

ARTICLE

Received 10 Aug 2015 | Accepted 12 Jan 2016 | Published 18 Feb 2016

DOI: 10.1038/ncomms10708

OPEN

# Cryo-EM structures of the eukaryotic replicative helicase bound to a translocation substrate

Ferdos Abid Ali<sup>1,\*</sup>, Ludovic Renault<sup>1,2,\*</sup>, Julian Gannon<sup>1,†</sup>, Hailey L. Gahlon<sup>3</sup>, Abhay Kotecha<sup>4</sup>, Jin Chuan Zhou<sup>1</sup>, David Rueda<sup>3</sup> & Alessandro Costa<sup>1</sup>

The Cdc45-MCM-GINS (CMG) helicase unwinds DNA during the elongation step of eukaryotic genome duplication and this process depends on the MCM ATPase function. Whether CMG translocation occurs on single- or double-stranded DNA and how ATP hydrolysis drives DNA unwinding remain open questions. Here we use cryo-electron microscopy to describe two subnanometre resolution structures of the CMG helicase trapped on a DNA fork. In the predominant state, the ring-shaped C-terminal ATPase of MCM is compact and contacts single-stranded DNA, via a set of pre-sensor 1 hairpins that spiral around the translocation substrate. In the second state, the ATPase module is relaxed and apparently substrate free, while DNA intimately contacts the downstream amino-terminal tier of the MCM motor ring. These results, supported by single-molecule FRET measurements, lead us to suggest a replication fork unwinding mechanism whereby the N-terminal and AAA + tiers of the MCM work in concert to translocate on single-stranded DNA.

<sup>1</sup>Macromolecular Machines, Clare Hall Laboratory, The Francis Crick Institute, Blanche Lane, South Mimms EN6 3LD, UK. <sup>2</sup>National Institute for Biological Standards and Control, Microscopy and Imaging, Blanche Lane, South Mimms EN6 3QG, UK. <sup>3</sup>Section of Virology and Single Molecule Imaging Group, Department of Medicine, MRC Clinical Centre, Imperial College London, London W12 0NN, UK. <sup>4</sup>Division of Structural Biology, Wellcome Trust Centre for Human Genetics, University of Oxford, Roosevelt Drive, Oxford OX3 7BN, UK. \* These authors contributed equally to this work. † Present address: Synthetic and Systems Biochemistry of the Microtubule Cytoskeleton, Lincoln's Inn Fields Laboratory, The Francis Crick Institute, 44 Lincoln's Inn Fields, London WC2A 3LY, UK. Correspondence and requests for materials should be addressed to A.C. (email: [alessandro.costa@crick.ac.uk](mailto:alessandro.costa@crick.ac.uk)).

**D**NA replication onset requires an initiator that loads a set of two helicases for double-helix unwinding. This provides the single-stranded DNA template for the replicative polymerases. In eukaryotic cells, helicase recruitment and origin activation are temporally separated<sup>1</sup>. The origin recognition complex partakes in loading an inactive dimer of ring-shaped MCM helicase motors that encircle double-stranded DNA<sup>2–4</sup>. Origin firing depends on the recruitment of a set of replication factors<sup>5</sup>, including the GINS and Cdc45 activators that bind to each MCM ring in the dimer, forming a pair of multisubunit Cdc45-MCM-GINS (CMG) holo-helicases<sup>6,7</sup>. Upon origin activation, the two CMG particles are believed to separate and move in opposite directions to unwind DNA<sup>8</sup>, however, the molecular basis of MCM double-ring uncoupling is unknown. The MCM helicase motor is a ring-shaped hetero-hexamer containing six homologous polypeptides belonging to the superfamily of AAA + ATPases. The N-terminal domain (NTD) of the MCM forms a DNA-binding collar and a co-axial carboxy-terminal ATPase motor powers substrate translocation through the ring central channel<sup>9</sup>. Whether DNA unwinding involves MCM motor translocation on duplex- or single-stranded DNA remains unclear<sup>4,10–13</sup>. DNA fork progression depends on the ATPase function of the MCM motor<sup>5,14</sup>, however, it is unknown how the energy derived from ATP hydrolysis is converted into motion and fork unwinding<sup>15</sup>. To start to address these outstanding questions, we have determined two cryo-electron microscopy (cryo-EM) structures of the CMG helicase trapped on a model DNA fork (by incubation with the slowly hydrolysable ATP analogue, ATP $\gamma$ S). We have also obtained two similar basic structures of the CMG helicase in the absence of DNA, imaged in conditions that allow for ATP turnover. Combined with single-molecule FRET analysis of DNA deformation by the CMG, our data provide important novel insights into the mechanism of replication fork progression in eukaryotic cells.

### Subnanometre resolution structure of the CMG

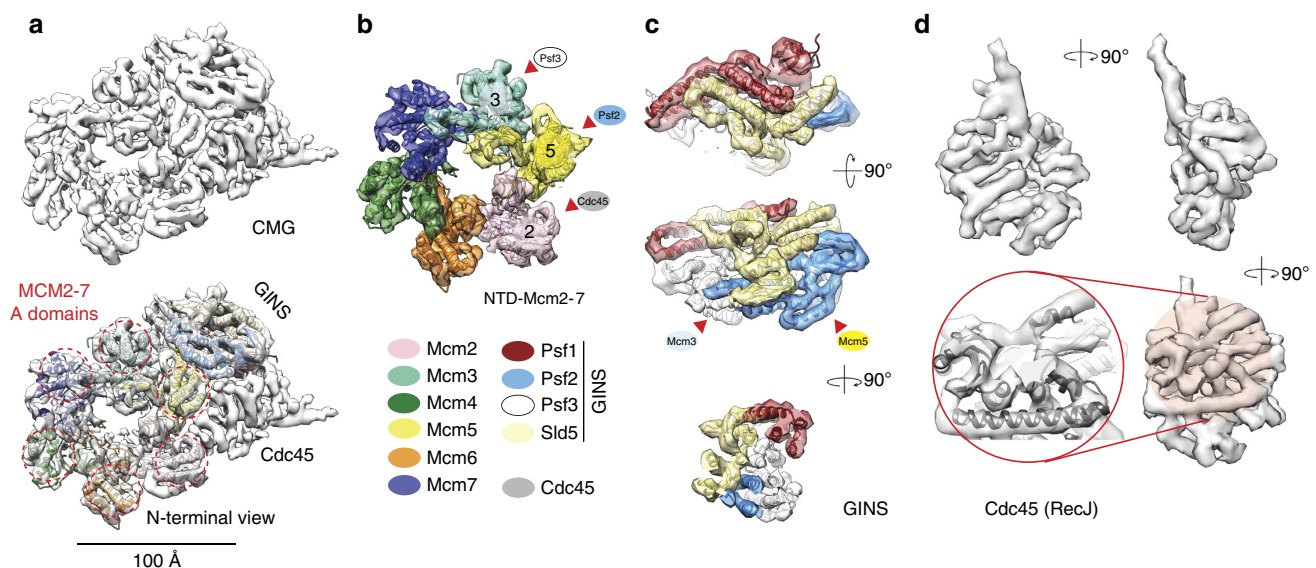
Catalytically active, baculovirus-expressed *Drosophila melanogaster* CMG was incubated with a model replication fork in the presence

of ATP $\gamma$ S, required for stable DNA binding<sup>6</sup>. Particles embedded in vitrified ice were imaged on a FEI Polara electron microscope equipped with an energy filter and a K2 Summit direct electron detector (Gatan, Inc.; Supplementary Fig. 1). Following two-dimensional (2D) and three-dimensional (3D) classification, a first structure was refined to 7.4 Å resolution (Supplementary Fig. 2). Atomic docking was employed to interpret the cryo-EM map, using the coordinates of known holo-helicase components. These efforts provide an exhaustive description of the CMG intersubunit interaction network. The structure contains a closed hexameric ring face that matches the N-terminal DNA-interacting collar of yeast MCM<sup>4</sup> (Fig. 1a,b, PDB entry 3JA8), albeit with significant inter-domain rearrangements (Supplementary Fig. 3 and Supplementary Movie 1). Combined with previous subunit mapping studies<sup>16,17</sup>, our data confirm that GINS components Psf2 and Psf3 (PDB entry 2Q9Q) interact with the outer perimeter of MCM subunits 5 and 3 (Fig. 1a–c). Remarkably, Psf2  $\alpha$ -helices 3 and 5 (as defined in the human GINS structure<sup>18</sup>) contact a region of the Mcm5 N-terminal ‘A domain’ that is protected by the N-terminal extension of the MCM subunit 7 from the opposing ring in the double hexamer, as described in the atomic resolution yeast structure<sup>4</sup> (Supplementary Fig. 4).

As previously proposed<sup>16,17,19</sup>, unoccupied density mapping next to GINS is assigned to Cdc45 and indeed matches the secondary structure elements of RecJ<sup>20</sup> (PDB entry 1IR6), a distant prokaryotic exonuclease homologue (Fig. 1d and Supplementary Fig. 5)<sup>21,22</sup>. In addition, a C-terminal protrusion projects from the RecJ-like catalytically defunct core of Cdc45 and wedges in between the Mcm5 and Mcm2 A domains<sup>16</sup>, as previously observed at low resolution<sup>19</sup> (Figs 1a and 2a,b).

