Molecular insights into the MCM8/9 helicase complex in DNA unwinding and translocation

Zhuangfeng Weng[1,†], Jun Li[1,†], Jiefu Zheng[1,†], Huanhuan Li[2,†], Zuer ,Lu[3], Yiyi Zhou[1], Dongyi, Xu[3], Huanhuan Liang[4,*], and Yingfang Liu[1,2,*]

[1] School of Medicine, Shenzhen Campus of Sun Yat-sen University, Shenzhen, Guangdong 518107, China.

[2] Department of Colorectal Surgery, The Sixth Affiliated Hospital, Sun Yat-sen University, Guangdong Institute of Gastroenterology, Guangdong Provincial Key Laboratory of Colorectal and Pelvic Floor Diseases, Guangzhou, Guangdong 510655, China.

[3] State Key Laboratory of Protein and Plant Gene Research, School of Life Sciences, Peking University, Beijing, China.

[4] Pharmaceutical Sciences (Shenzhen), Sun Yat-sen University, Shenzhen, Guangdong 518107, China.

[†] These authors contributed equally to this work.

* Correspondence should be addressed to Yingfang Liu (liuyingf5@mail.sysu.edu.cn) or Huanhuan Liang (lianghh26@mail.sysu.edu.cn)

Running titer: Structure of MCM8/9 helicase complex

Abstract

The minichromosome maintenance protein 8 and 9 (MCM8 and MCM9) form a functional helicase complex (MCM8/9) that plays an essential role in DNA homologous recombination (HR) repair for DNA double-strand break. However, the assembly mechanism and the structural characterization of MCM8/9 remain unclear. Here, we report structures of the MCM8/9 complex combining X-ray crystallography and cryo-electron microscopy reconstruction analysis. The structures reveal that MCM8/9 is arranged by a 3-fold symmetry axis to form a heterohexamer with a central channel to accommodate DNA. Multiple characteristic hairpins from the N-terminal oligosaccharide/oligonucleotide (OB) domains of MCM8/9 hexameric ring protrude to the central channel and serve to unwind the duplex DNA. Importantly, the MCM8/9 hexamer is consisted of an N-tier ring and a C-tier ring that are connected by the flexible linkers (N-C linkers) which are much longer than those of MCM2-7. Our structural dynamic analyses based on the cryo-EM data reveal that the flexible C-tier ring showed rotating motions relative to the N-tier ring, which maybe mediated by the N-C linkers as shorten the length of the N-C linkers will greatly decrease the helicase activity and chemoresistance of MCM8/9. Our structural and biochemical results implicate an unusual DNA unwinding and translocation mechanism of MCM8/9 helicase complex in HR.

Key words:

MCM8/9; helicase complex; hexamer; homologous recombination; cryo-EM

Introduction

DNA helicases play critical roles in many cellular processes including DNA replication, transcription, recombination and repair. They translocate along the DNA and unwind DNA duplex in an ATP-dependent manner, which generates transient single-stranded DNA intermediates required for the above cellular processes. Dysfunction of multiple DNA helicases have been correlated with various human diseases including cancers, Bloom Syndrome, Werner Syndrome, Fanconi Anemia, reproductive deficiencies and infertility, *etc* (1,2).

The minichromosome maintenance (MCM) proteins comprise one of the most extensively studied subfamilies of DNA helicases. In human genome, the MCM family contains two subgroups of AAA+ ATPase/helicase complexes, consisted of eight MCM members (MCM2-9) (3). Among them, the hexameric MCM2-7 complex, which was known as a replicative helicase, play vital roles in the initiation and elongation during eukaryotic chromosome replication as it formed into different intermediates including the helicase-loading intermediate Orc-Cdc6-Cdt1-Mcm2-7 (OCCM), the inactive Double MCM2-7 Hexamer (DH) and the active replicative helicase complex Cdc45-Mcm2-7-GINS (CMG) (4-6). The origin DNA was encircled into the central channel of the MCM2-7 hexamer via the MCM2-MCM5 entry "gate" (7). Inside the chamber of MCM2-7 hexamer, layers of conserved hairpin loops, including the helix-2 insertion loops (H2I), the presensor-1 (PS1) as well as β-turn motifs of oligosaccharide/oligonucleotide (OB) domains, contribute to DNA binding and unwinding (8). In addition, another loop allosteric communication loop (ACL) conserved in all MCM helicases was also important for helicase activity by mediating inter-subunit interactions between the N-terminal DNA processing domain and the C-terminal AAA+ domain (9,10).

