CMG–Pol epsilon dynamics suggests a mechanism for the establishment of leading-strand synthesis in the eukaryotic replisome

Jin Chuan Zhou[a,1], Agnieszka Janska[b,1], Panchali Goswami[a], Ludovic Renault[a,2], Ferdos Abid Ali[a], Abhay Kotecha[c], John F. X. Diffley[b], and Alessandro Costa[a,3]

[a]Macromolecular Machines Laboratory, The Francis Crick Institute, London, NW1 1AT, United Kingdom; [b]Chromosome Replication Laboratory, The Francis Crick Institute, London, NW1 1AT, United Kingdom; and [c]Division of Structural Biology, Wellcome Trust Centre for Human Genetics, University of Oxford, Oxford, OX3 7BN, United Kingdom

Edited by Thomas Kelly, Memorial Sloan-Kettering Cancer Center, New York, NY, and approved March 6, 2017 (received for review January 13, 2017)

The replisome unwinds and synthesizes DNA for genome duplication. In eukaryotes, the Cdc45–MCM–GINS (CMG) helicase and the leading-strand polymerase, Pol epsilon, form a stable assembly. The mechanism for coupling DNA unwinding with synthesis is starting to be elucidated, however the architecture and dynamics of the replication fork remain only partially understood, preventing a molecular understanding of chromosome replication. To address this issue, we conducted a systematic single-particle EM study on multiple permutations of the reconstituted CMG–Pol epsilon assembly. Pol epsilon contains two flexibly tethered lobes. The noncatalytic lobe is anchored to the motor of the helicase, whereas the polymerization domain extends toward the side of the helicase. We observe two alternate configurations of the DNA synthesis domain in the CMG-bound Pol epsilon. We propose that this conformational switch might control DNA template engagement and release, modulating replisome progression.

DNA replication | CMG helicase | DNA polymerase | single-particle electron microscopy

DNA replication is catalyzed by the replisome, a molecular machine that coordinates DNA unwinding and synthesis (1). These two functions must be tightly coordinated to prevent the rise of genome instability, which is a major cause of cancer. DNA unwinding by a replicative helicase involves single-strand translocation of a hexameric motor, whereas DNA synthesis requires template priming by a primase and extension by dedicated replicative DNA polymerases (2). In eukaryotes, the helicase function is performed by the Cdc45–MCM–GINS (CMG) complex (3, 4) and the primase function is played by Pol alpha (5), whereas DNA synthesis is catalyzed by two specialized DNA polymerases, Pol epsilon and delta. According to the consensus view, Pol epsilon synthesizes the leading and Pol delta the lagging strand (6–11). However, recent studies indicate that the division of labor between replicative polymerases might be more promiscuous than originally thought (12, 13). In in vitro-reconstituted DNA replication reactions, Pol delta can support leading-strand duplication (11, 14), but switching from Pol delta to epsilon is necessary for efficient establishment of leading-strand synthesis (14). The mechanism of substrate handoff between the two polymerases is currently unknown.

Recent breakthroughs in structural biology begin to provide an architectural framework to understand the interaction between helicase and polymerases at the replication fork. For example, studies on the CMG helicase and its subcomplexes have established that the MCM is a six-member ring with an N-terminal domain that serves as a processivity collar (15) and a C-terminal ATPase motor domain that provides the DNA unwinding function (16–21). High-resolution cryo-EM analysis has shown that the ATPase motor translocates on the leading-strand template (22), in agreement with work on *Xenopus* embryo extracts (23). The GINS and Cdc45 components of the CMG bind to the side and stabilize the N-terminal domain of the MCM ring (closing a dynamic Mcm5-2 gate), allowing for the motor to move on DNA (16, 20, 22).

Although previous work from us and others established that Pol alpha maps in proximity to the N-terminal face of the MCM ring (24, 25), recent data on the reconstituted yeast replisome indicate that Pol epsilon stably anchors onto the ATPase face of the helicase (24). Therefore, MCM not only functions as the motor that catalyzes fork progression but also is a central nexus around which the replication machinery is organized.

Pol epsilon is a heterotetramer consisting of two main modules with distinct functions. The first module is the Pol2 N-terminal catalytic domain, which is dispensable for viability (26, 27). The second module is the noncatalytic portion of the assembly comprising the essential Pol2 C-terminal domain (a catalytically defunct polymerase), Dpb2 (a defunct exonuclease), and the Dpb3 and Dpb4 ancillary factors (28–33). Coordinated action of the CMG and Pol epsilon supports leading-strand synthesis (14, 34–36), and emerging evidence indicates that the noncatalytic module of Pol epsilon plays a separate role in replication, being essential for CMG formation in cells and perhaps by stimulating the DNA unwinding function of the CMG helicase (14, 28, 33, 37, 38).

Significance

Faithful and efficient genome duplication is essential for the propagation of life. Aberrant DNA replication can lead to genomic instability and cancer. In eukaryotes, the replication machinery is composed of the DNA-unwinding enzyme Cdc45–MCM–GINS (CMG) and dedicated DNA synthesis factors. Three different polymerases act sequentially on the leading-strand template to establish DNA replication. We describe the architecture and dynamics of the main leading-strand polymerase bound to the CMG helicase, and we propose a mechanism for the establishment of efficient leading-strand synthesis. Our findings provide important insights into how the eukaryotic replication machinery functions to ensure that genome integrity is maintained during replication.

Author contributions: J.C.Z., A.J., J.F.X.D., and A.C. designed research; J.C.Z., A.J., P.G., L.R., and A.K. performed research; F.A.A. contributed new reagents/analytic tools; J.C.Z., P.G., L.R., and A.C. analyzed data; and J.C.Z. and A.C. wrote the paper.

The authors declare no conflict of interest.

This article is a PNAS Direct Submission.

Freely available online through the PNAS open access option.

Data deposition: The EM 3D maps have been deposited in the EMDataBank, www.emdatabank.org [EMD-3642 (CMGE), EMD-3643 (CMGE ΔCat.), and EMD-3644 (CMG)].

[1]J.C.Z. and A.J. contributed equally to this work.