### AAA + ATPase in the activated MCM helicase motor

DNA unwinding by hexameric helicases involves the ATP hydrolysis-driven reconfiguration of the ATPase centres, causing allosteric changes that promote substrate translocation<sup>23–25</sup>. Seeking to identify various intermediates in the DNA unwinding reaction, we refined all intact CMG 3D classes and identified two similar, higher-resolution CMG forms that differ only in their ATPase module. The first form has been introduced



**Figure 1 | CMG helicase structure at subnanometre resolution.** (a) Resolution density (7.4 Å) map of the CMG viewed from the MCM N-terminal face, without or with docked MCM and GINS atomic structures. (b) Detailed view of the MCM N-terminal DNA-interacting collar. Psf3 contacts Mcm3, Psf2 contacts Mcm5 and Cdc45 contacts Mcm2. (c) Detailed view of the GINS assembly with docked human atomic structure (PDB entry 2Q9Q). (d) Density assigned to Cdc45. The Cdc45 core matches the secondary structure elements of the bacterial RecJ exonuclease (PDB entry 1IR6).

before, while a second form was refined to 9.8 Å resolution (Fig. 2a,b, Supplementary Figs 2,6 and 7, and Supplementary Movie 2). Inspection of the two helicase structures reveals that the ATPase tier is found as either topologically closed ('compact ATPase') or notched ('relaxed ATPase'). To define the interdomain arrangement in the helicase motor, we compared the recently established atomic model of yeast (double-hexameric) MCM<sup>4</sup> with our newly determined CMG structures. The compact ATPase configuration is highly reminiscent of the yeast MCM, with a closed ATPase ring and nearly parallel NTD and AAA + tiers. In this configuration, MCM protomers are aligned and NTD interacts with AAA + in *cis*, with a slight left-handed twist<sup>4</sup>, confirming a previous model proposed by us<sup>16,19</sup> and others<sup>17</sup> (Fig. 2b). Conversely, in the relaxed ATPase form, the GINS-Cdc45 elements wedge in between the NTD and AAA + tiers, a configuration previously observed at low resolution (Fig. 2a) (ref. 19). In this form, individual AAA + domains become rearranged with respect to the cognate NTDs, approximating a conformer captured in a recent crystallographic study on an archaeal MCM hexamer<sup>26</sup>. DNA density can be recognized in both CMG structures and the nature of the DNA contacts will be discussed below.

### Remodelling of the ATPase motor

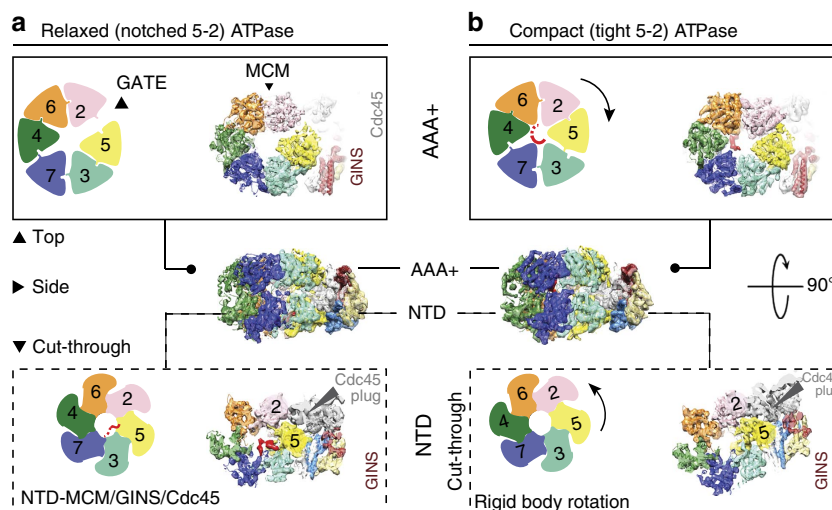
To accurately describe ATPase motor dynamics, the yeast Mcm2-7 AAA + hexamer was fitted into the CMG structure and further energy minimization was allowed for the individual docked ATPase modules into the two refined 3D classes. This docking exercise highlights a remarkable plasticity of the Mcm2-7 protomers, with relative tilting and rocking of the AAA + domains, occurring both within the six Mcm2-7 ring subunits and between the two CMG conformers (Figs 2 and 3a). As previously proposed<sup>16,19</sup>, we can assign a discontinuity in the AAA + domain to the Mcm5-2 ATPase interface (Fig. 2a,b). By comparing the compact and relaxed structures, a conformational transition can be modelled in the ATPase tier. This reveals two types of ATPase interfaces that behave differently. Movement mainly depends on the loosening and tightening of the Mcm5-2 interface (root mean squared deviation (r.m.s.d.) 8.2 Å between the open and closed configurations; Fig. 3b,c). Mcm5-2 loosening is chaperoned by compensatory tightening of the neighbouring Mcm3-5 and 2-6 interfaces (r.m.s.d.: 5.3 and 8.3 Å, respectively;

Fig. 3b,c), while the rest of the subunits undergo more subtle reconfigurations (with Mcm4-7 and Mcm7-3 r.m.s.d.: 3.1 and 3.3 Å, respectively; Fig. 3b,c). Coherent with these asymmetric rearrangements, *in vitro* studies of the *Drosophila* CMG have shown that certain Mcm2-7 ATPase sites (found at the static hinge in our structures) can be inactivated with minimal effects on DNA unwinding by the CMG, while other sites (the active Mcm5-2 ATPase and its immediate neighbors) are required for translocation<sup>6</sup> (Fig. 3c).

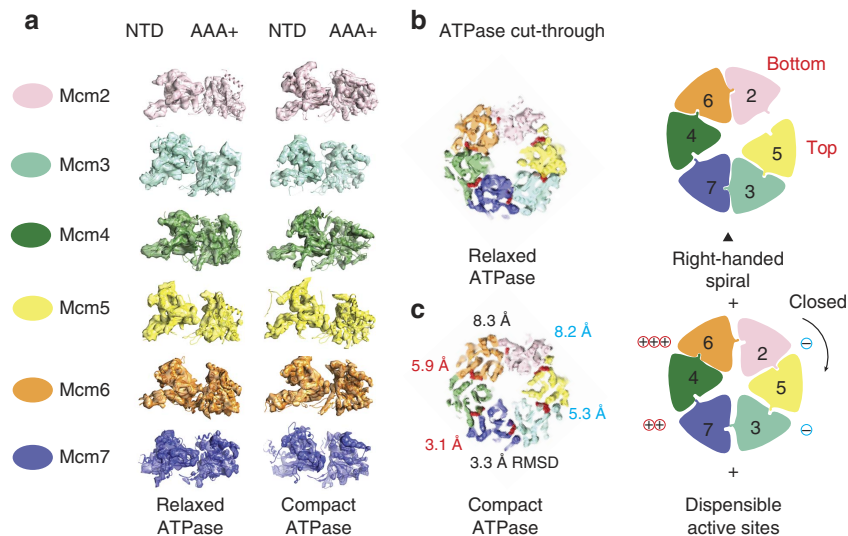
We have previously observed a compact Mcm5-2 interface at low resolution for a CMG in the presence of a non-hydrolysable ATP analogue (with or without DNA), while the nucleotide-free CMG contained a notched 5-2 interface<sup>16,19</sup>. These data support the notion that the relaxed ATPase form observed in our new, subnanometre resolution structure might contain an empty (post hydrolysis) ATPase site at the 5-2 interface. In line with this notion, the compact ATPase state is the predominant form (74% of particles that reach subnanometre resolution) in the ATP $\gamma$ S-CMG-DNA preparation, while the relaxed structure is underrepresented (the remaining 26%), as would be expected when using a slowly hydrolysable ATP analogue. To test whether ATP hydrolysis affects the particle distribution of the two conformers, we have solved the cryo-EM structure of the CMG incubated with ATP (hydrolysis rate is unaffected by the presence or absence of DNA<sup>6</sup>; imaging performed on a Titan Krios electron microscope, equipped with a Falcon II direct electron detector, FEI). Three-dimensional classification led to the identification of two CMG forms highly similar to the ATP $\gamma$ S/DNA-bound structures (although DNA free; Supplementary Figs 8–10 and Supplementary Movies 2 and 3). Remarkably, when the ATP-treated CMG is imaged, the notched Mcm5-2 structure is more represented than the compact state (72% versus 28%), supporting the notion that ATP hydrolysis by the Mcm5-2 site could convert the compact ATPase into a relaxed configuration (Supplementary Fig. 11). Although we cannot rule out the possibility that one of the two forms might be an inactive state of the helicase<sup>12</sup>, it is possible that the two structures described here represent two intermediates of the DNA translocation reaction.