The other helicase complex consists of MCM8 and MCM9, which are homolog of MCM2-7 (11). Both proteins have a similar architecture consisting of an MCM domain at N-terminus (NTD) and an AAA+ ATPase domain at C-terminus (CTD). The NTD and CTD are connected by flexible linkers (N-C linkers), which were reported to stabilize the relative position of NTD and CTD of archaea MCM helicase and permit

the movement of the AAA+ domain during unwinding (12). In addition, MCM8 possesses an additional WHD domain in CTD while MCM9 contains a putative HTH domain and an extensional C-tail at C-terminus (13). Different from the replicative helicase MCM2-7, MCM8 and MCM9 serve as a DNA helicase complex (MCM8/9) in HR-mediated DNA repair for DNA double-strand break (DSB) and DNA interstrand crosslinks (ICLs) (14). Several studies have suggested that cells lacking *MCM8* and *MCM9* were highly sensitive to DNA cross-linking agents (15,16). The *MCM8* and *MCM9* knock-out mice were sterile and revealed gametogenesis deficiency as well as chromosomal damage due to impaired HR (14). During HR repair, MCM8/9 was rapidly recruited to the DNA damage sites and colocalized with the recombinase Rad51 (17). It also interacted with the nuclease complex MRN (MRE11-RAD50-NBS1) and was required for DNA resection at DSBs to facilitate the HR repair (18). Recently, the C17orf53 protein (also called HROB) was identified as an important factor in loading MCM8/9 complex to the sites of DNA damage and stimulating its helicase activity to promote replication fork progression during HR (19,20). Though fruitful progress has been made for MCM8/9 regarding its role in HR during the past decade, the structure/function relationship, especially the structural characterization of MCM8/9 helicase activity in HR remains unknown, which results in scientific gaps of knowledge of MCM8/9 compared with its homolog MCM2-7 helicase complex. It is intriguing to propose a structural model to dissect whether MCM8/9 complex has unusual assembly mechanism compared to MCM2-7 complex.

In the previous study, we reported the crystal structures of the NTDs of MCM8 and MCM9, respectively, and also depicted their DNA binding features and ion binding capability (21). Here, we provide a 3.7 Å resolution cryo-EM structure of chicken MCM8/9 complex and a 6.6 Å crystal structure of human MCM8/9 complex. The structures reveal that MCM8 and MCM9 assemble into a conserved heterohexamer with an N-tier ring and a C-tier ring that are connected by flexible N-C linkers. The architecture of the hexameric MCM8/9 is similar to those primary replicative MCM helicases from eukaryote, yeast and Archaea. But its alternate assemble mode is unique, at least so far. Importantly, the structural flexibility of the gMCM8/9 complex shows

the CTD AAA+ motor ring tilts right and left relative to the NTD, which suggests that MCM8/9 complex may translocate along the DNA to unwind substrates in an unusual rotation fashion.

Results

Cryo-EM structure of gMCM8/9 NTD reveals a hexameric ring

To study the assembly and unwinding mechanisms of the MCM8/9 helicase complex, we first determined the structure of MCM8/9 from chicken (*Gallus gallus*, gMCM8/9) using cryo-EM single particle reconstruction (Figure S1, Table 1, Materials and Methods). The cryo-EM analysis reveals that gMCM8 and gMCM9 alternately assemble into a single heterohexamer which contains a central channel wide enough to accommodate dsDNA. The gMCM8/9 is a two-tiered hexameric ring consisted of the N terminal domains ring (NTD ring) and the C terminal domains ring (CTD ring) (Figure 1A and 1B). As shown in Figure 1B, the electron density map of gMCM8/9 CTD is blurry and hard to define a structural model, indicating that the CTD ring is greatly flexible in solution.