[2]Present address: Netherlands Centre for Electron Nanoscopy (NeCEN), Gorlaeus Laboratory, Leiden 2333, the Netherlands.

[3]To whom correspondence should be addressed. Email: alessandro.costa@crick.ac.uk.

This article contains supporting information online at www.pnas.org/lookup/suppl/doi:10.1073/pnas.1700530114/-/DCSupplemental.

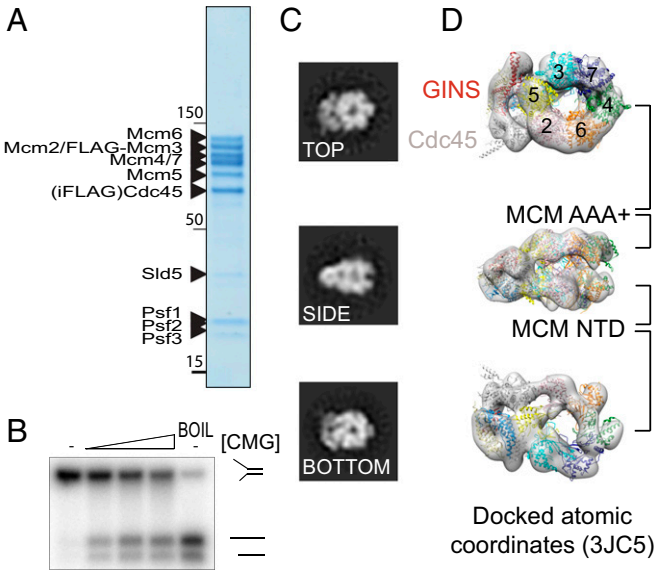

Fig. 1. Catalytically active yeast CMG. (*A*) Coomassie-stained gel of the yeast CMG obtained from strain yJCZ2. (*B*) A DNA unwinding assay shows that the overexpressed yeast CMG is catalytically active. CMG quantities used in the reactions are (from left to right) 0, 159, 397, and 793 fmol. (*C*) 2D image analysis showing that our CMG preparations contain stable CMG particles, suitable for 3D reconstruction. (*D*) 3D reconstruction with atomic docking of the 11 different CMG subunits (PDB ID code 3JC5) showing that the purified CMG is very similar to the published yeast and *Drosophila* CMG complexes.

To understand how Pol epsilon and the CMG work together to duplicate the leading strand, we have reconstituted the intact CMG–Pol epsilon assembly for electron microscopy analysis. We describe here the complete structure of the helicase-leading strand polymerase complex. Using a combination of subunit dropout, domain deletion, and MBP fusion mutants, we can orient Pol epsilon with respect to the CMG helicase and define unexpected architectural features in the eukaryotic replisome. We uncover a conformational change of the DNA synthesis domain of Pol epsilon that (*i*) suggests a structural mechanism for the polymerase switch important during the establishment of leading-strand synthesis and (*ii*) provides a first insight into how replisome processivity might be regulated.

Results

Pure, Active, and Homogeneous Yeast CMG from a Diploid Overexpression Strain. Catalytically active CMG has been previously purified from (*i*) *Drosophila melanogaster* embryo extracts (3), (*ii*) baculovirus-infected insect cells overexpressing fly or human proteins (4, 34), and (*iii*) a haploid *Saccharomyces cerevisiae* strain overexpressing yeast proteins (11). Some of these methods can be tedious, and preparation yields are often variable. Purification of a reconstituted, stoichiometric helicase–polymerase assembly, however, requires a reproducible approach to isolate large quantities of the CMG complex. To this end, we have devised a reliable strategy to produce yeast CMG, built on the established integration of plasmids bearing two codon-optimized genes under the control of a bidirectional galactose-inducible promoter (39) (Fig. S1*A*). Coexpressing as many as 11 different genes with this system is challenging due to the limited choice of selection markers required for plasmid integration. To circumvent this problem, we have mated a Mat a Mcm2–7 overexpression strain with a Mat α GINS–Cdc45 strain (Fig. S1*B*), yielding a diploid strain that produces yeast CMG. A similar method was recently used to produce the 15-member Ino80 complex (40). Purification strategies involved FLAG affinity [to capture FLAG-Mcm3/Cdc45$^{\text{internal FLAG}}$, strain yJCZ2 (Fig. S1*C*) or alternatively Cdc45$^{\text{internal FLAG}}$, strain yJCZ3 (Fig. S1*D*)], followed by anion exchange steps. This approach reproducibly yields Coomassie-stainable amounts of CMG helicase (Fig. 1*A*), in the low micromolar concentration range, which is necessary for reconstitution and further purification of a polymerase cocomplex (see *Materials and Methods*). Importantly, we can show that our yeast CMG is as vigorous a helicase as the baculovirus-expressed *Drosophila* CMG (4, 22) (Fig. 1*B*). Likewise, 2D (Fig. 1*C*) and 3D (Fig. 1*D*) EM image analysis demonstrates that our yeast CMG structure is virtually identical to the published *Drosophila* (16) and yeast EM volumes (24) (Fig. S1). Docking the atomic coordinates of the yeast CMG [Protein Data Bank (PDB) ID code 3JC5 (20)] into our 3D EM map allows the unambiguous identification of each of the 11 helicase subunits in our structure (Fig. 1*D*).