### The CMG helicase is a single-stranded DNA translocase

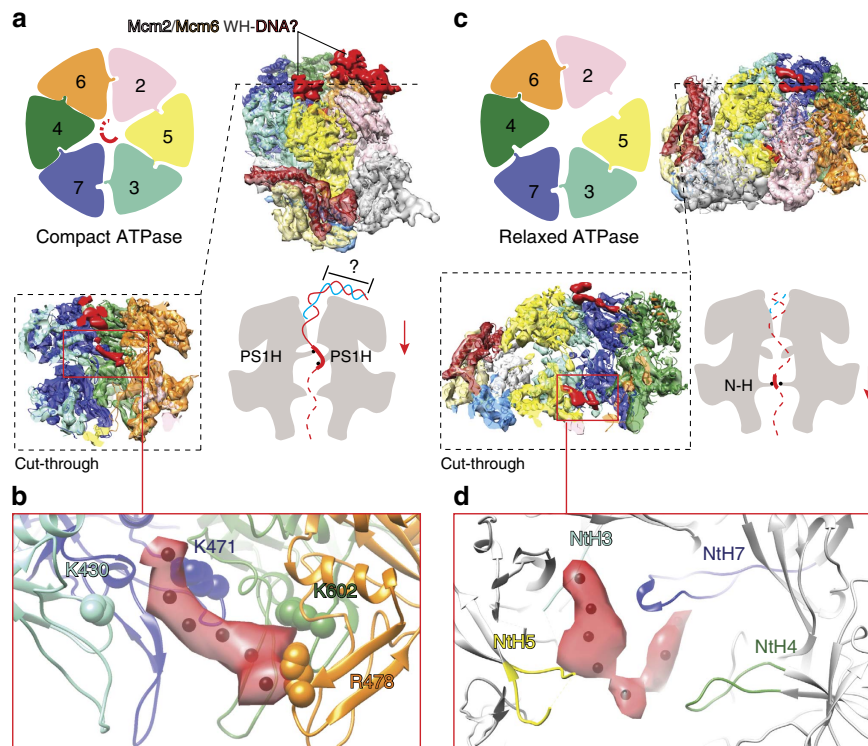
Comparison of the compact ATPase DNA-CMG structure with the DNA-free counterpart reveals the presence of apparently bent, rod-shaped density features contacting the C-terminal face



**Figure 2 | Two configurations in the ATP $\gamma$ S-CMG-DNA complex.** (a) AAA +, side and cut-through view of the CMG in a relaxed ATPase configuration. Cdc45 topologically locks the Mcm5-2 gate. (b) AAA +, side and cut-through N-terminal view of the CMG in a compact ATPase configuration. DNA density surmounting the AAA + domain has been removed for visualization purposes (also see Fig. 4).



**Figure 3 | The MCM ATPase centres.** (a) Segmented density and docked atomic structures of the six Mcm2-7 protomers in the two conformers. The NTD and AAA + domains tilt and rock with respect to one another. (b) ATPase sites in the relaxed ATPase conformer. (c) ATPase sites in the compact ATPase conformer. One side of the MCM ring is more static and the corresponding active sites are dispensable for viability (– means Walker A mutation kills DNA unwinding, + means tolerated mutation).



**Figure 4 | DNA-bound form of the CMG helicase.** (a) The compact ATPase form contains rod-shaped, bent density features surmounting the ATPase face tentatively assigned to duplex DNA engaged by the MCM winged helix (WH) C-terminal extensions. An MCM slice through the side view reveals an extended density feature, which we assign to single-stranded DNA, traversing the AAA + pore. (b) Single-stranded DNA contacts conserved positively charged residues on the AAA + PS1H hairpins that have a key role in DNA unwinding. (c) The relaxed ATPase form contains a thin density feature surmounting the AAA + ring, which we assign to the flexible C-terminal MCM WH extensions. Although the AAA + tier appears substrate free, a well-resolved elongated density threads through the MCM N-terminal collar and (d) contacts a set of N-terminal hairpins important for DNA binding and helicase activity.

of Mcm2-6 and surmounting the pore of the MCM motor (Fig. 4a and Supplementary Fig. 11). The local resolution is low in this region of the cryo-EM map (compared with the core protein components) and not suitable for atomic docking

(Supplementary Figs 6 and 7). Nonetheless, the dimensions and overall shape of the additional density feature matches that of duplex DNA, possibly engaged by one or more C-terminal MCM extensions (containing a winged helix fold)<sup>4</sup>. The notion that

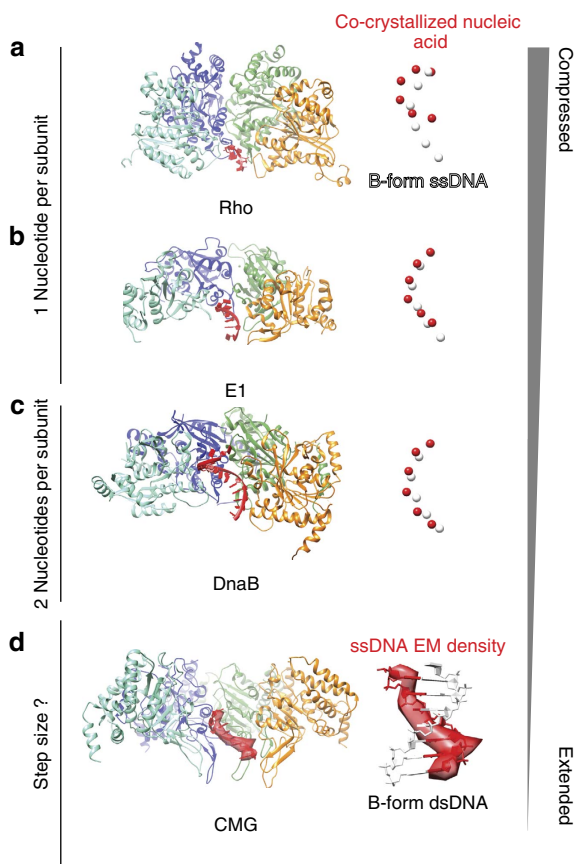


duplex DNA enters through the AAA +, and not the N-terminal tier of MCM, is in agreement with our earlier streptavidin labelling of the DNA fork<sup>19</sup>. The inner core of the AAA + domain is better defined and shows continuous density spanning three MCM subunits (Mcm7-4-6) and contacting the pre-sensor 1 (PS1) and helix-2 insert pore loops (Fig. 4b). This density can accommodate single- but not double-stranded DNA (Fig. 5), indicating that the AAA + domain of the CMG helicase is a single-stranded DNA translocase (Supplementary Movie 3). A set of positively charged residues on the PS1 hairpins (Mcm7 K471, Mcm4 K602 and Mcm6 R478) contact and spiral around the nucleic acid (Fig. 4b). These DNA-binding elements are highly conserved from archaea through eukarya and their mutation affects DNA binding and abrogate DNA unwinding<sup>11,12,27</sup>. Noticeably, Mcm3 K430 is also poised in close proximity to the DNA density (Fig. 4b). This is a known DNA translocation element that is essential for viability in yeast<sup>27</sup>.

Comparison of the relaxed ATPase structures in the presence or absence of DNA shows more subtle differences (Supplementary Fig. 11). In both structures, thin density features

surmount the C-terminal face of the AAA + pore and probably correspond to DNA-free winged helix MCM appendices<sup>4</sup>, raising the question of whether DNA occupancy in this state is lower than in the compact ATPase state. Remarkably, although the AAA + pore loops appear DNA free, a second set of N-terminal pore loops encircle an elongated density feature visible in the nucleic-acid-treated CMG but not in the DNA-free CMG. Thus, we assign the N-terminal contacting density to DNA. Similar to the PS1 hairpins, these ‘N-terminal hairpins’ are also conserved and are essential for DNA binding and have a role in unwinding<sup>28–30</sup> (Fig. 4d). It remains to be established whether the compact and relaxed DNA-binding modes in the ring channel correspond to a translocation and a paused state<sup>12</sup>, respectively, or represent two intermediates in the DNA translocation reaction.

