six DNA-binding loops in the AAA+ regions of the CTD motor domains<sup>8–10</sup>. DNA-binding loops in AAA+ domains have previously been established in AAA+ clamp-loader heteropentamers, and we propose that these loops function as an amalgam of one DNA-binding site in CMG, as in the clamp loaders<sup>48–50</sup>. We propose that the second DNA-binding site is in the NTD and is formed by the OB folds, which are well-established DNA-binding elements in the archaeal Mcm NTD and are required for yeast Mcm2–7 function *in vivo*<sup>25</sup>. Bacterial Rho also binds ssRNA through six OB folds located in the NTD<sup>51</sup>. However, other elements may compose the DNA-binding sites in CMG, for example the WHDs of Mcms 3, 4, 6 and 7. Clearly, additional structural information of CMG bound to DNA is needed to identify possible DNA sites for structure-directed mutagenesis.

Translocation requires that DNA-binding sites not only change intersite distance but also switch affinity for DNA during an ATP-hydrolysis cycle. The CTD DNA-binding site in the ATPase motor domains seems to be the most likely candidate site that changes affinity in response to the ATP state. In its most simple form, the binding affinities between the CTD and DNA would alternate between tight and loose relative to an intermediate fixed-affinity NTD DNA-binding site. For example, in the ATP state, tight CTD binding relative to the NTD site would release DNA from the NTD faster than from the CTD. In the ADP or apo state, weak CTD binding relative to the NTD site would release DNA from the CTD faster than from the NTD. Together, these phenomena would result in translocation along DNA with cycles of ATP hydrolysis. Reverse polarity of tracking may result from binding ssDNA in the opposite orientation or from reversing the relative affinity of ATP versus ADP states in the CTD.

Why is CMG circular if it doesn't act in a rotary fashion? In fact, many (nonreplicative) helicases are monomers or dimers, and thus helicase action does not require circularity<sup>37–42</sup>. Perhaps replicative helicases are ring shaped for processivity, as has been proposed earlier<sup>10</sup>. It should be noted that the results to date do not specifically exclude a linear pumpjack model for the homohexameric helicases. In fact, hexameric archaeal helicases can function when they are 'doped' with up to three mutated subunits<sup>52</sup>, thus suggesting a nonsequential and possibly a nonrotational mechanism. In summary, the CMG structure supports unexpected and specific predictions about its mechanism; these predictions may possibly extend to other replicative helicases and provide a foundation for further dissection of the eukaryotic replisome.

## METHODS

Methods and any associated references are available in the [online version of the paper](#).

**Accession codes.** The 3D cryo-EM maps of CMG at 3.7-Å, 4.8-Å and 4.7-Å resolution have been deposited in the Electron Microscopy Data Bank under accession codes EMD-6534, EMD-6536 and EMD-6535, respectively. The corresponding atomic models have been deposited in the Protein Data Bank under accession codes PDB 3JC6, PDB 3JC7, PDB 3JC5, respectively.

*Note: Any Supplementary Information and Source Data files are available in the online version of the paper.*

## ACKNOWLEDGMENTS

Cryo-EM data were collected on a Titan Krios I at the Howard Hughes Medical Institute, Janelia Farm. We also collected a cryo-EM data set on an FEI Polara with a K2 detector at the University of Texas Health Science Center. We thank the staff at these facilities for help with data collection. We also thank L. Pellegrini (University of Cambridge) for sharing the structure of human Cdc45 before publication. This work was funded by the US National Institutes of Health (GM111472 and OD12272 to H.L. and GM115809 to M.E.O'D.) and the Howard Hughes Medical Institute (M.E.O'D.).

## AUTHOR CONTRIBUTIONS

Z.Y., L.B., J.S., R.G., M.E.O'D. and H.L. designed experiments. Z.Y., L.B., J.S., R.G. and J.L. performed experiments. Z.Y., L.B., J.S. and H.L. analyzed the data. M.E.O'D. and H.L. wrote the manuscript.

## COMPETING FINANCIAL INTERESTS

The authors declare no competing financial interests.