Pol Epsilon Architecture. Wild-type yeast polymerase epsilon has been characterized in an early cryo-EM single-particle study

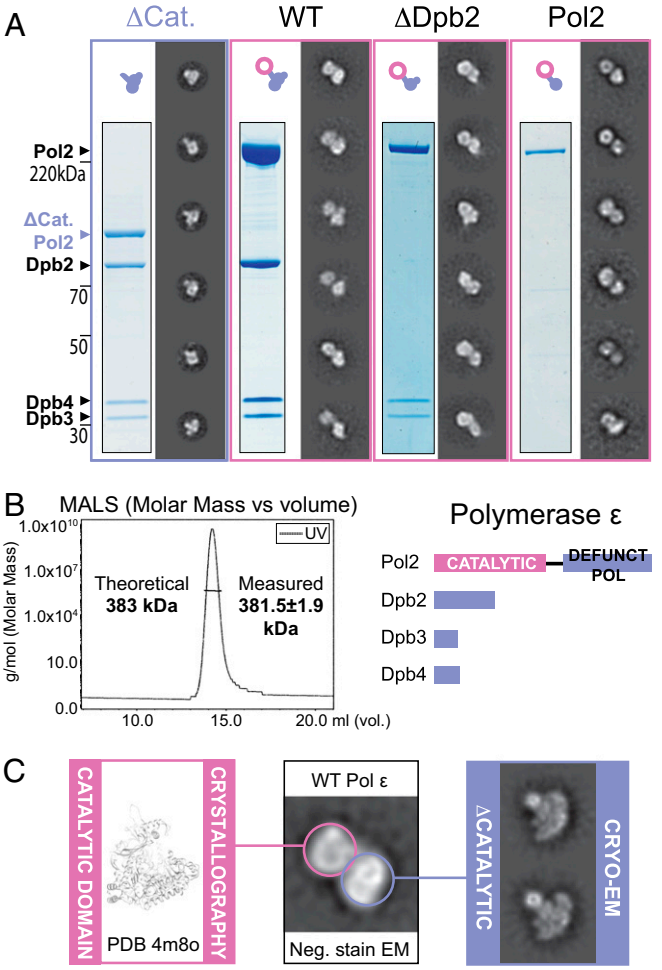

Fig. 2. Architecture of the Pol epsilon complex. (*A*) From the left: A tetrameric complex lacking the catalytic domain of Pol2 (Δcat) forms a singly lobed assembly. The WT complex forms a bilobed assembly. A Dpb2 subunit dropout trimeric complex forms a bilobed assembly. Surprisingly, the isolated Pol2 is also bilobed. (*B*, *Left*) Size-exclusion chromatography with multiangle light scattering shows that the wild-type Pol epsilon complex is a single homotetramer in solution and not a dimer of tetramers. (*Right*) A stick diagram representing the Pol epsilon tetramer. (*C*) The Pol epsilon assembly is formed of two lobes. One lobe comprises the N-terminal catalytic domain of Pol2 (for which a crystal structure is available; PDB ID code 4m8o), and the second lobe comprises the noncatalytic portion of the complex (for which we determined a 2D cryo-EM structure).

(41), however the high-resolution X-ray structure of the catalytic domain (27) cannot be easily integrated with the lower resolution EM data. In an attempt to reconcile the two studies, we have decided to image yeast Pol epsilon by negative stain EM. To this end, we have characterized the active polymerase preparation used in the reconstituted yeast DNA replication system (38) (Fig. S2). We observed that wild-type Pol epsilon is composed of two lobes that appear connected by an isthmus of electron density (Fig. 2*A* and Movie S1). To establish whether the two lobes represent a dimeric form of the protein complex or rather a monomer, we have analyzed the same preparation by size-exclusion chromatography with multiangle light scattering. According to our measurements, the absolute molecular mass of Pol epsilon is 381.5 ± 1.9 kDa, in striking agreement with a predicted molecular mass of 383 kDa for a monomeric, CBP-tagged Pol epsilon complex (Fig. 2*B*). A maltose-binding protein (MBP) fused to one of four Pol epsilon subunits produces a polymerase particle decorated with one, not two, bright density feature proximal to the polymerase particle, in further support of the notion that Pol epsilon forms a single, not a double, hetero-tetramer. We conclude that Pol epsilon is formed by two spatially separated lobes, probably connected via a linker. To establish the identity of the two lobes, we have characterized a Pol epsilon complex lacking the Dpb2 subunit, to find that this assembly still exists as a bilobed entity (Fig. 2*A*, Fig. S3, and Movie S2). We then characterized the isolated Pol2 subunit of Pol epsilon and discovered that this polypeptide alone also contains a bilobed structure (Fig. 2*A*, Fig. S4, and Movie S3). We tentatively assign one lobe to the N-terminal catalytic domain and the second lobe to a C-terminal, catalytically defunct polymerase repeat. Our results on Pol2 differ from an earlier EM study on the same isolated subunit, which had been described as a singly lobed entity (41). To validate our results, we analyzed a deletion mutant of Pol epsilon that contains all subunits but lacks the N-terminal catalytic domain of Pol2 ("Δcat"; Fig. S5). As predicted, Δcat forms one singly lobed structure (Fig. 2*A* and Movie S4), and cryo-EM 2D analysis confirms that this is a large and compact protein assembly (Fig. 2*C* and Fig. S6), in contrast to earlier work that describes the non-catalytic portion of Pol epsilon as a poorly structured entity (41).

In summary, we conducted negative stain and cryo-EM studies on wild-type, subunit dropout, and domain deletion mutants. We find that Pol epsilon contains a bilobed structure, where the catalytic domain of Pol2 constitutes one lobe and the noncatalytic modules in the assembly form the second lobe (Fig. 2*C*). Our findings support the notion that functionally separated modules in Pol epsilon are indeed spatially separated.

CMG–Pol Epsilon Reconstitution. To establish how leading/lagging-strand segregation is achieved at the eukaryotic replication fork, an exhaustive description of the CMG–Pol epsilon complex is needed (42). Recent structural work on the eukaryotic replisome shows that Pol epsilon is anchored to the ATPase side of the helicase ring (24). To experimentally locate the catalytic domain of Pol epsilon in the helicase–polymerase complex, we have developed a protocol to reconstitute the CMG–Pol epsilon assembly, followed by mild cross-linking (XL) and purification over a glycerol gradient (see *Materials and Methods*). Our preparation yielded homogeneous, monodisperse, stabilized particles that are suitable for EM analysis (Fig. S7). As discernible in 2D averages of side views, CMG bound to wild-type Pol epsilon appears decorated with a bilobed feature. One lobe is proximal to the ATPase tier of the MCM motor, whereas the second lobe is more peripheral. Our observation agrees with the notion that the isolated, monomeric Pol epsilon is a bilobed entity, indicating a CMG:Pol epsilon stoichiometry of 1:1 (Fig. 3*A*). To establish the orientation of Pol epsilon bound to the CMG, we repeated the helicase–polymerase reconstitution experiment using the Pol epsilon Δcat mutant (Fig. S8). As expected from our characterization of the isolated deletion mutant, the CMG appears decorated with one lone polymerase lobe (Fig. 3*A*). The 3D reconstruction (Fig. S7 *D–F*) of the full wild-type assembly yields a recognizable CMG structure, bound to polymerase density that departs from the ATPase tier and extends toward the outer perimeter of the helicase structure, contacting the peripheral helicase component Cdc45 (Fig. 3 *B–D*). Comparing the 3D structure of the full assembly with the 2D averages and the 3D structure of the reconstituted CMG–Δcat Pol epsilon, the catalytic domain of Pol epsilon can be unambiguously located at a 90° offset to the CMG helicase ring pore and not proximal to the ATPase tier (Fig. 3 and Fig. S8 *C–F*).