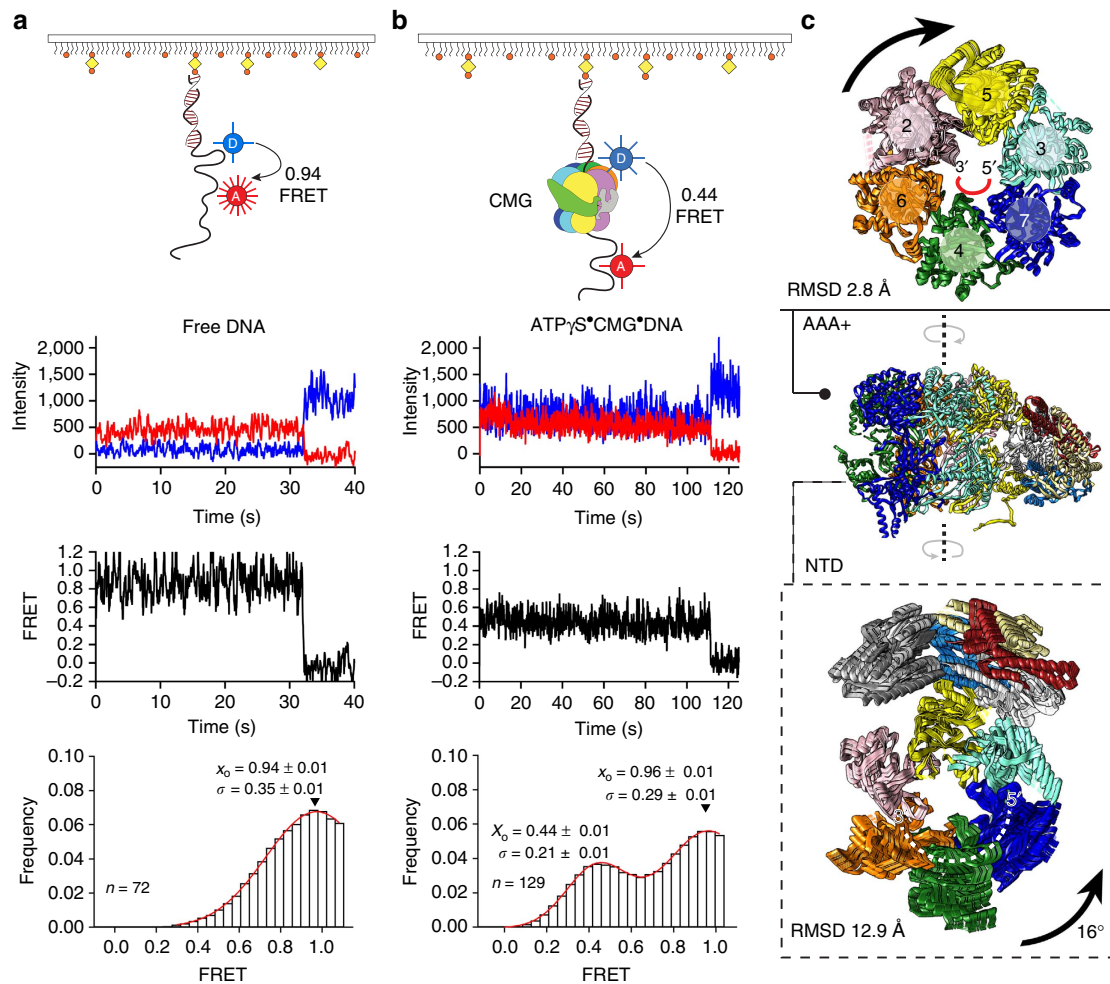
Comparison of the DNA-engaged MCM ATPase tier with other structures of nucleic acid-bound hexameric helicases highlights a mixture of similarities and differences (Fig. 5). A universal feature appears to be the presence of a right-handed spiral formed by the ATPase pore loops, which follows the helical character of the nucleic acid substrate<sup>15</sup>. Whereas in previously reported helicase structures (Rho, E1 and, to a lesser extent, DnaB) the translocation substrate is more compressed than that of canonical B-form DNA<sup>24,31,32</sup> (Fig. 5a–c), the ATP $\gamma$ S–CMG–DNA complex matches the structure of one single strand extracted from a B-form double helix (Fig. 5d). Noticeably, the Rho and E1 helicases, which significantly compress the translocation substrate as they encircle it, contact one base per ATPase protomer, while DnaB contacts two bases per protomer. These observations are compatible with an unwinding step size of 1- (Rho and E1) and 2-bp (DnaB) per ATP hydrolysed. Although the resolution of our compact ATP $\gamma$ S–CMG–DNA complex is not high enough to count bases in the single-stranded nucleic acid, the more extended state of the DNA substrate suggests that the CMG might unwind at least two base pairs per ATP molecule hydrolysed. To confirm the notion that single-stranded DNA is stabilized in an extended configuration on CMG binding, we performed single-molecule experiments, in the presence of either ATP $\gamma$ S or ATP. CMG–DNA binding was monitored by measuring FRET efficiency between a Cy3 donor and a Cy5 acceptor fluorophore, separated by seven nucleotides in surface-immobilized single-stranded DNA. This is an established method to measure DNA stretching by multi-subunit, DNA-encircling ATPases<sup>33</sup> (Fig. 6a,b). CMG binding to DNA in the presence of ATP $\gamma$ S results in a dramatic decrease from  $0.94 \pm 0.01$  (naked DNA; Fig. 6a) to  $0.44 \pm 0.01$  (DNA + CMG; Fig. 6b and Supplementary Fig. 12) in mean FRET ( $x_0$ ), due to both a reduction in molecular flexibility relative to naked single-stranded DNA and CMG-mediated single-stranded DNA stretching. Similar results were obtained with ATP ( $x_0 = 0.96 \pm 0.01$  for naked DNA and  $x_0 = 0.48 \pm 0.01$  for CMG), although binding/stretching events occurred with lower frequency (Supplementary Fig. 12). These data agree with the observation that use of a slowly hydrolysable ATP analogue is required to stabilize the interaction between the MCM AAA + domain and single-stranded DNA<sup>6,12</sup>, and support the notion that CMG binding stabilizes and stretches single-stranded DNA. We note that the lower frequency of DNA binding/stretching observed when CMG is ATP incubated correlates with the predominance of the relaxed ATPase form observed in the ATP–CMG cryo-EM data set.



**Figure 5 | Comparison of the ATP $\gamma$ S–CMG–DNA structure with available helicase–nucleic acid assemblies.** (a) The Rho termination factor (PDB entry 3ICE) contacts and compresses single-stranded RNA. The RNA structure is compared with one single DNA strand extracted from a B-form double helix. Spheres represent the centre of mass of each nucleotide. (b) The Papillomavirus E1 replicative helicase (PDB entry 2GXA) contacts and compresses single-stranded DNA. (c) The bacterial DnaB replicative helicase contacts and slightly compresses single-stranded DNA (approximating A-form DNA, PDB entry 4ESV). (d) The MCM motor of the CMG harbours a more extended form of single-stranded DNA, which appears to match B-form DNA. The DNA density matches the structure of one single strand in the double helix. This supports the notion that the CMG is a single-stranded DNA translocase.

#### ATPase controlled rotation of the CMG DNA-binding domain

How does the ATPase state of the MCM motor affect the rest of the CMG structure? To address this question we have aligned the AAA + rings of the two ATP–CMG conformers and used



**Figure 6 | DNA engagement and deformation by the CMG helicase.** Single-molecule FRET analysis of (a) isolated DNA and (b) DNA + ATP $\gamma$ S CMG. Top to bottom: cartoon schematic of the single-molecule FRET experiment indicating FRET between Cy3 (Donor, blue sphere) and Cy5 (Acceptor, red sphere, 7-nt separation) attached to biotinylated 3'-tailed duplex DNA, immobilized on a neutravidin-coated biotin-PEG surface; donor (blue) and acceptor (red) intensity trajectories are anti-correlated until single-step photobleaching of the acceptor. FRET trajectories (black) between Cy3 and Cy5 exhibit a sharp drop to zero FRET when the acceptor photobleaches; histogram of FRET values collected from all molecules with a Gaussian fit are shown in red, where  $x_0$  is the mean FRET and  $\sigma$  is the distribution width. ATPase controlled ring rotation and DNA stretching by the CMG helicase. (c) The AAA + tier of the Mcm2-7 motor rotates clockwise as Mcm2 moves towards the Mcm5 protomer, to close a gap in the motor domain. The DNA-interacting NTD rotates anticlockwise as a rigid body, together with GINS and Cdc45.

molecular morphing to model a transition from a relaxed to a compact ATPase conformation. In our model, the motor undergoes minor subunit rearrangements, with an overall r.m.s.d. of 2.8 Å, mainly due to the breathing of the Mcm5-2 interface (Fig. 6c). Conversely, the non-catalytic portion of the CMG complex, formed by Cdc45, NTD-Mcm2-7 and GINS, rotates by 16° as a rigid body with respect to the motor (overall r.m.s.d. of 12.9 Å between the two forms; Figs 2a,b and 6c, and Supplementary Movie 4). Remarkably, the NTD and AAA + tiers of Mcm2-7 rotate in opposite directions, as they transition from a relaxed to a compact ATPase form (Fig. 6c). This rotation is compatible with a nucleotide-state-controlled inter-subunit movement, previously observed using double electron-electron resonance in the archaeal MCM motor<sup>29</sup>.

Previous crystallographic studies in archaea have shown that single-stranded DNA can line the NTD-MCM pore. This interaction (preserved in the eukaryotic Mcm6-4-7) is essential for DNA replication initiation in yeast and probably important for the initial melting of the origin DNA duplex (Fig. 6c). According to our model, as the ATPase motor tightens to engage the translocation strand, the AAA + domain rotates clockwise,

while the NTD DNA-interacting collar of the MCM rotates anticlockwise. Considering that DNA interacts with the inner perimeter of NTD-MCM running 3' → 5' in an anticlockwise manner<sup>30</sup>, we note that a further anticlockwise NTD rotation would result in DNA duplex underwinding, indeed compatible with origin DNA opening.

## Discussion

DNA replication start sites need to be licensed for initiation and this process involves the deposition of a pair of MCM helicases that form a two-fold symmetric, double-hexameric assembly<sup>2,3</sup>. GINS/Cdc45 recruitment is necessary for origin firing and multiple lines of evidence suggest that two CMG holo-helicases separate and move in opposite directions during replication. For example, single-molecule studies have demonstrated that two replisomes need not be physically linked for efficient DNA replication to occur<sup>8</sup>. Further evidence of helicase separation is derived from EM studies on the CMG, which primarily form monomeric assemblies<sup>16</sup> (although loosely tethered dimeric CMGs have also been observed)<sup>19</sup>. A recent high-resolution