Conversely, the NTD ring of gMCM8/9, which can be well reconstructed at 3.7 Å resolution with a *C3* symmetry axis, clearly shows a pinwheel-like architecture and protrudes its six helical domains around the central channel (Figure 1A). The NTDs of gMCM8 and gMCM9 have a similar domain organization with all the MCM proteins and can be divided into three domains: the zinc finger domain (ZF), the helical domain (HD) and the OB domain. The ZF domain is embedded in the OB domain and usually binds to zinc ions (Figure 1C and 3B). Previous gel filtration study suggested that MCM8/9 can exist as dimer or hexamer (22), which is consistent with our observation that MCM8/9 is a trimer of heterodimer (Figure 1A). This leads to the formation of three putative trimer interfaces between each MCM8-MCM9 heterodimer (Figure 1C). The dimer interface between NTDs of gMCM8 and gMCM9 is largely composed of the β-strands from the OB domain of gMCM9 and the flanking ZF and OB domain of gMCM8. Both the polar and hydrophobic interactions between respective β-strands are responsible for the binding of the MCM8/9 dimer. For example, the side chains of

$R220_{gMCM9}$ and $Q248_{gMCM9}$ form hydrogen bonds with the carbonyl groups in main chains of $I172_{gMCM8}$ and $T226_{gMCM8,}$ respectively. Besides, the main chain of $T119_{gMCM8}$ forms hydrogen bond with the side chain of $D228_{gMCM8}$. Furthermore, the benzene ring of $F253_{gMCM9}$ inserts into a hydrophobic pocket surrounded by residues L179, P217, F181, L201 and Y199 from gMCM8 and W250 from gMCM9, which further stabilize the interaction between gMCM8 and gMCM9. Likewise, the trimer interface is dominated by the electrostatic interaction and hydrogen-bonding interactions between respective strands and loops of gMCM8 and gMCM9. The side chain of $R152_{gMCM8}$ forms salt bridges with the side chains of $D200_{gMCM9}$ and $D232_{gMCM9.}$ Besides, the side chain of $K158_{gMCM8}$ form a hydrogen bond with the side chain of $E37_{gMCM9}$ (Figure 1D and 1E). These interactions facilitate and stabilize the assembly of the NTD hexameric ring of gMCM8/9. Additionally, when we performed structural superposition, the NTD ring of gMCM8/9 superposes well to that of human MCM8/9 (hMCM8/9) with an overall root mean square deviation (RMSD) of about 2.2 Å, indicating that eukaryotic MCM8 and MCM9 form a conserved hexamer complex during evolution (Figure 1F, S2 and S3).

Spatial distribution and functional role of OB fold hairpins in DNA unwinding

As in many hexameric helicase complexes, hairpin loops of OB fold domains, are essential for DNA binding and/or unwinding (8). In the NTD ring of gMCM8/9 complex, there are two kinds of hairpin loops from OB domains. The first kind of OB hairpins of gMCM8 and gMCM9 (refer to OB-hp1s) protrude to the central chamber in a three-fold symmetry manner and narrow down the channel, which are consistent with the previous structures of archaeal MCM (Figure 2A). Particularly, the three OB-hp1s of gMCM8 form the upper layer and the three gMCM9 OB-hp1s comprise the lower layer (Figure 2C). Importantly, the tip of MCM8 OB-hp1 contains several positively charged residues (R345, K347, K350) while the MCM9 OB-hp1 contains continuous polar residues (Q252, Q253, D254) (Figure 2D). As the recent cryo-EM structure of the *S. cerevisiae* CMG on a forked DNA reported that some conserved residues, like K364 and K67 from MCM7 OB hairpin, R449 and R451 from MCM4 OB hairpin loop and

polar residue Q308 and N307 from MCM3 OB fold domain are involved in forked DNA interaction and may contribute to its separation (23). We docked a double-stranded DNA (dsDNA) into the channel and found that the OB-hp1s of gMCM8 contact closely with it and insert their highly positively charged terminal loops into the major or minor grooves of the dsDNA, suggesting that they could play a regulatory role in substrate DNA processing and/or unwinding (Figure 2C).

The OB domains of gMCM8 and gMCM9 harbor another kind of interesting hairpin loops known as the allosteric communication loops (ACL), which were referred here as the OB-hp2s (Figure 2B). These loops locate on the opposite side of the OB-hp1s but extend to the CTD of gMCM8/9. As several studies has found that the ACL act to facilitate the communications between NTD and AAA+ catalytic domain of hexameric helicases as well as modulate the positioning of the OB-hp1s and then regulate the helicase activities (9,10). We propose that the OB-hp2s may also play such functional roles because the OB-hp2s show high degree of conservation when performed superposition of gMCM8/9 NTD to the previous structure of *M. thermautotrophicus* MCM (mtMCM) as well as performed sequences alignment among them (Figure 2E and 2F).