We wondered whether short exposure to mild XL might have introduced artefacts in our preparations. Three lines of evidence argue against this notion. First, with the exception of the recovered density of the catalytic domain, our structure of the CMG–Pol epsilon appears similar to the previously published structure (with particular emphasis on the Pol epsilon anchor; Movie S5) (42). Second, the bilobed feature in the cross-linked CMG–Pol epsilon is highly reminiscent of the non–cross-linked, isolated wild-type Pol epsilon assembly (Fig. 3*A*). Third, our structure agrees with the published XL–mass spectrometry (XL-MS) characterization of the CMG–Pol epsilon architecture (24). According to the XL-MS study, one major CMG contact made by the Pol epsilon catalytic domain is with Cdc45 helix α6 (20, 24, 43). This Cdc45 element projects radially from the CMG core and intimately contacts the catalytic domain of Pol epsilon

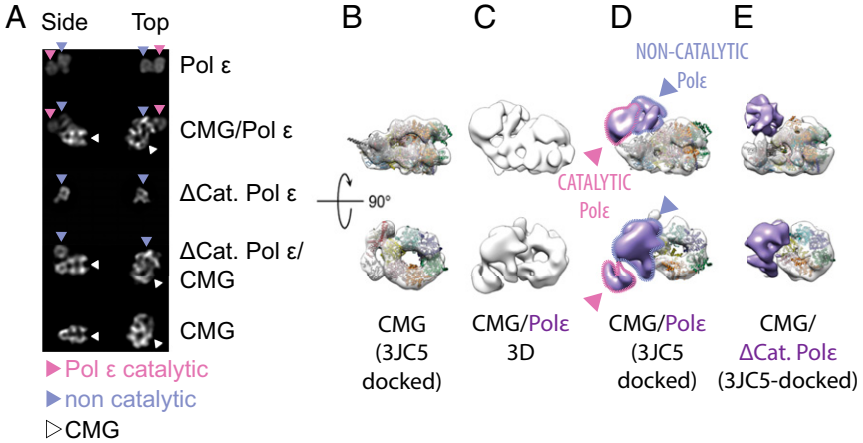

Fig. 3. CMG–Pol epsilon reconstitution and 3D structure. (*A*) The isolated Pol epsilon has a bilobed structure (top row). In complex with Pol epsilon, the CMG is decorated with a bilobed feature (second row). The isolated Δcat Pol epsilon is a singly-lobed entity (third row). In complex with Δcat Pol epsilon, the CMG is decorated with a singly lobed feature (fourth row). Characteristic side and top views of the CMG (bottom row). (*B*) Yeast CMG reconstruction with docked atomic coordinates (PDB ID code 3JC5). (*C*) 3D structure of the CMG–Pol epsilon complex. (*D*) 3D structure of CMG–Pol epsilon complex with docked atomic coordinates of the CMG and assigned catalytic domain and noncatalytic portion of Pol epsilon. The catalytic domain departs radially from the core particle. Density corresponding to the polymerase is highlighted in purple. (*E*) 3D structure of the CMG–Δcat Pol epsilon complex color-coded as in *C*, with docked atomic coordinates.

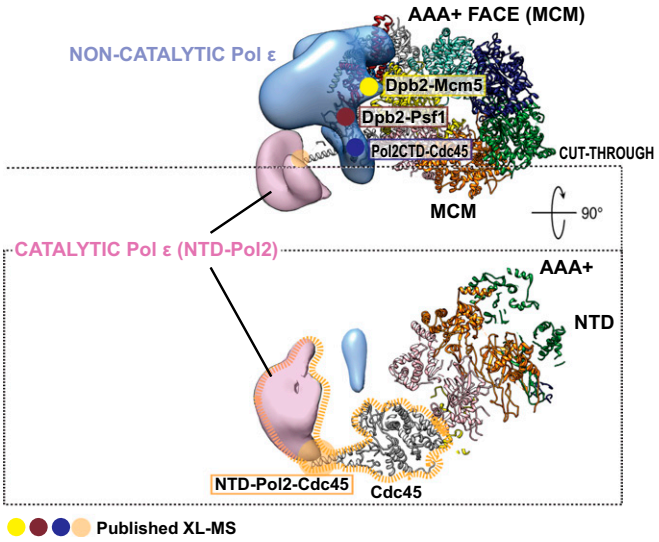

Fig. 4. Integration of the 3D EM structure of the CMG–Pol epsilon with published XL-MS data. The noncatalytic portion of Pol epsilon sits on top of the ATPase tier of the MCM. Dpb2 contacts Mcm5 (yellow) and Psf1 (brown), and Pol2-CTD contacts Cdc45 (blue). The catalytic domain of Pol2 contacts the tip of helix α6 in Cdc45 (orange).

in our structure (Fig. 4). Other detected contacts between the noncatalytic portion of Pol epsilon and the CMG (including Pol2-CTD·Cdc45, Dpb2·Psf1, and Dpb2·Mcm5) are summarized in Fig. 4. Altogether, our data indicate that Pol epsilon can exist in an extended bilobed configuration when CMG-bound, with the catalytic domain of Pol epsilon departing radially from the globular core of the complex. The previous structure reported by O'Donnell, Li and colleagues might have captured a distinct state of the polymerase, where the catalytic module is markedly flexible (hence invisible in the averaged 3D structure) or alternatively very tightly compacted against the noncatalytic polymerase-anchor domain (24).