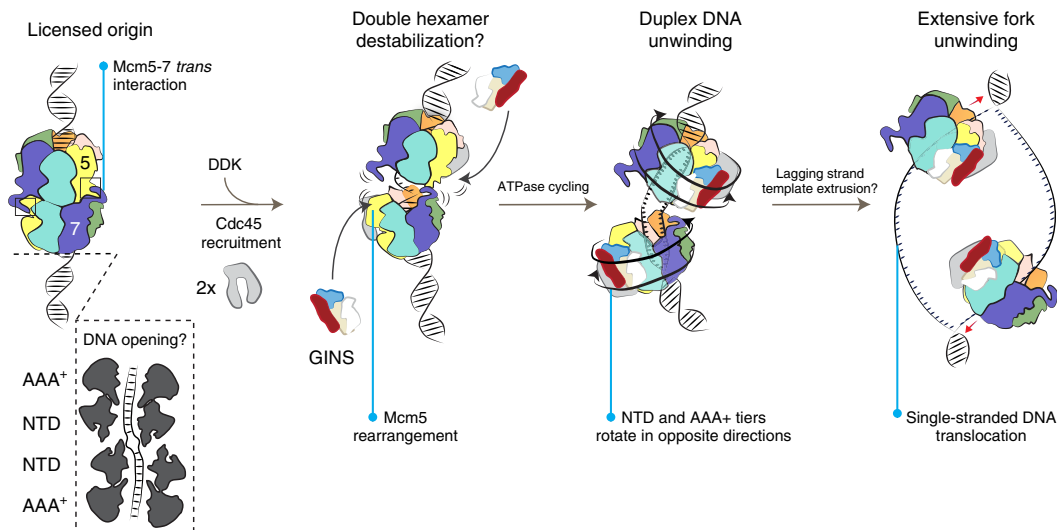
study describes the dimerization interface of the yeast MCM double ring, which is in part formed by an N-terminal Mcm7 helical insert that latches onto the Mcm5 A domain in the opposing hexamer<sup>4</sup>. Our subnanometre resolution CMG structure reveals that the A domain of Mcm5 undergoes a reconfiguration in the CMG (Supplementary Movie 1), which would be incompatible with Mcm7 engagement (at least in yeast). Remarkably, the highly conserved Mcm5 site that is generally protected by Mcm7 in the double hexamer is found GINS associated in the CMG (Supplementary Fig. 4). GINS recruitment depends on the DDK phosphorylation of Mcm4, 6 and 7 (refs 5,34); however, DDK requirement in yeast can be bypassed by a mutation of Mcm5 (ref. 35) that promotes a reconfiguration of the Mcm5 A domain<sup>28</sup>. DDK phosphorylation probably promotes a similar conformational change in Mcm5 that could disrupt the (yeast specific) Mcm5-7 double hexamer interface and enable GINS binding. These events, we suggest, might help destabilize the double hexamer and contribute to replication initiation (Fig. 7).

Our new CMG structures better explain the role of Cdc45, containing a globular extension emanating from the RecJ-like core. This extension inserts in between the A domains of Mcm5 and 2 (refs 16,19,36), plugging a gate that is generally found open in the isolated MCM complex<sup>16,36</sup> and used for DNA loading<sup>37</sup>. During elongation, ring opening is thought to allow for the bypass of a roadblock on the translocation strand in certain viral replicative helicases<sup>38</sup>, while the eukaryotic replisome stalls on meeting an obstruction on the translocation strand<sup>10,39</sup>. A topologically sealed Mcm2-7 ring (the CMG structure described here) provides a rationale to explain this observation. With Cdc45 plugging the Mcm5-2 gate, full ring opening is energetically disfavoured in the holo-helicase (Fig. 2a,b), possibly explaining helicase stalling.

Whether the CMG helicase is a single-stranded or a duplex DNA translocase is still debated<sup>4,9,10,40</sup>. As the Mcm2-7 is initially loaded onto duplex DNA, replication fork establishment has been suggested to occur via two alternative mechanisms. According to the ‘strand extrusion’ model duplex, DNA enters

the AAA + ATPase tier of MCM and the two strands become spatially segregated before exiting the helicase toroid<sup>9,13</sup>. A second model, called ‘strand exclusion’, envisages translocation on one single strand and steric exclusion of the other strand<sup>9-11</sup> (Supplementary Fig. 13). Recent structural characterization of the yeast MCM double hexamer provides some evidence in support of the strand extrusion model<sup>4</sup>. In the double-ring structure, a unique lateral channel exists between the Mcm6 and 2 subunits, which might be used for extrusion of the lagging strand. Given the position of this side pore, mapping in between the AAA + and NTD tiers, strand extrusion would occur after the double strand has been spooled through the AAA + motor (Supplementary Fig. 13). This hypothesis is supported by the notion that the CMG helicase contains a central pore that is large enough to accommodate double-stranded DNA<sup>16,17,19</sup>. Our new ATP $\gamma$ S-CMG-DNA structure, however, provides robust evidence against a duplex translocase model and favours the steric exclusion mechanism. In fact, in the compact ATPase state we observe thin continuous density that spirals inside the AAA + MCM hexamer and contacts a set of positively charged residues known to be important for DNA binding and translocation. This density can accommodate single- but not double-stranded DNA (Fig. 5d), indicating that the MCM motor unwinds a replication fork by translocating on the leading strand template, whereas the lagging strand template is sterically excluded (Fig. 7). Support for our model derives from studies on *Xenopus* egg extracts, according to which the replisome can bypass a road block on the lagging strand but not on the translocation (leading) strand<sup>10</sup>. Future, high-resolution studies on higher-order assemblies will be needed to define the replication fork trajectory in the eukaryotic replisome<sup>17</sup>.

Three main mechanisms have been proposed for coupling nucleotide turnover with substrate translocation in hexameric NTPases, involving either stochastic, concerted or rotary firing<sup>15</sup>. The stochastic model, mainly supported by studies on protein unfoldases, suggests that the NTPase firing does not need to follow an ordered sequence of events for substrate translocation



**Figure 7 | Origin activation and replication fork unwinding by the CMG helicase.** Origin licensing involves the loading of a head-to-head double hexameric ring that encircles duplex DNA, which might become partially deformed. Cdc45 is loaded onto the double hexamer, in a process that requires DDK phosphorylation of MCM. DDK phosphorylation might cause a rearrangement in Mcm5 and disrupt the (yeast specific) Mcm5-7 *trans* interaction. This rearrangement would expose a GINS interacting element in Mcm5. ATP hydrolysis by the MCM promotes the relative rotation of the NTD and AAA + tiers of the helicase, in a movement that might promote duplex DNA underwinding. Following a poorly understood lagging-strand extrusion process, the CMG helicase extensively unwinds the replication fork, by translocating on single-stranded DNA.



to occur<sup>41,42</sup>. The concerted model envisions six NTPase subunits that simultaneously bind, hydrolyse or release the nucleotide. This mechanism is supported by symmetrized crystallographic structures of the SV40 large T-antigen replicative helicase<sup>23</sup>. According to the rotary mechanism, nucleotide hydrolysis occurs sequentially from one hydrolase centre to the next, similar to a Mexican wave around a stadium<sup>15,24,31</sup>. The rotary cycling model is supported by biochemical and crystallographic studies of the T7 gp4 helicase<sup>43</sup> and by three helicase structures, the Rho transcription termination<sup>24</sup> factor, the Papillomavirus E1 (ref. 31) and the bacterial DnaB replicative helicases<sup>32</sup>, which have all been imaged bound to their translocation substrate (single-stranded nucleic acid). Importantly, in the E1 and Rho translocation–substrate complexes, the hydrolase centres are radially distributed in an order that sequentially visits the steps of the NTPase cycle<sup>24,31</sup>. In both structures, a staircase formed by the ATPase pore loops follows the right-handed spiral of a single-stranded nucleic acid and the position of the pore loops in the assembly correlates with the nucleotide occupancy state of the ATPase centres. An ATPase-controlled rotary pore-loop movement would therefore promote substrate translocation through the ATPase channel. Owing to the homo-oligomeric nature of the E1 helicase complex however, a rotary cycling movement in the helicase hexamer can only be modelled and not visually proven. Because of its inherently asymmetric nature, we reasoned that the Mcm2-7 hetero-hexamers would be a useful tool to test the rotary cycling model. The asymmetric distribution of ATPase centres and the ordered spiral organization of the DNA-interacting PS1 hairpins indeed favour a sequential rotary over a concerted firing or stochastic mechanism for substrate translocation by the CMG. However, to our surprise, we were unable to characterize multiple rotational states of the ATPase, with DNA contacting different ATPase protomers. Rather, the ATPase exists in two states: compact (DNA gripping) and relaxed (DNA ungrapping; Supplementary Fig. 14). Our results lend themselves to two possible explanations. A first possibility is that one of six rotational states of the MCM is more stable than the other, short-lived ATPase ring permutations. However, we favour a second possibility. The ATPase motor could fluctuate between the compact state (where DNA is stabilized/extended by Mcm7-4-6) and the relaxed state where the AAA+ motor releases the substrate. In this second configuration, DNA could be handed off to the N-terminal collar and our observation of stable NTD–DNA association in the relaxed ATPase structure supports this model. Alternatively, the relaxed ATPase configuration could represent a stalled form of the helicase, as could be found, for example, at a pausing replication fork (Supplementary Fig. 14).

Notably, not all ATPase sites in the MCM motor of the CMG equally contribute to DNA translocation<sup>6,44</sup> and this feature is shared by other hetero-hexameric AAA+ ATPases such as the Rpt1-6 assembly<sup>45</sup> of the proteasome or the dynein motor (which in some organisms contains ATPase centres that have become inactivated during evolution)<sup>46</sup>. Both the proteasome and dynein have been imaged in a compact and a relaxed ATPase form, with movement dependent on the breathing of one lone hydrolase interface and chaperoning activity of the neighbouring interfaces<sup>47,48</sup> (such as the Mcm5-2 site and its immediate neighbours in our CMG structure). Although these ATPase motors are functionally very distinct (a DNA, a polypeptide and a microtubule translocase, respectively), phylogenetic<sup>49,50</sup> and mechanochemical<sup>19</sup> kinship has been noticed before for the three systems, suggesting a shared mechanism of substrate translocation for hetero-hexameric AAA+ ATPases. Taken together, our results suggest that the CMG helicase might translocate on single-stranded DNA via a mechanism distinct

from strictly sequential rotary cycling<sup>6,44</sup> and probably involves the interplay between the N-terminal and AAA+ tiers of the MCM motor<sup>29</sup>.