To determine the significant role of the OB domain hairpins, we performed helicase assay using a plasmid-based substrate that usually used to test MCM helicase activity (24,25). The wild type (WT) MCM8/9 complex only exhibited limited ability to unwind the DNA substrate. To investigate the helicase activities of the MCM8/9 and its mutants, we introduced the newly identified regulatory protein—HROB that could increase the helicase activity of the WT MCM8/9 complex (20). As shown in Figure 2G and 2H, deleting the tip residues of the MCM8 OB-hp1s or mutating that of MCM9 OB-hp1s, both decreased the helicase activities of MCM8/9 in the helicase assay. The MCM8 OB-hp2s mutants (MCM8_E295R, MCM8_P308G) corresponding to that in mtMCM slightly decreased the helicase activity in the presence of HROB. Collectively, the above structural analyses and biochemical experiments, together with the repeated reports of the functional roles of OB hairpin loops in NTD of hexameric machines, suggest that the OB hairpins of MCM8/9 play a regulatory role in DNA unwinding.

The overall structure of hMCM8/9 helicase complex with 3-fold symmetric axis

To propose the whole structural model of MCM8/9 helicase complex, we further determined the crystal structure of human core MCM8/9 complex (amino acid, hMCM8: 61-840; hMCM9:1-684) at a resolution of 6.6 Å (Figure 3 and S4, Table 2, Materials and Methods). The phased was solved using the molecular replacement (MR) method by providing the hMCM8/9 NTD heterodimer structure and the hMCM8/9 CTD predictive model as searching model (see Methods for detail). The atomic modules of hMCM8/9, which consists of the NTD domain and the CTD AAA+ motor domain, fit well into their electron density map (Figure S4B and S4C). But the WHD domain of hMCM8, the putative HTH domain of hMCM9, and the N-C liners are not well defined due to the low-resolution electron density map.

The crystal adopts $P2_13$ space group and there is one hMCM8/9 heterodimer in an asymmetric unit. As shown in Figure S4B, the density map shows clear profile of hexamer suggests that the hMCM8/9 heterodimer assemble into a hexameric architecture based on the 3-fold symmetric axis of crystal. This observation is in line with our cryo-EM analyses that the hexameric hMCM8/9 complex is a trimer of heterodimer (Figure 1A). hMCM8 and hMCM9 arrange on an alternative pattern and form a pinwheel-like single hexamer, decorated by six zinc ions binding to the ZFs of hMCM8/9 (Figure 3C). Its architecture is quite similar to that of the eukaryotic MCM2-7 heterohexamer and archaeal MCM homohexamer, except that MCM8 contains a long α5 helix in the HD domain protruding around the hexameric ring. In addition, the OB-hp2 loops from OB domains of hMCM8 and hMCM9 are in close proximity to the H2I and PS1 hairpins within the AAA+ domains, which may mediate communications between the NTD and CTD of the helicase complex as it is equivalent to the ACL loop in archaeal MCM helicases (Figure 3A). At last, the 3-fold symmetric assembly pattern of hMCM8/9 complex indicates its unique features and molecular mechanism.

Structural refinement and comparison indicate a rotational motion of MCM8/9

Taking inspiration from the stimulation effect of HROB on MCM8/9 helicase activity, we propose that MCM8/9 unwind DNA would be tightly controlled as it is for

MCM2-7. If this is the case, the structure of MCM8/9 should undergo conformational change upon DNA binding or helicase activation. To study the dynamics of MCM8/9 complex, we performed structural flexibility analyses using multi-body refinement based on our cryo-EM dataset. This method uses separate focused refinements and generates movies to describe the most important motions on the relative orientations of the rigid bodies, which allows us to gain insight into the molecular motions of gMCM8/9, particularly in the structural flexibility of its CTD ring (26-28). As expected, twelve components with each including a discrete number of independently moving bodies were generated, providing a characterization of the motions in gMCM8/9 CTD ring. As shown in the movie 1, the CTD AAA+ motor ring rotates right and left reciprocally relative to the NTD and the motion drives the conformational change of the N-C linkers. This kind of motion may underlie the unwinding mechanism of MCM8/9 helicase.