CMG–Pol Epsilon Dynamics. We have so far established that Pol epsilon is a bilobed entity, with two globular domains that appear to be connected by a linker (Fig. 2). Using a protocol that involved reconstitution, XL, glycerol gradient purification, and imaging, we have described the intact helicase–polymerase complex, containing two recognizable globular domains for Pol epsilon (Fig. 3). Taken together, these data point to an inherent flexible nature of Pol epsilon, which might play an important role in replisome dynamics, in particular during leading-strand replication establishment (see *Discussion*). To further characterize the different conformational states of Pol epsilon, we revisited the 2D image analysis of the isolated wild-type enzyme. By aligning all particles to one lone Pol epsilon domain, it appears that the second domain exists in two alternate states, either compact or extended (Movie S6). To verify that what we observe is a real conformational switch and not distinct 2D views of the same 3D object, we have repeated the analysis using a Pol epsilon derivative that contains a C-terminal MBP fusion of the Dpb3 subunit (Fig. S9). Our 2D analysis indicates that the MBP maps either in close proximity to the interface between the two lobes (equator) or alternatively at the tip of one lobe (south pole; Fig. 5*A* and Movie S7). This result supports the notion that the Dpb3-containing noncatalytic portion of Pol epsilon rotates with respect to the catalytic domain. We cannot rule out, however, from this experiment alone, a second scenario whereby Dpb3 has two alternate binding sites on Pol epsilon, resulting in two possible locations for the MBP tag. To exclude this second hypothesis, we repeated the MBP fusion experiment, tagging the C-terminal domain of Pol2, which is the only subunit that spans the two lobes of Pol epsilon (Fig. S10). As in the previous experiment, the MBP density can be observed either at the equator or at the south pole of the assembly (Fig. 5*B* and Movie S7). Collectively, our data establish that Pol epsilon undergoes a large-scale movement with one globular domain rotating with respect to the other. In conclusion, the isolated Pol epsilon undergoes a conformational change reminiscent of a switchblade knife being ejected from its handle.

Integrating the rotating Pol epsilon complex into the helicase–polymerase superassembly could provide new important insights into replisome dynamics. We have therefore extended the analysis on polymerase flexibility to the CMG–Pol epsilon complex. We found that the catalytic domain of Pol epsilon moves with respect to the core particle and exists in two alternate states (Fig. 5 *C* and *D* and Movie S8), either projecting outward from Cdc45 or bent inwards and in closer proximity to the helicase ring (two conformations captured in a recent XL/MS analysis) (24). The inherent flexibility of Pol epsilon hence persists in the helicase–polymerase assembly. We postulate that the two Pol epsilon configurations reflect distinct functional states, and DNA engagement might select for one of these two forms.

Discussion

We have reconstituted and imaged a 15-member assembly of the eukaryotic replisome, comprising the CMG helicase and the leading-strand polymerase, Pol epsilon. The DNA synthesis

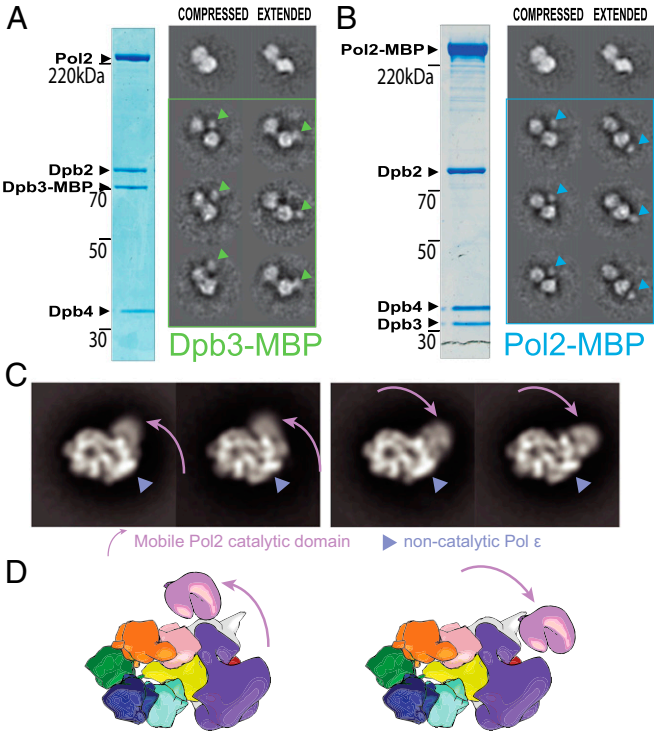

Fig. 5. CMG–Pol epsilon dynamics. (*A*) The isolated Pol epsilon complex with a fused MBP tag at the C terminus of Dpb3 contains two configurations: compressed (with the MBP tag mapping at the interface between lobes, equator) or extended (with the MBP tag mapping at the tip of one lobe, south pole). (*B*) The isolated Pol epsilon complex with a fused MBP tag at the C terminus of Pol2 also contains two configurations: compressed (MBP tag at the equator) or extended (MBP-tag at the south pole). (*C*) The CMG–Pol epsilon contains a flexible catalytic domain of Pol2. This can be found in proximity to the MCM ring (*Left*) or departing radially from the core particle (*Right*). (*D*) Cartoon representation depicting the flexibility of the CMG–Pol epsilon.