Reprints and permissions information is available online at <http://www.nature.com/reprints/index.html>.

1. Evrin, C. *et al.* A double-hexameric MCM2-7 complex is loaded onto origin DNA during licensing of eukaryotic DNA replication. *Proc. Natl. Acad. Sci. USA* **106**, 20240–20245 (2009).
2. Remus, D. *et al.* Concerted loading of Mcm2-7 double hexamers around DNA during DNA replication origin licensing. *Cell* **139**, 719–730 (2009).
3. Botchan, M. & Berger, J. DNA replication: making two forks from one prereplication complex. *Mol. Cell* **40**, 860–861 (2010).
4. Yeeles, J.T., Deegan, T.D., Janska, A., Early, A. & Diffley, J.F. Regulated eukaryotic DNA replication origin firing with purified proteins. *Nature* **519**, 431–435 (2015).
5. Moyer, S.E., Lewis, P.W. & Botchan, M.R. Isolation of the Cdc45/Mcm2-7/GINS (CMG) complex, a candidate for the eukaryotic DNA replication fork helicase. *Proc. Natl. Acad. Sci. USA* **103**, 10236–10241 (2006).
6. Ilves, I., Petojevic, T., Pesavento, J.J. & Botchan, M.R. Activation of the MCM2-7 helicase by association with Cdc45 and GINS proteins. *Mol. Cell* **37**, 247–258 (2010).
7. Makarova, K.S., Koonin, E.V. & Kelman, Z. The CMG (CDC45/RecJ, MCM, GINS) complex is a conserved component of the DNA replication system in all archaea and eukaryotes. *Biol. Direct* **7**, 7 (2012).
8. Slaymaker, I.M. & Chen, X.S. MCM structure and mechanics: what we have learned from archaeal MCM. *Subcell. Biochem.* **62**, 89–111 (2012).
9. Brewster, A.S. *et al.* Crystal structure of a near-full-length archaeal MCM: functional insights for an AAA+ hexameric helicase. *Proc. Natl. Acad. Sci. USA* **105**, 20191–20196 (2008).
10. Enemark, E.J. & Joshua-Tor, L. Mechanism of DNA translocation in a replicative hexameric helicase. *Nature* **442**, 270–275 (2006).
11. Singleton, M.R., Sawaya, M.R., Ellenberger, T. & Wigley, D.B. Crystal structure of T7 gene 4 ring helicase indicates a mechanism for sequential hydrolysis of nucleotides. *Cell* **101**, 589–600 (2000).
12. Miller, J.M., Arachea, B.T., Epling, L.B. & Enemark, E.J. Analysis of the crystal structure of an active MCM hexamer. *eLife* **3**, e03433 (2014).
13. Li, N. *et al.* Structure of the eukaryotic MCM complex at 3.8 Å. *Nature* **524**, 186–191 (2015).
14. Costa, A. *et al.* DNA binding polarity, dimerization, and ATPase ring remodeling in the CMG helicase of the eukaryotic replisome. *eLife* **3**, e03273 (2014).
15. Costa, A. *et al.* The structural basis for MCM2–7 helicase activation by GINS and Cdc45. *Nat. Struct. Mol. Biol.* **18**, 471–477 (2011).
16. Sun, J. *et al.* The architecture of a eukaryotic replisome. *Nat. Struct. Mol. Biol.* **22**, 976–982 (2015).
17. Lyubimov, A.Y., Strycharska, M. & Berger, J.M. The nuts and bolts of ring-translocase structure and mechanism. *Curr. Opin. Struct. Biol.* **21**, 240–248 (2011).
18. Bochman, M.L. & Schwacha, A. The Mcm complex: unwinding the mechanism of a replicative helicase. *Microbiol. Mol. Biol. Rev.* **73**, 652–683 (2009).
19. Itsathitphaisarn, O., Wing, R.A., Eliason, W.K., Wang, J. & Steitz, T.A. The hexameric helicase DnaB adopts a nonplanar conformation during translocation. *Cell* **151**, 267–277 (2012).
20. Fu, Y.V. *et al.* Selective bypass of a lagging strand roadblock by the eukaryotic replicative DNA helicase. *Cell* **146**, 931–941 (2011).
21. Chen, B. & Frank, J. Two promising future developments of cryo-EM: capturing short-lived states and mapping a continuum of states of a macromolecule. *Microscopy (Oxf.)* doi:10.1093/jmicro/dfv344 (31 October 2015).
22. Chang, Y.P., Wang, G., Bermudez, V., Hurwitz, J. & Chen, X.S. Crystal structure of the GINS complex and functional insights into its role in DNA replication. *Proc. Natl. Acad. Sci. USA* **104**, 12685–12690 (2007).
23. Choi, J.M., Lim, H.S., Kim, J.J., Song, O.K. & Cho, Y. Crystal structure of the human GINS complex. *Genes Dev.* **21**, 1316–1321 (2007).