As multiple conformations of gMCM8/9 were generated from cryo-EM study, we firstly supposed that our present atomic model of hMCM8/9 was conformationally similar to the inactive DH of MCM2-7 as the WT hMCM8/9 also has very low helicase activity. When superimposing the hexameric hMCM8/9 complex to one hexamer of the MCM2-7 DH by aligning their respective C-tier ring, we found that the two C-tier rings aligned well while their N-tier rings has a ~28° clockwise rotation (Figure 4A and 4B). Such rotating structural differences may arise from the different assemble patterns of the two hexameric complexes. However, a more reasonable explanation is that the DH of MCM2-7 induces the conformational change of its N-tier ring. As during double-hexamer assembly, the N-tier ring of MCM2-7 also rotated by nearly 30° when aligned to the MCM2-7 hexamer in its loading intermediate OCCM (29). We then compared the MCM8/9 structure with that of MCM2-7 in the OCCM complex by aligning their respective C-tier ring. Contrasting to the above alignment with the double hexamer MCM2-7, the NTDs of MCM8/9 fitted well to that of MCM7, MCM3 and MCM5, but rotated slightly in the following MCM2, MCM6 and MCM4 (Figure 4C). This small variance may result from the opening gate of MCM5-MCM2 interface in the OCCM

complex. Thus, the superposition to the OCCM complex also indicates that the present structure of hMCM8/9 is resemble to that of MCM2-7 in a loading intermediate state. And similar to MCM2-7 in different intermediates, we proposed that the MCM8/9 complex may have multiple conformations with rotational motions between NTD and CTD rings during helicase activation.

The long N-C linkers of MCM8/9 paly essential role in MCM8/9 helicase activity

When hexameric helicase unwinds the duplex DNA, great conformational change occurred between N-tier ring and C-tier motor ring, which are bridged by flexible N-C linkers. To study whether the N-C linker is important for MCM8/9 helicase activity, we dissected the N-C linkers of the major eukaryotic MCM proteins (MCM2-9). The lengths of the N-C linkers are variable and can be divided into two groups. The first one composed of MCM2, MCM4, MCM6 and MCM8 possesses a 40-residues length linker while the other group including MCM3, MCM5, MCM7 and MCM9 contains a linker with about 30 residues (Figure S5C and S5D). Although their sequences are varied, the secondary structures of the N-C linkers are more conservative, which possess an N-terminal random loop and a parallel short helix at C terminus (Figure S5A and S5B). Notably, the structures of the N-C linkers could not be well assigned neither in our crystal nor cryo-EM structures (Figure 1A and 3B), indicating that the linkers are quite flexible. We tried to shorten the N-C linkers by deleting several amino acids in loop regions of MCM8 (MCM8Δ369-377) or MCM9 (MCM9Δ283-287) alone to prepare single deletion mutants or combined them into double deletion complex. In the helicase assay *in vitro*, the unwinding ability of all the deletion mutants are dramatically decreased compared to WT MCM8/9 complex. In accordance with the previously observed stimulating effect of HROB, all the mutants retained quite low helicase activities compared to WT complex in the present of HROB, indicating that the N-C linkers are required for MCM8/9 helicase activity (Figure 5A and 5B).

Furthermore, MCM8 and MCM9 had been reported to mediate cellular resistance to DNA cross-linking agent cisplatin that induces DNA interstrand crosslinks and mostly repaired by HR (15,30). We therefore sought to determine whether the N-C linkers were

important for the chemoresistance of MCM8 or MCM9. To this end, we targeted MCM8 or MCM9 with sgRNAs using the CRISPR-Cas9 system in chicken DT40 cells to produce knockout (KO) cell lines. As expected, the *MCM8*- or *MCM9*- deficient cells significantly sensitized to cisplatin relative to the WT DT40 cells. Importantly, expression of WT MCM8 or MCM9 cDNAs in their respective KO cells almost complemented chemoresistance, while expression of their N-C linker deletion mutants had little or no such complementary effect (Figure 5C and 5D). These results suggest that the integral N-C linkers of MCM8/9 are critical for its helicase activity, thus promoting its cellular resistance to DNA damaging agents.