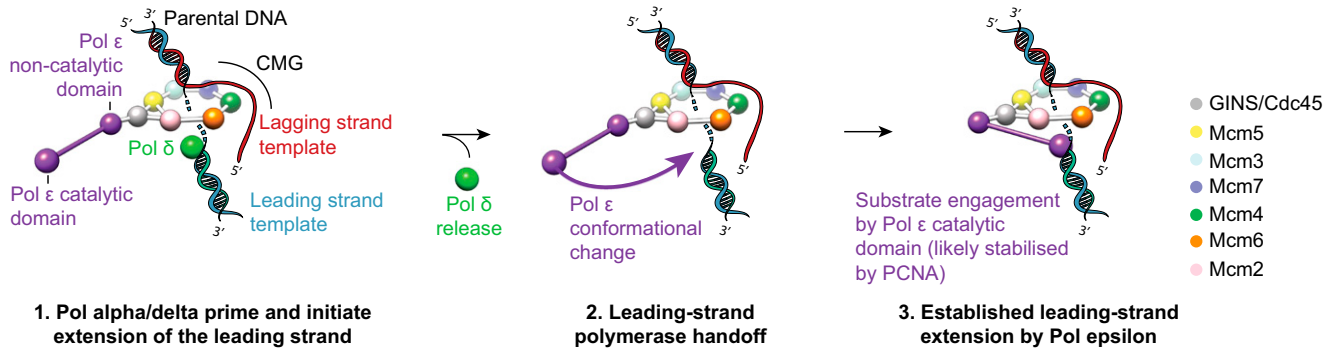

Fig. 6. A polymerase-switch mechanism for the establishment of leading-strand synthesis, facilitated by a reconfiguration of the Pol epsilon catalytic domain.

domain in this complex is highly dynamic, providing important insights into the function of the DNA replication machinery.

According to the commonly accepted model of the replication fork, lagging-strand synthesis occurs discontinuously, with cycles of priming and extension of Okazaki fragments, catalyzed by Pol alpha and delta, respectively. Conversely, the leading-strand template is copied continuously by Pol epsilon (44). Lagging-strand synthesis is highly dynamic, and Pol alpha and delta might be recycled as the replication fork progresses. In agreement with this notion, no stable association of Pol delta with the replisome core has been reported to date, and loss of the (Ctf4-mediated) association between Pol alpha and the CMG helicase does not seem to affect fork progression rates (14, 43, 45, 46). A dynamic interplay between different polymerases might not, however, be restricted to the duplication of the lagging strand (12, 13). Recent findings, in fact, suggest that two subsequent polymerase switching events—(*i*) Pol alpha to Pol delta and (*ii*) Pol delta to Pol epsilon—might both be required to establish leading-strand duplication (14). This highly choreographed process could revolve around a stable, however dynamic, CMG and Pol epsilon assembly.

An overview of our current knowledge on Pol epsilon processivity is useful to understand how the CMG helicase and Pol epsilon work together at the replication fork. In all characterized replication systems, processivity factors help tether the replicative polymerase to newly duplicated DNA, often by topologically enclosing the double helix (1, 47, 48). However, the requirement for a dedicated, canonical processivity factor by Pol epsilon has been a matter of debate. Crystallographic analysis on the isolated catalytic domain of Pol2 has revealed a unique insertion that allows the polymerase to encircle the nascent duplex DNA, providing what appears to be an in-built processivity collar (27). Furthermore, in work reported by Langston, O'Donnell, and colleagues, the CMG helicase and Pol epsilon could be purified from yeast cells as a stable protein complex. Because the MCM motor component of the CMG helicase is a ring that encircles DNA, the authors suggested that the helicase itself might act as the main processivity factor that links Pol epsilon to the replication fork (35). These findings are in line with early observations that the DNA sliding clamp PCNA (the processivity factor in the eukaryotic replication fork) only mildly stimulates processivity in a minimal replisome system (11, 14, 49). DNA tethering by the helicase alone, however, is not sufficient to ensure optimal Pol epsilon processivity. In fact, recent reconstitution studies of a more complete replisome that includes the Mrc1, Tof1, and Csm3 fork stabilization factors indicate that cellular rates of fork progression can only be achieved when Pol epsilon is PCNA-associated (14). Therefore, one would predict that the DNA synthesis domain of Pol epsilon is functionally separated and only a PCNA link can lead to stable association between the Pol2 catalytic domain and the DNA template (whereas the helicase-anchor module would play an important role in the correct assembly of the CMG helicase and possibly stimulate DNA unwinding) (14). Using negative stain EM, O'Donnell, Li, and colleagues have recently analyzed a helicase–polymerase supercomplex produced by mixing DNA-coincubated yeast CMG with Pol epsilon. The derived structure appears highly compact, and the proposed assignment placed the catalytic domain of Pol2 next to the ATPase ring of the CMG helicase. The Pol epsilon density, however, appears not to account for the expected molecular mass of Pol epsilon, and the authors entertained the possibility that the helicase–polymerase assembly might be incomplete (24). Using an overexpression strain to produce the CMG and a multistep protocol that combines salt-dialysis reconstitution, protein XL, and purification over a glycerol gradient for electron microscopy sample preparation, we have now obtained a complete structure of Pol epsilon bound to the CMG helicase and experimentally located the DNA synthesis domain in the superassembly. Our results indicate that the ATPase-anchor module does not contain the catalytic domain, but rather the polymerase domain extends radially from the side of the CMG helicase and is free to switch between two positions around the equator of the protein assembly. We speculate that the dynamic nature of the DNA-synthesis module in Pol epsilon might relate to DNA binding and release, with only one form being competent for substrate binding. In turn, PCNA engagement by Pol epsilon might stabilize the interaction between the DNA synthesis domain of Pol2 and the leading strand, for fully processive leading-strand synthesis. We note that DNA binding and release by the catalytic domain of Pol epsilon might facilitate substrate handoff between Pol delta and epsilon, while Pol epsilon remains physically tethered to the moving fork via the CMG helicase (14). Such a handoff mechanism has been invoked during the establishment of leading-strand synthesis in the early stages of replisome progression and in reestablishing coupled leading-strand synthesis after DNA damage repair (Fig. 6). This model will be tested in the future, using higher resolution cryo-EM on a CMG–Pol epsilon complex, imaged in the act of processively duplicating the leading strand.

Materials and Methods

Full details of experimental procedures are described in *SI Materials and Methods*.

Protein Purification. Yeast strains for the expression of either CMG or Pol epsilon were first induced with galactose, then harvested and lysed with a freezer mill. CMG extracts were first purified by affinity column followed by successive anion-exchange steps. Pol epsilon extracts were first purified by affinity column followed by Heparin and size-exclusion steps.

CMGE Reconstitution. Purified CMG and Pol epsilon were dialyzed together in the reconstitution buffer, cross-linked with 0.05% glutaraldehyde, and then

BIOCHEMISTRY

separated by glycerol gradient sedimentation. Fractions were collected and used for EM preparation.