24. Kamada, K., Kubota, Y., Arata, T., Shindo, Y. & Hanaoka, F. Structure of the human GINS complex and its assembly and functional interface in replication initiation. *Nat. Struct. Mol. Biol.* **14**, 388–396 (2007).
25. Froelich, C.A., Kang, S., Epling, L.B., Bell, S.P. & Enemark, E.J. A conserved MCM single-stranded DNA binding element is essential for replication initiation. *eLife* **3**, e01993 (2014).
26. Sun, J. *et al.* Cryo-EM structure of a helicase loading intermediate containing ORC-Cdc6-Cdt1-MCM2-7 bound to DNA. *Nat. Struct. Mol. Biol.* **20**, 944–951 (2013).
27. Samel, S.A. *et al.* A unique DNA entry gate serves for regulated loading of the eukaryotic replicative helicase MCM2-7 onto DNA. *Genes Dev.* **28**, 1653–1666 (2014).
28. Petojevic, T. *et al.* Cdc45 (cell division cycle protein 45) guards the gate of the Eukaryote Replisome helicase stabilizing leading strand engagement. *Proc. Natl. Acad. Sci. USA* **112**, E249–E258 (2015).
29. Rothenberg, E., Trakselis, M.A., Bell, S.D. & Ha, T. MCM forked substrate specificity involves dynamic interaction with the 5'-tail. *J. Biol. Chem.* **282**, 34229–34234 (2007).
30. Simon, A.C. *et al.* A Ctf4 trimer couples the CMG helicase to DNA polymerase  $\alpha$  in the eukaryotic replisome. *Nature* **510**, 293–297 (2014).
31. Onesti, S. & MacNeill, S.A. Structure and evolutionary origins of the CMG complex. *Chromosoma* **122**, 47–53 (2013).
32. Krastanova, I. *et al.* Structural and functional insights into the DNA replication factor Cdc45 reveal an evolutionary relationship to the DHH family of phosphoesterases. *J. Biol. Chem.* **287**, 4121–4128 (2012).
33. Sanchez-Pulido, L. & Ponting, C.P. Cdc45: the missing RecJ ortholog in eukaryotes? *Bioinformatics* **27**, 1885–1888 (2011).
34. Lairson, L.L., Henrissat, B., Davies, G.J. & Withers, S.G. Glycosyltransferases: structures, functions, and mechanisms. *Annu. Rev. Biochem.* **77**, 521–555 (2008).
35. Wakamatsu, T. *et al.* Structure of RecJ exonuclease defines its specificity for single-stranded DNA. *J. Biol. Chem.* **285**, 9762–9769 (2010).
36. Yamagata, A., Kakuta, Y., Masui, R. & Fukuyama, K. The crystal structure of exonuclease RecJ bound to  $Mn^{2+}$  ion suggests how its characteristic motifs are involved in exonuclease activity. *Proc. Natl. Acad. Sci. USA* **99**, 5908–5912 (2002).
37. Niedziela-Majka, A., Chesnik, M.A., Tomko, E.J. & Lohman, T.M. *Bacillus stearothermophilus* PcrA monomer is a single-stranded DNA translocase but not a processive helicase *in vitro*. *J. Biol. Chem.* **282**, 27076–27085 (2007).