Discussion

In this study, we determined the structures of MCM8/9 helicase complex at near-atomic resolution, which provides many structural characterizations of MCM8/9 helicase. Firstly, MCM8 and MCM9 form a single heterohexameric ring with a centric channel that allows the DNA to pass through. Secondly, inside the chamber, layers of OB hairpin loops may serve to bind and unwind duplex DNA. Thirdly, the N-C linkers of MCM8/9 play important roles in its helicase activity and chemoresistance against DNA cross-linking agents. Notably, the rotary motion of the flexible MCM8/9 CTD AAA+ motor ring may underlie the strand separation and extrusion mechanism.

MCM8/9 forms a single hexameric helicase complex. Both MCM8/9 and MCM2-7 belong to the MCM protein family with highly conserved domain organization and sequence similarity (3,31). At replication origins during G1 phase, MCM2-7 assembles one by one into an inactive DH with the help of loading factors Orc, Cdc6 and Cdt1 (32). While in our study, the MCM8/9 complex mainly exits as a single hexamer in the purified recombinant proteins. In addition, the MCM8/9 hexamer is arranged in an alternate mode with 3-fold symmetry axis, which is unique compared to MCM2-7 and those homohexameric MCM helicase. Upon entry into the S phase, the MCM2-7 DH dissembles into two single hexamer that further form two active CMG helicases with the accessory factors Cdc45 and GINS to advance bidirectional replication forks (33,34). Comparatively, the MCM8/9 complex coordinate the processing of broken

DNA ends and D-loop disruption in homologous recombination repair pathway (13). It is still unknown whether two MCM8/9 hexamers are required for these processes at the same time or whether the assembly of double hexamer of MCM8/9 requires additional loading factors. The only known regulator of MCM8/9 complex is HROB. As an ssDNA binding factor, HROB interacts with MCM8/9 complex and stimulates its helicase activity by enhancing the DNA binding ability of MCM8/9 (20). The potential for HROB to promote double hexamer assembly of MCM8/9 remains to be addressed. Furthermore, the MCM2-7 double hexamer encircles the double-stranded origin DNA. The MCM5-MCM2 interface characterizes as a "gate" that undergoes open and closed conformation change to engage the DNA entry (7). It is still unknown how MCM8/9 complex engage the DNA into the central channel. The interaction of MCM8/9 with DNA require more investigations, especially high-resolution structures of MCM8/9-DNA complex.

Multiple OB hairpins are involved in DNA unwinding. In the central channel of the MCM2-7 complex, several structural elements involved in DNA binding and strand separation were illustrated by a series of CMG structures complexed with different types of DNA(35). The recent studies reported that the OB hairpin loops of MCM3, MCM4, MCM6 and MCM7 forms "dam and diversion tunnel" to block and divert the lagging stand from the leading stand (23). What's more, the highly conserved residue F363 from OB hairpin loop of MCM7 serves as a strand separation pin to unwind forked DNA via making a π-π interaction with DNA(36). For MCM8/9 complex, the OB-hp1s of MCM8 and MCM9 have potential functions in DNA unwinding in the helicase assay. But instead of spirally-configured OB hairpin loops in MCM2-7 complex, the OB-hp1s form a two-layer structure arranged in a 3-fold symmetry manner. In which, three OB-hp1s of MCM9 form the upper layer while three OB-hp1s of MCM8 form the lower layer, which suggests a unique unwound mechanism of MCM8/9. Additionally, highly conserved phenylalanine also could be found on the OB-hp1s of MCM8 and MCM9 (F353 of MCM8, F251 of MCM9) (Figure 2D). It is also interesting to dissect the function of these phenylalanine in strand separation.