EM Sample Preparation and Data Collection. For negative staining, protein sample was applied to glow-discharged, carbon-coated copper grids; stained with successive drops of 2% uranyl formate solution; and then blotted dry. Grids were either imaged on a Tecnai G2 Spirit transmission electron microscope (FEI) at 120 kV (for all Pol epsilon constructs and CMG–Δcat) or using a JEM-2100 LaB6 electron microscope (JEOL) at 120 kV (for CMG and CMG–Pol epsilon). For cryopreparation of Δcat Pol epsilon, purified sample was applied to glow-discharged C-flat grids, blotted, and plunge-freezed. Data were collected on a Tecnai F30 Polara electron microscope at 300 kV with a K2 Summit direct electron detector (Gatan, Inc.).

Single-Particle Image Processing. Particles were picked semiautomatically in EMAN2 (50). Contrast transfer function parameters were estimated using CTFFIND4 (51). All further processing was performed in RELION (52).

Note Added in Proof. After submission of this manuscript, a paper describing new CMG–DNA structures was published by O'Donnell, Li, and colleagues (53). This study supports the notion that the ATPase motor of the CMG is a single-stranded DNA translocase.

ACKNOWLEDGMENTS. We thank Gideon Coster, Ann Early, and Lucy Drury for help with yeast strain; Ali Alidoust and Namita Patel for yeast cultures; Svend Kjaer for support with the multi-angle light scattering measurements; Raffaella Carzaniga and Lucy Collinson (The Francis Crick Institute) and Kirsty MacLellan-Gibson (National Institute for Biological Standards and Control) for technical support with EM; and Hasan Yardimci and Hazal Kose for generous help with helicase assay. This work was supported by The Francis Crick Institute [core funding from Cancer Research UK, the UK Medical Research Council, and the Wellcome Trust (WT); to J.F.X.D. and A.C.] and two PhD fellowships from the Boehringer Ingelheim Fonds (to A.J. and F.A.A.). J.F.X.D. is the recipient of European Research Council Grant 669424-CHROMOREP and Wellcome Senior Investigator Award 106252/Z/14/Z. A.K. is supported by the WT. The Division of Structural Biology–Particle Imaging Center Electron Microscopy Facility at the University of Oxford was founded by the WT JIF Award 060208/Z/00/Z and is supported by WT Equipment Grant 093305/Z/10/Z. The WT, Medical Research Council, and Biotechnology and Biological Sciences Research Council also support the National EM facility, which enabled provision of the K2 detector at Oxford.