38. Velankar, S.S., Soultanas, P., Dillingham, M.S., Subramanya, H.S. & Wigley, D.B. Crystal structures of complexes of PcrA DNA helicase with a DNA substrate indicate an inchworm mechanism. *Cell* **97**, 75–84 (1999).
39. Fischer, C.J., Maluf, N.K. & Lohman, T.M. Mechanism of ATP-dependent translocation of *E. coli* UvrD monomers along single-stranded DNA. *J. Mol. Biol.* **344**, 1287–1309 (2004).
40. Lee, J.Y. & Yang, W. UvrD helicase unwinds DNA one base pair at a time by a two-part power stroke. *Cell* **127**, 1349–1360 (2006).
41. Rad, B. & Kowalczykowski, S.C. Efficient coupling of ATP hydrolysis to translocation by RecQ helicase. *Proc. Natl. Acad. Sci. USA* **109**, 1443–1448 (2012).
42. Bjornson, K.P., Wong, I. & Lohman, T.M. ATP hydrolysis stimulates binding and release of single stranded DNA from alternating subunits of the dimeric *E. coli* Rep helicase: implications for ATP-driven helicase translocation. *J. Mol. Biol.* **263**, 411–422 (1996).
43. Singleton, M.R., Dillingham, M.S. & Wigley, D.B. Structure and mechanism of helicases and nucleic acid translocases. *Annu. Rev. Biochem.* **76**, 23–50 (2007).
44. Eisenmesser, E.Z. *et al.* Intrinsic dynamics of an enzyme underlies catalysis. *Nature* **438**, 117–121 (2005).
45. Frauenfelder, H., Sligar, S.G. & Wolynes, P.G. The energy landscapes and motions of proteins. *Science* **254**, 1598–1603 (1991).
46. Henzler-Wildman, K. & Kern, D. Dynamic personalities of proteins. *Nature* **450**, 964–972 (2007).
47. Lyubimov, A.Y., Costa, A., Bleichert, F., Botchan, M.R. & Berger, J.M. ATP-dependent conformational dynamics underlie the functional asymmetry of the replicative helicase from a minimalist eukaryote. *Proc. Natl. Acad. Sci. USA* **109**, 11999–12004 (2012).
48. Bowman, G.D., O'Donnell, M. & Kuriyan, J. Structural analysis of a eukaryotic sliding DNA clamp-clamp loader complex. *Nature* **429**, 724–730 (2004).
49. Simonetta, K.R. *et al.* The mechanism of ATP-dependent primer-template recognition by a clamp loader complex. *Cell* **137**, 659–671 (2009).
50. Kelch, B.A., Makino, D.L., O'Donnell, M. & Kuriyan, J. How a DNA polymerase clamp loader opens a sliding clamp. *Science* **334**, 1675–1680 (2011).
51. Skordalakes, E. & Berger, J.M. Structure of the Rho transcription terminator: mechanism of mRNA recognition and helicase loading. *Cell* **114**, 135–146 (2003).
52. Moreau, M.J., McGeoch, A.T., Lowe, A.R., Itzhaki, L.S. & Bell, S.D. ATPase site architecture and helicase mechanism of an archaeal MCM. *Mol. Cell* **28**, 304–314 (2007).