## Methods

**Purification of the CMG helicase.** pFastBac1 plasmids containing genes that encode the individual subunits of the CMG complex (a gift from Dr Michael R. Botchan) were used to produce bacmids, which were subsequently transfected into Sf9 cells using the Invitrogen Bac-to-Bac Baculovirus Expression System methods<sup>6</sup>. P3 baculoviruses used in protein expression experiments were freshly amplified from the P2 stocks for 5 days in 100 ml of Sf9 cells grown in 250 ml flasks in Grace's medium supplemented with 10% FCS. The infections for protein purification were carried out by inoculating 41 of Hi5 cells at  $10^6 \text{ ml}^{-1}$  with a multiplicity of infection of 5 by adding 200 ml of each virus stock at a titre  $10^8 \text{ pfu ml}^{-1}$ . Infected cells incubated for 72 h at 27 °C were harvested by centrifugation and washed with PBS + 5 mM MgCl<sub>2</sub>. The following steps were performed either on ice or at 4 °C, unless indicated otherwise. The collected cells were resuspended in 200 ml of buffer C (25 mM Hepes pH 7.6, 0.02% Tween-20, 10% glycerol, 1 mM EDTA, 1 mM EGTA) supplemented with 15 mM KCl, 2 mM MgCl<sub>2</sub>, 0.4 mM phenylmethylsulfonyl fluoride, 2 mM 2-mercaptoethanol and the complete protease inhibitors cocktail from Roche Diagnostics. The cell suspension was snap frozen in 10 ml aliquots and stored at –80 °C. To purify CMG complexes from the extract, the infected cell suspension was thawed and cells were broken in a Dounce homogenizer. KCl was added to 100 mM and the extract was cleared by centrifugation at 14,000 r.p.m. in an Avanti J-26S XP centrifuge for 10 min. The cleared extract was incubated with 2 ml of anti-FLAG M2-agarose beads (Sigma-Aldrich) for 2–3 h with continuous end-over-end mixing. The beads were then collected in a 20-ml Poly-Prep disposable chromatography column (BioRad) and washed in 30 ml C-100 buffer (buffer C with 100 mM KCl and 1 mM dithiothreitol (DTT)) before transferring to a 10 ml column (BioRad). The column was washed twice with 5 ml C-100 buffer. Bound complexes were eluted with 200 µg ml<sup>-1</sup> flag peptide in C-100 supplemented with complete protease inhibitor cocktail, the first elution step was performed in 5 ml of elution buffer for 15 min at room temperature with end-over-end mixing, then collecting the flowthrough and repeating the step with 4 ml of elution buffer for 10 min. Both fractions were pooled, cooled on ice and pumped through a Mono S HiTrap SP FF column equilibrated in C-100 buffer. The flowthrough and a further 4 ml of wash was collected and injected onto a MonoQ HR 5/5 column equilibrated in C-100 buffer<sup>6</sup>. The column was washed in 15 ml of C-100 and bound complexes were eluted with 20-ml 100–550 mM KCl gradient in buffer C supplemented with 1 mM DTT. Fractions (0.5 ml) were collected, the CMG peak fractions (eluted at 410–440 mM KCl) were diluted to 150 mM KCl in buffer C and injected onto a Mono Q Pur 1.6/5 column connected to AKTAmicro purification system equilibrated in buffer D (25 mM Hepes pH 7.6, 1 mM EDTA, 1 mM EGTA, 1 mM DTT) supplemented with 150 mM KCl. Elution was performed using a 2-ml 150–550 mM KCl linear gradient in buffer D and 75 µl fractions were collected. The CMG peak fractions were pooled and dialysed for 16 h into buffer A (25 mM Hepes pH 7.6, 50 mM sodium acetate, 10 mM magnesium acetate, 1 mM DTT). The CMG preparation was supplemented with 1 mM ATP or treated as described in the 'CMG–DNA complex reconstitution' section. Protein concentration was measured by using known MCM3 protein standards serving as a reference on a silver stain SDS–PAGE gel. A yield of 150 µl with a concentration of 1 µM was achieved.

**Helicase assay.** Assays were performed in 20 µl reactions contacting 25 mM Hepes pH 7.5, 75 mM NaCl, 0.5 mM ATP, 10 mM magnesium acetate, 1 mM DTT, 0.1 mg ml<sup>-1</sup> BSA. Purified CMG (123 to 410 fmol) was incubated in the presence of 27 µM ATPγS and 1.8 pmol of [<sup>32</sup>P]-ATP-labelled fork DNA substrate for 3 h at 30 °C, to allow for fork loading. Unwinding was induced by adding ATP to 7 mM for 5 min at 30 °C and stopped with 6 × stop buffer (150 mM Tris pH 8.0, 3% SDS, 120 mM EDTA). The unwinding product was separated using 8% PAGE in 0.1% SDS 1 × TBE.

**CMG–DNA complex reconstitution.** The oligonucleotides used for CMG–DNA reconstitution were synthesized by Integrated DNA Technology. Sequences of the two partially complementary oligos used are shown below:

Leading strand template: 5'-CACTCGGGCTCGTTTTTACAACGTCGTGACTGGGCACTTGATCGGCCAACCTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTTT TT TTTTTT-3'

Lagging strand template: 5'-CTGCGTCGGTCCGGCGTTGGCCGATCAAGTGCCAGTCACGACGTTGTAA AACGAGCCCGAGTG-3'

When annealed together, this model replication fork contains a 50-bp double-stranded region with two single-stranded fork overhangs (16-nt on the 5'-end and 40-nt poly-T on the 3'-end, based on Petojevic *et al.*<sup>12</sup>).

Annealing of the two strands was performed by mixing each of the oligos in equimolar amounts, heating at 95 °C for 3 min and slow cooling to room temperature for 60 min. To reconstitute the nucleoprotein complex, forked DNA substrate was added to dialysed ~300 nM CMG in a 2:1 molar ratio in the presence of 0.1 mM ATPγS (Sigma). The CMG–DNA–ATPγS mixture was left to incubate for 2 h at 30 °C, to allow for complex association before cryo-grid preparation.



**Cryo-grid preparation and data collection for CMG-DNA.** Four microlitres of reconstituted CMG-DNA assembly at 300 nM concentration were applied onto freshly glow-discharged Quantifoil 1.2/1.3 or C-flat 1/1 grids. After a 30-s incubation in 100% humidity, the Quantifoil grid was double-side blotted for 4 s using a Vitrobot (FEI) and plunged into liquid ethane. For the C-flat grid, the sample was incubated for 2 min and double-side blotted for 3 s in a Cp3 (Gatan, Inc.), operating at 90% humidity and plunged into liquid ethane. Cryo-grids were screened for ice quality using on a JEOL-2100 or a FEI Spirit LaBe operated at 120 kV and equipped with a  $4\text{ k} \times 4\text{ k}$  or  $2\text{ k} \times 2\text{ k}$  Ultrascan charge-coupled device camera (Gatan, Inc.), respectively. Data were collected on a Tecnai F30 Polara electron microscope operated at 300 kV and equipped with a K2 Summit direct electron detector (Gatan, Inc.) and an energy filter in zero-loss mode (GIF Quantum, Gatan, Inc.). Twenty-five-frame movies (2,098) were manually collected, with a single frame duration of 0.4 ms. Movies were acquired using SerialEM, in single-electron counting mode with a total dose of  $48\text{ e}^- \text{ \AA}^{-2}$  and a  $-3.5$  to  $1.7$  defocus range, at a constant nominal magnification of  $\times 37,037$ , yielding a  $1.35\text{-\AA}$  pixel size.

**Cryo-grid preparation and data collection for ATP-CMG.** The ATP-CMG sample at 330 nM concentration was applied onto freshly glow-discharged open-holes Quantifoil 1.2/1.3. After 30 s incubation, grids were double-side blotted for 5 s in a Vitrobot (FEI) at 100% humidity and plunge frozen into liquefied ethane. Grids were loaded onto the LMB Cambridge Titan Krios electron microscope (FEI) operated at 300 kV for automated data collection with the EPU software (FEI). Images were recorded on a FEI Falcon II detector at a nominal magnification of  $\times 47,000$  (yielding a pixel size of  $1.77\text{ \AA}$ ). An in-house built system described in ref. 51 was used to collect 17 frames per second. Five hundred and thirty-six movies of 51 frames were recorded using a  $-2$  to  $-4\text{ \mu m}$  defocus range with an electron dose of  $51\text{ e}^- \text{ \AA}^{-2}$ .

**CMG-DNA image processing.** To correct for beam-induced drift, whole frame alignment and averaging was performed for each movie using MotionCorr (<http://cryoem.ucsf.edu/software/driftcorr.html>)<sup>52</sup>. Particles (340,573) were picked semi-automatically in EMAN2 (ref. 53). Contrast transfer function parameters were estimated using CTFIND4 (ref. 54) and low-quality integrated movies were excluded. All further processing was performed in RELION 1.4 (ref. 55). Extracted particles were binned by 2, yielding a pixel size of  $2.7\text{ \AA}$  per pixel. Two-dimensional classification allowed for the isolation of a set of 60,287 high-quality CMG particles (Supplementary Fig. 2). A first 3D refinement was performed using a  $50\text{-\AA}$  low-pass filtered compact ATP-CMG structure as a starting model, resulting in an initial  $8.1\text{ \AA}$  structure. Postprocessing (using automatically estimated B factor and an arbitrarily chosen mask) improved the resolution to  $7.2\text{ \AA}$  but resulted in a poorly defined ATPase tier (Supplementary Fig. 2). To separate various ATPase conformers in the data set, the refined 3D volume was filtered to  $50\text{ \AA}$  and used as a model for 3D classification without particle alignment (ten classes). After excluding poor-quality classes, 2 structures were isolated, referred to as 'compact ATPase' (38,792 particles; Supplementary Fig. 6) and 'relaxed ATPase' (13,692 particles; Supplementary Fig. 7, also refer to Figs 2 and 4, and Supplementary Fig. 2 for comparisons). Each particle subset was separately refined and postprocessed, resulting in a  $7.4\text{-\AA}$  resolution structure (compact ATPase) and  $9.8\text{ \AA}$  (relaxed ATPase).