1. Zhang D, O'Donnell M (2016) The eukaryotic replication machine. *Enzymes* 39: 191–229.
2. Yao NY, O'Donnell ME (2016) Evolution of replication machines. *Crit Rev Biochem Mol Biol* 51:135–149.
3. Moyer SE, Lewis PW, Botchan MR (2006) 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.
4. Ilves I, Petojevic T, Pesavento JJ, Botchan MR (2010) Activation of the MCM2-7 helicase by association with Cdc45 and GINS proteins. *Mol Cell* 37:247–258.
5. Pellegrini L (2012) The Pol α-primase complex. *Subcell Biochem* 62:157–169.
6. Clausen AR, et al. (2015) Tracking replication enzymology in vivo by genome-wide mapping of ribonucleotide incorporation. *Nat Struct Mol Biol* 22:185–191.
7. Nick McElhinny SA, Gordenin DA, Stith CM, Burgers PM, Kunkel TA (2008) Division of labor at the eukaryotic replication fork. *Mol Cell* 30:137–144.
8. Pursell ZF, Isoz I, Lundström EB, Johansson E, Kunkel TA (2007) Yeast DNA polymerase epsilon participates in leading-strand DNA replication. *Science* 317:127–130.
9. Miyabe I, Kunkel TA, Carr AM (2011) The major roles of DNA polymerases epsilon and delta at the eukaryotic replication fork are evolutionarily conserved. *PLoS Genet* 7:e1002407.
10. Yu C, et al. (2014) Strand-specific analysis shows protein binding at replication forks and PCNA unloading from lagging strands when forks stall. *Mol Cell* 56:551–563.
11. Georgescu RE, et al. (2014) Mechanism of asymmetric polymerase assembly at the eukaryotic replication fork. *Nat Struct Mol Biol* 21:664–670.
12. Johnson RE, Klassen R, Prakash L, Prakash S (2015) A major role of DNA polymerase δ in replication of both the leading and lagging DNA strands. *Mol Cell* 59:163–175.
13. Stillman B (2015) Reconsidering DNA polymerases at the replication fork in eukaryotes. *Mol Cell* 59:139–141.
14. Yeeles JT, Janska A, Early A, Diffley JF (2017) How the eukaryotic replisome achieves rapid and efficient DNA replication. *Mol Cell* 65:105–116.
15. Barry ER, McGeoch AT, Kelman Z, Bell SD (2007) Archaeal MCM has separable processivity, substrate choice and helicase domains. *Nucleic Acids Res* 35:988–998.
16. Costa A, et al. (2011) The structural basis for MCM2-7 helicase activation by GINS and Cdc45. *Nat Struct Mol Biol* 18:471–477.
17. Pape T, et al. (2003) Hexameric ring structure of the full-length archaeal MCM protein complex. *EMBO Rep* 4:1079–1083.
18. Fletcher RJ, et al. (2003) The structure and function of MCM from archaeal *M. thermoautotrophicum*. *Nat Struct Biol* 10:160–167.
19. Li N, et al. (2015) Structure of the eukaryotic MCM complex at 3.8 Å. *Nature* 524: 186–191.
20. Yuan Z, et al. (2016) Structure of the eukaryotic replicative CMG helicase suggests a pumpjack motion for translocation. *Nat Struct Mol Biol* 23:217–224.
21. McGeoch AT, Trakselis MA, Laskey RA, Bell SD (2005) Organization of the archaeal MCM complex on DNA and implications for the helicase mechanism. *Nat Struct Mol Biol* 12:756–762.
22. Abid Ali F, et al. (2016) Cryo-EM structures of the eukaryotic replicative helicase bound to a translocation substrate. *Nat Commun* 7:10708.
23. Fu YV, et al. (2011) Selective bypass of a lagging strand roadblock by the eukaryotic replicative DNA helicase. *Cell* 146:931–941.
24. Sun J, et al. (2015) The architecture of a eukaryotic replisome. *Nat Struct Mol Biol* 22: 976–982.
25. Simon AC, et al. (2014) A Ctf4 trimer couples the CMG helicase to DNA polymerase α in the eukaryotic replisome. *Nature* 510:293–297.
26. Feng W, D'Urso G (2001) Schizosaccharomyces pombe cells lacking the amino-terminal catalytic domains of DNA polymerase epsilon are viable but require the DNA damage checkpoint control. *Mol Cell Biol* 21:4495–4504.
27. Hogg M, et al. (2014) Structural basis for processive DNA synthesis by yeast DNA polymerase ε. *Nat Struct Mol Biol* 21:49–55.
28. Johansson E, Macneill SA (2010) The eukaryotic replicative DNA polymerases take shape. *Trends Biochem Sci* 35:339–347.
29. Tahirov TH, Makarova KS, Rogozin IB, Pavlov YI, Koonin EV (2009) Evolution of DNA polymerases: An inactivated polymerase-exonuclease module in Pol epsilon and a chimeric origin of eukaryotic polymerases from two classes of archaeal ancestors. *Biol Direct* 4:11.
30. Araki H, Hamatake RK, Johnston LH, Sugino A (1991) DPB2, the gene encoding DNA polymerase II subunit B, is required for chromosome replication in Saccharomyces cerevisiae. *Proc Natl Acad Sci USA* 88:4601–4605.
31. Dua R, Levy DL, Campbell JL (1999) Analysis of the essential functions of the C-terminal protein/protein interaction domain of Saccharomyces cerevisiae pol epsilon and its unexpected ability to support growth in the absence of the DNA polymerase domain. *J Biol Chem* 274:22283–22288.
32. Isoz I, Persson U, Volkov K, Johansson E (2012) The C-terminus of Dpb2 is required for interaction with Pol2 and for cell viability. *Nucleic Acids Res* 40:11545–11553.
33. Kesti T, Flick K, Keränen S, Syväoja JE, Wittenberg C (1999) DNA polymerase epsilon catalytic domains are dispensable for DNA replication, DNA repair, and cell viability. *Mol Cell* 3:679–685.
34. Kang YH, Galal WC, Farina A, Tappin I, Hurwitz J (2012) Properties of the human Cdc45/Mcm2-7/GINS helicase complex and its action with DNA polymerase epsilon in rolling circle DNA synthesis. *Proc Natl Acad Sci USA* 109:6042–6047.
35. Langston LD, et al. (2014) CMG helicase and DNA polymerase ε form a functional 15-subunit holoenzyme for eukaryotic leading-strand DNA replication. *Proc Natl Acad Sci USA* 111: 15390–15395.
36. Georgescu RE, et al. (2015) Reconstitution of a eukaryotic replisome reveals suppression mechanisms that define leading/lagging strand operation. *eLife* 4:e04988.
37. Hogg M, Johansson E (2012) DNA polymerase ε. *Subcell Biochem* 62:237–257.
38. Yeeles JT, Deegan TD, Janska A, Early A, Diffley JF (2015) Regulated eukaryotic DNA replication origin firing with purified proteins. *Nature* 519:431–435.
39. Frigola J, Remus D, Mehanna A, Diffley JF (2013) ATPase-dependent quality control of DNA replication origin licensing. *Nature* 495:339–343.
40. Kurat CF, Yeeles JT, Patel H, Early A, Diffley JF (2017) Chromatin controls DNA replication origin selection, lagging-strand synthesis, and replication fork rates. *Mol Cell* 65:117–130.
41. Asturias FJ, et al. (2006) Structure of Saccharomyces cerevisiae DNA polymerase epsilon by cryo-electron microscopy. *Nat Struct Mol Biol* 13:35–43.
42. Sun J, Yuan Z, Georgescu R, Li H, O'Donnell M (2016) The eukaryotic CMG helicase pumpjack and integration into the replisome. *Nucleus* 7:146–154.
43. Simon AC, Sannino V, Costanzo V, Pellegrini L (2016) Structure of human Cdc45 and implications for CMG helicase function. *Nat Commun* 7:11638.
44. Burgers PM, Gordenin D, Kunkel TA (2016) Who is leading the replication fork, Pol ε or Pol δ? *Mol Cell* 61:492–493.
45. Pellegrini L, Costa A (2016) New insights into the mechanism of DNA duplication by the eukaryotic replisome. *Trends Biochem Sci* 41:859–871.
46. Gambus A, et al. (2009) A key role for Ctf4 in coupling the MCM2-7 helicase to DNA polymerase alpha within the eukaryotic replisome. *EMBO J* 28:2992–3004.
47. Lee SJ, Richardson CC (2011) Choreography of bacteriophage T7 DNA replication. *Curr Opin Chem Biol* 15:580–586.
48. Lewis JS, Jergic S, Dixon NE (2016) The E. coli DNA replication fork. *Enzymes* 39:31–88.
49. Chilkova O, et al. (2007) The eukaryotic leading and lagging strand DNA polymerases are loaded onto primer-ends via separate mechanisms but have comparable processivity in the presence of PCNA. *Nucleic Acids Res* 35:6588–6597.
50. Tang G, et al. (2007) EMAN2: An extensible image processing suite for electron microscopy. *J Struct Biol* 157:38–46.
51. Rohou A, Grigorieff N (2015) CTFFIND4: Fast and accurate defocus estimation from electron micrographs. *J Struct Biol* 192:216–221.
52. Scheres SH (2012) RELION: Implementation of a Bayesian approach to cryo-EM structure determination. *J Struct Biol* 180:519–530.
53. Georgescu R, et al. (2017) Structure of eukaryotic CMG helicase at a replication fork and implications to replisome architecture and origin initiation. *Proc Natl Acad Sci USA* 114:E697–E706.