## ONLINE METHODS

**Sample preparation and electron microscopy.** *S. cerevisiae* CMG was purified as previously described<sup>16</sup>. To prepare EM grids, we first diluted each sample with 20 mM Tris-acetate, pH 7.5, 40 mM K-glutamate, 2 mM DTT and 0.1 mM EDTA. Before EM-grid preparation of CMG, we verified the sample homogeneity by negative-stain EM. Then we applied 3  $\mu$ l of CMG sample at a final concentration of 0.6 mg/ml to glow-discharged C-flat 1.2/1/3 holey carbon grids, which were then incubated for 10 s at 6 °C and 90% humidity, blotted for 3 s, and plunged into liquid ethane with an FEI Vitrobot IV. We loaded the grids into an FEI Titan Krios electron microscope at 300 keV and collected images automatically in low-dose mode at a magnification of  $\times 29,000$  and a pixel size of 1.01 Å per pixel. A Gatan K2 Summit direct electron detector was used for image recording with an underfocus range from 1.5 to 3.5  $\mu$ m under super-resolution mode. The dose rate was ten electrons per angstroms squared per second, and the total exposure time was 5 s. The total dose was divided into a 25-frame movie, and each frame was exposed for 0.2 s.

**Image processing and 3D reconstruction.** Approximately 8,000 raw movie micrographs were collected. The movie frames were first aligned and superimposed with motioncorr<sup>53</sup>. Contrast-transfer-function parameters of each aligned micrograph were calculated with CTFIND4 (ref. 54). All the remaining steps, including particle autoselection, 2D classification, 3D classification, 3D refinement, and density-map post-processing were performed with Relion-1.3 or Relion-1.4 (refs. 55,56). We manually picked  $\sim 10,000$  particles from different views to generate 2D averages, which were used as templates for subsequent automatic particle selection. Automatic particle selection was then performed for the entire data set. Approximately one million particles were initially selected. Particles were then sorted on the basis of similarity to the 2D references; the 10% of particles with the lowest z scores were deleted from the particle pool. 2D classification of all remaining particles was performed, and particles in unrecognizable classes on the basis of visual inspection were removed. A total of 687,794 particles were used for 3D classification. We derived six 3D models from the data set and found three models that were similar to each other, whose associated particles were combined for further refinement; the other three models were distorted, and those particles were discarded, thus yielding a data-set size of 469,818 particles. This final data set was used for further 3D refinement, resulting in a 3.8-Å 3D density map. The resolution of the map was estimated by the gold-standard Fourier-shell correlation, at a correlation cutoff value of 0.143. The 3D density maps were corrected for the detector modulation transfer function and sharpened by application of a negative B factor of  $-144 \text{ \AA}^2$ . The particles had some preference for end-on views, but because of the large number of particles used, virtually all of the angular space was well sampled. To improve the density of the flexible C-terminal AAA+ ring of Mcm2-7, we performed a 'focused' 3D classification procedure. We first generated two soft-edge masks, one from the Mcm2-7 CTD-ring region and the other from the entire CMG complex. Subtraction of the Mcm2-7 CTD ring mask from the entire CMG complex mask was performed with `relion_image_handler`, from which we obtained a mask that covered only the Mcm2-7 NTD ring, GINS, and Cdc45. This mask was then applied to the 3.8-Å density map. The resulting masked map without the Mcm2-7 CTD ring was used to subtract that region from the original particles, and 3D classification was performed on the Mcm2-7 CTD ring-only particles without further image alignment. Two distinct conformations were found: the CTD ring of conformer I was 'tilted' compared to the NTD ring, whereas conformer II had roughly 'parallel' CTD and NTD rings. These two maps were further refined, thus resulting in two 3D maps with an estimated resolution of 4.7 Å and 4.8 Å, respectively. Estimation of local resolution was calculated with ResMap.

**Model building, refinement, and validation.** The initial model of *S. cerevisiae* Cdc45 was built in Rosetta, on the basis of the extracted density map, with the homolog RecJ as a reference<sup>35,57</sup>. The fold of this model was essentially the same as that in the crystal structure of human Cdc45 (L. Pellegrini, University of Cambridge, personal communication). The initial models of the *S. cerevisiae* GINS subunits were generated from the crystal structure of the human GINS

complex (PDB 2Q9Q)<sup>22</sup> with the SWISS-MODEL server<sup>58</sup>. Six NTDs of the yeast Mcm2-7 were directly extracted from the cryo-EM structure of the yeast Mcm2-7 double hexamer (PDB 3JA8)<sup>13</sup>. These models were docked into the 3.7-Å-resolution 3D density map of the Cdc45-GINS-Mcm2-7 NTD-tier ring in COOT<sup>59</sup> and fitted into the density with Chimera<sup>60</sup>. After assignment of the models, remaining density corresponded to the C-terminal B domain of Psf1, which was absent from the crystal structure. We manually built this domain *de novo* from a polyaniline model. Then the entire atomic model was manually adjusted and built in COOT. Clearly resolved bulky residues such as phenylalanine, tyrosine, tryptophan, and arginine were crucial for sequence registration.