**ATP-CMG image processing.** Similar to the CMG-DNA data set, beam-induced drift was corrected and averaging was performed for each movie as described in ref. 52. Contrast transfer function parameters were estimated using CTFIND3 (ref. 56) and the best 471 out of 536 integrated movies were selected for particle picking in XMIPP3 (ref. 57) and subsequent processing in RELION 1.3 (ref. 55). Semi-manually picked particles (160,401) were 2D-classified, to get rid of bad particles, resulting in a clean data set containing 78,601 particles. A first ATP-CMG 3D volume was determined to  $9.5\text{ \AA}$ , using a  $50\text{-\AA}$ -filtered version of EMDB entry 2,772 as an initial model (Supplementary Fig. 8). Using the  $9.5\text{-\AA}$  structure filtered to  $50\text{ \AA}$ , 3D classification was performed to identify four recognizable CMG classes<sup>58</sup> as the new, '3D-cleaned' data set. Using statistical movie processing<sup>59</sup>, an  $8.3\text{-\AA}$  reconstruction (Supplementary Fig. 8) was then obtained from these merged (29,772) particles. As the resulting structure showed a disordered AAA + motor domain, further focused 3D classification<sup>55,58</sup> was performed to identify different motor conformers. This resulted in two classes (compact ATPase, 5,111 particles and relaxed ATPase, 13,182 particles), which were independently refined and postprocessed as described above (Supplementary Figs 8–10).

**Resolution estimation and model building.** Resolution of the obtained cryo-EM maps was estimated using the 'gold-standard' Fourier Shell Correlation (FSC) method, using the 0.143 FSC criterion. Local resolution was estimated using ResMap<sup>60</sup> and visualized using UCSF Chimera<sup>61</sup>. Automated atomic docking was performed using UCSF Chimera and further model manipulation performed using The PyMOL Molecular Graphics System. The A and B-C domains of the NTD tier of the yeast Mcm2-7 structure (PDB entry 3JA8 (ref. 4)) were docked as isolated rigid bodies. GINS (PDB entry 2Q9Q<sup>18</sup>) and RecJ (PDB entry 1IR6 (ref. 20)) were docked as rigid bodies. The C-terminal Psf1 domain GINS was modelled as

described in ref. 19. The AAA + MCM domains were docked as six individual rigid bodies after docking of the cryo-EM MCM AAA + tier (PDB entry 3JA8 (ref. 4)). FSC to compare the CMG atomic models and the cryo-EM maps were computed using RELION 1.4 and visualized with the PDBe FSC server. Figures and Movies were generated using UCSF Chimera<sup>61</sup>.

**Single-molecule FRET.** DNA oligonucleotides were purchased from Operon and labelled with fluorescent dyes Cy3 and Cy5. Reactions were carried out using an amino-modified C6-dT oligonucleotide and a mono-reactive Cy3 or Cy5 dye (GE Healthcare). Labelling reactions were performed with 1 nmol of DNA in 44  $\mu\text{l}$  of 100 mM sodium carbonate buffer pH 8.5 and 10 nmol of Cy3 or Cy5 dye dissolved in 7  $\mu\text{l}$  of dimethyl sulfoxide. Reactions were performed overnight at room temperature. Reverse-phase HPLC purification was performed on an analytical C8-column (Sigma-Aldrich Supelco Discovery BIO wide pore C8,  $25\text{ cm} \times 4.6\text{ mm} \times 5\text{ \mu m}$ ), to separate labelled and unlabelled DNA; fractions containing labelled DNA were collected and stored at  $-20\text{ }^\circ\text{C}$  in 10 mM Tris-HCl pH 8.0. DNA sequences are as follows; for the Cy3 strand 5'-Cy3-CGCGAGGA ATGGATGTAGGG-biotin-3' and for the Cy5 strand 5'-CCCTACATCCA TTCCTCGCGTTTTT (Cy5-T)(T)<sub>65</sub>-3'. Quartz slides and cover slips were prepared following established protocols<sup>62</sup>. Briefly, quartz slides and coverslips were passivated with methoxy-PEG-SVA ( $M_r = 5,000$ , Laysan Bio, Inc.) containing 10% biotin-PEG-SVA ( $M_r = 3,400$ , Laysan Bio, Inc.) in 100 mM sodium bicarbonate. Reaction chambers were first incubated with  $0.2\text{ mg ml}^{-1}$  BSA (Sigma-Aldrich) in T50 buffer (10 mM Tris-HCl pH 7.0 and 50 mM NaCl) for 10 min. Next, BSA was washed with T50 buffer and neutravidin ( $0.2\text{ mg ml}^{-1}$  in T50 buffer) was injected and incubated for 10 min. Excess neutravidin was removed by washing with buffer E (25 mM Hepes pH 7.6, 50 mM sodium acetate, 10 mM magnesium acetate, 10% glycerol, 1 mM DTT and 2 mM Trolox). DNA was surface-immobilized by incubation for 10 min with an annealed biotinylated DNA duplex (25 pM). Excess DNA was then washed with imaging buffer containing the CMG protein (Buffer E supplemented with  $50\text{ \mu M}$  ATP $\gamma$ S (Sigma-Aldrich) or ATP (Sigma-Aldrich), 7 nM CMG protein and an oxygen scavenging system containing 10 mM 3,4-dihydroxybenzoic acid (Sigma-Aldrich) and 120 nM protocatechuate dioxygenase (Sigma-Aldrich; the final concentration of protocatechuate dioxygenase is corrected for the presence of 40% stabilizer), to minimize dye photobleaching). DNA molecules were imaged on a home-built, prism-based total internal reflection fluorescence microscope. All single-molecule measurements were recorded at room temperature using continuous green excitation (532 nm laser) at  $\sim 1.0\text{ mW}$  and 30 ms time resolution. Apparent FRET efficiencies were calculated as  $\text{FRET} = I_A / (I_D + I_A)$ , where  $I_A$  and  $I_D$  are the acceptor and donor intensities, respectively. Acceptor intensity ( $I_A$ ) was corrected for donor emission in the acceptor channel (11%), no direct excitation of the acceptor was observed.

## References

- Remus, D. & Diffley, J. F. Eukaryotic DNA replication control: lock and load, then fire. *Curr. Opin. Cell Biol.* **21**, 771–777 (2009).
- Remus, D. *et al.* Concerted loading of Mcm2-7 double hexamers around DNA during DNA replication origin licensing. *Cell* **139**, 719–730 (2009).
- Evvin, 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).
- Li, N. *et al.* Structure of the eukaryotic MCM complex at 3.8 Å. *Nature* **524**, 186–191 (2015).
- 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).
- 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).
- 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).
- Yardimci, H., Loveland, A. B., Habuchi, S., van Oijen, A. M. & Walter, J. C. Uncoupling of sister replisomes during eukaryotic DNA replication. *Mol. Cell* **40**, 834–840 (2010).
- Costa, A. & Onesti, S. Structural biology of MCM helicases. *Crit. Rev. Biochem. Mol. Biol.* **44**, 326–342 (2009).
- Fu, Y. V. *et al.* Selective bypass of a lagging strand roadblock by the eukaryotic replicative DNA helicase. *Cell* **146**, 931–941 (2011).
- McGeoch, A. T., Trakselis, M. A., Laskey, R. A. & Bell, S. D. Organization of the archaeal MCM complex on DNA and implications for the helicase mechanism. *Nat. Struct. Mol. Biol.* **12**, 756–762 (2005).
- Petojevic, T. *et al.* Cdc45 (cell division cycle protein 45) guards the gate of the eukaryotic replisome helicase stabilizing leading strand engagement. *Proc. Natl Acad. Sci. USA* **112**, E249–E258 (2015).
- Takahashi, T. S., Wigley, D. B. & Walter, J. C. Pumps, paradoxes and ploughshares: mechanism of the MCM2-7 DNA helicase. *Trends Biochem. Sci.* **30**, 437–444 (2005).

14. Heller, R. C. *et al.* Eukaryotic origin-dependent DNA replication *in vitro* reveals sequential action of DDK and S-CDK kinases. *Cell* **146**, 80–91 (2011).
15. 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).
16. Costa, A. *et al.* The structural basis for MCM2-7 helicase activation by GINS and Cdc45. *Nat. Struct. Mol. Biol.* **18**, 471–477 (2011).
17. Sun, J. *et al.* The architecture of a eukaryotic replisome. *Nat. Struct. Mol. Biol.* **22**, 976–982 (2015).
18. 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).
19. Costa, A. *et al.* DNA binding polarity, dimerization, and ATPase ring remodeling in the CMG helicase of the eukaryotic replisome. *Elife* **3**, e03273 (2014).
20. Yamagata, A., Kakuta, Y., Masui, R. & Fukuyama, K. The crystal structure of exonuclease RecJ bound to Mn<sup>2+</sup> ion suggests how its characteristic motifs are involved in exonuclease activity. *Proc. Natl Acad. Sci. USA* **99**, 5908–5912 (2002).
21. 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).
22. Sanchez-Pulido, L., Diffley, J. F. & Ponting, C. P. Homology explains the functional similarities of Treslin/Ticrr and Sld3. *Curr. Biol.* **20**, R509–R510 (2010).
23. Gai, D., Zhao, R., Li, D., Finkielstein, C. V. & Chen, X. S. Mechanisms of conformational change for a replicative hexameric helicase of SV40 large tumor antigen. *Cell* **119**, 47–60 (2004).
24. Thomsen, N. D. & Berger, J. M. Running in reverse: the structural basis for translocation polarity in hexameric helicases. *Cell* **139**, 523–534 (2009).
25. Enemark, E. J., Chen, G., Vaughn, D. E., Stenlund, A. & Joshua-Tor, L. Crystal structure of the DNA binding domain of the replication initiation protein E1 from papillomavirus. *Mol. Cell* **6**, 149–158 (2000).
26. 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).
27. Lam, S. K. *et al.* The PSI hairpin of Mcm3 is essential for viability and for DNA unwinding *in vitro*. *PLoS ONE* **8**, e82177 (2013).
28. Fletcher, R. J. *et al.* The structure and function of MCM from archaeal *M. thermoautotrophicum*. *Nat. Struct. Biol.* **10**, 160–167 (2003).
29. Barry, E. R., Lovett, J. E., Costa, A., Lea, S. M. & Bell, S. D. Intersubunit allosteric communication mediated by a conserved loop in the MCM helicase. *Proc. Natl Acad. Sci. USA* **106**, 1051–1056 (2009).
30. 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).
31. Enemark, E. J. & Joshua-Tor, L. Mechanism of DNA translocation in a replicative hexameric helicase. *Nature* **442**, 270–275 (2006).
32. Itsathithaisarn, 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).
33. Taylor, M. R. *et al.* Rad51 paralogs remodel pre-synaptic Rad51 filaments to stimulate homologous recombination. *Cell* **162**, 271–286 (2015).
34. Labib, K. How do Cdc7 and cyclin-dependent kinases trigger the initiation of chromosome replication in eukaryotic cells? *Genes Dev.* **24**, 1208–1219 (2010).
35. Hardy, C. F., Dryga, O., Seematter, S., Pahl, P. M. & Sclafani, R. A. mcm5/cdc46-bob1 bypasses the requirement for the S phase activator Cdc7p. *Proc. Natl Acad. Sci. USA* **94**, 3151–3155 (1997).
36. 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).
37. 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).
38. Yardimci, H. *et al.* Bypass of a protein barrier by a replicative DNA helicase. *Nature* **492**, 205–209 (2012).
39. Duxin, J. P., Dewar, J. M., Yardimci, H. & Walter, J. C. Repair of a DNA-protein crosslink by replication-coupled proteolysis. *Cell* **159**, 346–357 (2014).
40. Takahashi, K., Yamada, H. & Yanagida, M. Fission yeast minichromosome loss mutants mis cause lethal aneuploidy and replication abnormality. *Mol. Biol. Cell* **5**, 1145–1158 (1994).
41. Martin, A., Baker, T. A. & Sauer, R. T. Rebuilt AAA + motors reveal operating principles for ATP-fuelled machines. *Nature* **437**, 1115–1120 (2005).
42. Yakamavich, J. A., Baker, T. A. & Sauer, R. T. Asymmetric nucleotide transactions of the HslUV protease. *J. Mol. Biol.* **380**, 946–957 (2008).
43. 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).
44. 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).
45. Beckwith, R., Estrin, E., Worden, E. J. & Martin, A. Reconstitution of the 26S proteasome reveals functional asymmetries in its AAA + unfoldase. *Nat. Struct. Mol. Biol.* **20**, 1164–1172 (2013).
46. Roberts, A. J., Kon, T., Knight, P. J., Sutoh, K. & Burgess, S. A. Functions and mechanics of dynein motor proteins. *Nat. Rev. Mol. Cell. Biol.* **14**, 713–726 (2013).
47. Schmidt, H., Gleave, E. S. & Carter, A. P. Insights into dynein motor domain function from a 3.3-Å crystal structure. *Nat. Struct. Mol. Biol.* **19**, 492–497, S1 (2012).
48. Matyskiela, M. E., Lander, G. C. & Martin, A. Conformational switching of the 26S proteasome enables substrate degradation. *Nat. Struct. Mol. Biol.* **20**, 781–788 (2013).
49. Erzberger, J. P. & Berger, J. M. Evolutionary relationships and structural mechanisms of AAA + proteins. *Annu. Rev. Biophys. Biomol. Struct.* **35**, 93–114 (2006).
50. Gleave, E. S., Schmidt, H. & Carter, A. P. A structural analysis of the AAA + domains in *Saccharomyces cerevisiae* cytoplasmic dynein. *J. Struct. Biol.* **186**, 367–375 (2014).
51. Bai, X. C., Fernandez, I. S., McMullan, G. & Scheres, S. H. Ribosome structures to near-atomic resolution from thirty thousand cryo-EM particles. *Elife* **2**, e00461 (2013).
52. 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).
53. Tang, G. *et al.* EMAN2: an extensible image processing suite for electron microscopy. *J. Struct. Biol.* **157**, 38–46 (2007).
54. Rohou, A. & Grigorieff, N. CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *J. Struct. Biol.* **192**, 216–221 (2015).
55. Scheres, S. H. RELION: implementation of a Bayesian approach to cryo-EM structure determination. *J. Struct. Biol.* **180**, 519–530 (2012).
56. Mindell, J. A. & Grigorieff, N. Accurate determination of local defocus and specimen tilt in electron microscopy. *J. Struct. Biol.* **142**, 334–347 (2003).
57. Sorzano, C. O. *et al.* XMIPP: a new generation of an open-source image processing package for electron microscopy. *J. Struct. Biol.* **148**, 194–204 (2004).
58. Amunts, A., Brown, A., Toots, J., Scheres, S. H. & Ramakrishnan, V. Ribosome. The structure of the human mitochondrial ribosome. *Science* **348**, 95–98 (2015).
59. Scheres, S. H. Beam-induced motion correction for sub-megadalton cryo-EM particles. *Elife* **3**, e03665 (2014).
60. Kucukelbir, A., Sigworth, F. J. & Tagare, H. D. Quantifying the local resolution of cryo-EM density maps. *Nat. Methods* **11**, 63–65 (2014).
61. Pettersen, E. F. *et al.* UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* **25**, 1605–1612 (2004).
62. Lamichhane, R., Solem, A., Black, W. & Rueda, D. Single-molecule FRET of protein-nucleic acid and protein-protein complexes: surface passivation and immobilization. *Methods* **52**, 192–200 (2010).

## Acknowledgements

We thank Shaoxia Chen and Christos Savva (LMB Cambridge) for assistance with automatic data collection on the Titan Krios. We also thank Michael Botchan for the gift of the CMG pFastBacs and for useful discussion; Kristy MacLellan-Gibson (NIBSC), Raffaella Carzaniga and Lucy Collinson (The Crick) for technical support with cryo screening; and Bik Tye and Ning Gao for sharing EMDDB and PDB data before release. We thank Ivar Ilves and James Berger for interminable conversations; Hasan Yardimci for helpful discussion, advise and support with the DNA unwinding reaction; Peter Cherepanov, Giulia Zanetti and Adelina Davies for critical reading of the manuscript; and the members of the Costa lab for useful comments. This work was supported by The Francis Crick Institute (which receives its core funding from Cancer Research UK, the UK Medical Research Council and the Wellcome Trust, A.C.) and a PhD fellowship from the Boehringer Ingelheim Fonds (to F.A.A.). A.K. is supported by the Wellcome Trust. The Division of Structural Biology—Particle Imaging Center Electron Microscopy Facility at University of Oxford was founded by The Wellcome Trust JIF award (060208/Z/00/Z) and is supported by a WT equipment grant (093305/Z/10/Z). The Wellcome Trust, MRC and BBSRC also supports the National EM facility, which enabled provision of the K2 detector at Oxford. The D.R. laboratory is funded by the Clinical Sciences Center of the Medical Research Council (RCUK MC-A658-5TY10) and a startup grant from Imperial College London. H.L.G. is funded by a postdoctoral fellowship from the Swiss National Science Foundation.

## Author contributions

F.A.A. and J.G. purified the CMG complex. F.A.A. reconstituted the CMG–DNA complex. L.R., F.A.A. and A.C. prepared the cryo-EM grids. F.A.A., L.R., A.K. and A.C. performed cryo-EM data collection. F.A.A. and L.R. performed the cryo-EM reconstruction. H.G. and F.A.A. performed single-molecule FRET experiments supervised by D.R. F.A.A. and A.C. wrote the manuscript with inputs from all authors.

**Additional information**

**Accession codes:** ATP-CMG: compact ATPase EMDB entry EMD-3320, relaxed ATPase EMDB entry EMD-3321. ATP<sub>γ</sub>S-CMG-DNA: compact ATPase EMDB entry EMD-3318, relaxed ATPase EMDB entry EMD-3319.

**Supplementary Information** accompanies this paper at <http://www.nature.com/naturecommunications>

**Competing financial interests:** The authors declare no competing financial interests.

**Reprints and permission** information is available online at <http://npg.nature.com/reprintsandpermissions/>

**How to cite this article:** Abid Ali, F. *et al.* Cryo-EM structures of the eukaryotic replicative helicase bound to a translocation substrate. *Nat. Commun.* 7:10708 doi: 10.1038/ncomms10708 (2016).



This work is licensed under a Creative Commons Attribution 4.0 International License. The images or other third party material in this article are included in the article's Creative Commons license, unless indicated otherwise in the credit line; if the material is not included under the Creative Commons license, users will need to obtain permission from the license holder to reproduce the material. To view a copy of this license, visit <http://creativecommons.org/licenses/by/4.0/>