The manually built structure was then refined in real space with phenix.real\_space\_refine<sup>61</sup> and manually adjusted in COOT. The reciprocal-space refinement was also performed in Phenix<sup>62</sup>, with secondary structure and stereochemical constraints applied. The structure factors (including phases) were obtained from the Fourier transform of the experimental density map with `phenix.map_to_structure_factors`. The atomic models were validated with MolProbity and had a percentile score of 95–97%.

The final model was cross-validated with a method described previously<sup>63</sup>. We randomly added 0.1-Å noise to the coordinates of the final model with the PDB tools in Phenix, then refined the noise added against the first half map (Half1) that was produced from one-half of the particle data set during refinement by RELION. We performed one round of coordinate refinement, and this was followed by B-factor refinement. The refined model was then correlated with the 3D maps of the two half maps (Half1 and Half2) in Fourier space to produce two FSC curves: FSC<sub>work</sub> (model versus Half1 map) and FSC<sub>free</sub> (model versus Half2 map), respectively. A third FSC curve was calculated between the refined model and the final 3.7-Å-resolution density map produced from all particles. The general agreement of these curves was taken as an indication that the model was not overfitted.

To build the complete models for CMG conformers I (4.8 Å) and II (4.7 Å), we first used the SWISS-MODEL server to derive two WHD models of MCM5 and MCM6, based on the crystal structure of a GntR family transcriptional regulator (PDB 3C7J) and an NMR structure of WHD of the human MCM6 (PDB 2KLQ)<sup>64</sup>, respectively. We then fitted them along with the above-described built model of Cdc45-GINS-Mcm2-7 N-tier ring, and the six-CT AAA+-domain structure extracted from the cryo-EM structure of yeast Mcm2-7 (PDB 3JA8), in COOT and Chimera. The structures were then refined in Phenix and manually adjusted in COOT. All modeled structures were validated with MolProbity<sup>65</sup>. Structural figures were prepared in Chimera and PyMOL (<http://www.pymol.org/>).

53. Li, X. *et al.* Electron counting and beam-induced motion correction enable near-atomic-resolution single-particle cryo-EM. *Nat. Methods* **10**, 584–590 (2013).
54. Rohou, A. & Grigorieff, N. CTFIND4: fast and accurate defocus estimation from electron micrographs. *J. Struct. Biol.* **192**, 216–221 (2015).
55. Scheres, S.H. Semi-automated selection of cryo-EM particles in RELION-1.3. *J. Struct. Biol.* **189**, 114–122 (2015).
56. Scheres, S.H. RELION: implementation of a Bayesian approach to cryo-EM structure determination. *J. Struct. Biol.* **180**, 519–530 (2012).
57. Wang, R.Y. *et al.* *De novo* protein structure determination from near-atomic-resolution cryo-EM maps. *Nat. Methods* **12**, 335–338 (2015).
58. Biasini, M. *et al.* SWISS-MODEL: modelling protein tertiary and quaternary structure using evolutionary information. *Nucleic Acids Res.* **42**, W252–W258 (2014).
59. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132 (2004).
60. Pettersen, E.F. *et al.* UCSF Chimera: a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
61. Afonine, P.V. *et al.* Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr. D Biol. Crystallogr.* **68**, 352–367 (2012).
62. Adams, P.D. *et al.* PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221 (2010).
63. Amunts, A. *et al.* Structure of the yeast mitochondrial large ribosomal subunit. *Science* **343**, 1485–1489 (2014).
64. Wei, Z. *et al.* Characterization and structure determination of the Cdt1 binding domain of human minichromosome maintenance (Mcm) 6. *J. Biol. Chem.* **285**, 12469–12473 (2010).
65. Chen, V.B. *et al.* MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 12–21 (2010).