Potential DNA-unwinding mode by MCM8/9. Unexpectedly, profound conformational differences information can be drawn from the structural comparison of MCM8/9 with MCM2-7, from which the helicase functional mechanism of MCM8/9 and the potential DNA-unwinding mode by MCM8/9 is implicated. When superimposing the hexameric ring of MCM8/9 to those of MCM2-7-containing complexes based on all the atoms of their CTDs, the NTD ring of MCM8/9 has different rotary shifts (Figure 5B and 5C). These results allow us to propose that there is relative rotation between the NTD and CTD rings of MCM8/9 upon helicase activation. The movie 1 produced from the multibody refinement for the gMCM8/9 also confirms that the flexible C-tier ring of gMCM8/9 shows a repeat rotary motion relative to the N-tier ring. This newly discovered motion exhibits distinct feature in helicase working models. For homohexameric helicases, it is proposed that they translocate along the DNA in a rotary fashion with the six subunits permutating in a spiral staircase (37,38). While the MCM2-7 in the active CMG complex is unsymmetrical and a pumpjack model was proposed, in which the C-tier AAA+ domain nodded up and down like an oil rig pumpjack attached to a stable platform composed of the N-tier ring, Cdc45 and GINS (39). Nevertheless, based on the above observation, we propose here a potential unusual DNA unwinding model for MCM8/9 helicase. In which, the N-tier ring acts like a wedge, inserting into the forked DNA. The MCM8/9 C-tier AAA+ motor domain rotates reciprocally with ATP binding and hydrolysis. These repeat rotations of the CTD tilt the N-C linkers which attached to the N-tier ring and then propel the helicase advance. During unwinding, the hexameric MCM8/9 encircles the leading strand of the forked DNA which passes through its central channel while the lagging strand is extruded from the central channel in a steric extrusion model or through a “gate” inside the NTD of MCM8/9 under tightly control (Figure 6). There are compelling reasons supporting this model. Firstly, the MCM8/9 complex forms a 3:3 heterohexamer in an alternate pattern, which is very unique among helicase complexes. It is possible that three MCM8 or MCM9 subunits bind and hydrolyze ATP simultaneously, which produces great energy to power the rotation of the helicase. Secondly, the lengths of the MCM8 and MCM9 N-C linkers are surprisingly much longer than those in MCM2-7

complex. Our functional study also showed that the long linker is critical for MCM8/9 helicase activity. Shorten its length will impede the relative motion of N-tier and C-tier rings and then dramatically decrease the MCM8/9 helicase activity. This observation is also consistent with previous reports indicating that the N-C linker stabilizes the relative position of NTD and CTD of archaea MCM helicase and permit the movement of the AAA+ motor domain during unwinding. We propose that the long length of the linkers in MCM8/9 are involved in the relative rotation between N- and C-tiers. We also concede that further structural advance and biochemical experiments are needed to test the applicability of current model for MCM8/9 helicase.

In summary, the MCM8/9 complex structures provided in this study will serve as a rich source of information for exploring the mechanism of the DNA helicase which is implicated in homologous recombination. In addition, the structure of MCM8/9 provide information for the analysis of multiple disease mutations in MCM8/9 associated with premature ovarian failure as well as cancers.

Material and methods

Plasmids

The human MCM8/9 genes were obtained by reverse transcription of mRNA from HeLa cells. The gene fragments coding MCM8/9 of *Gallus gallus* were also from corresponding reverse transcripts and the total RNA was extracted from chicken DT40 cell line. Human MCM8_61-840 and MCM9_1-684 were inserted into pFastbac-Dual vector (Invitrogen). *Gallus gallus* MCM8_50-830 and MCM9_1-691 were inserted into pFastbac-1 vector, respectively. All the MCM8 were fused with 6× His tag at their C-terminus for subsequent affinity purification. The genes coding MCM8/9 NTDs and relative mutants were cloned into multiple cloning site I (MCS I) of pRSFDuet-1 vector (Novagen). The mutations were introduced by polymerase chain reaction (PCR).

Human HROB (NM_001171251.3) was obtained by reverse transcription of mRNA from 293T cells. HROB_391-580 was inserted into pET28S-SUMO vector which expresses N-terminal SUMO tag.

Protein expression and purification

To express the core hMCM8/9 and gMCM8/9 complex proteins, recombinant baculoviruses were prepared using the Bac-to-Bac expression system (Invitrogen). The proteins were expressed in *Trichoplusia ni* (BTI-Tn5B1–4, Hi5) insect cells for 60 hours at 27°C before harvest. The cell pellet was collected by centrifugation at 800× g and then freshly frozen by liquid nitrogen and stored at -80°C before use. For purification, the cell pellet was resuspended with lysis buffer (50 mM HEPES, pH 8.0, 350 mM KCl, 50 mM sodium glutamate, 10% glycerol and 30 mM imidazole) and lysed by cell homogenizer (Avestin Emulsiflex C3). Then the lysate was clarified by centrifugation at 16,000× rpm for 30 min. The supernatant was mixed with nickel-NTA resin (Novagen) and kept stirring for 1 hour at 4°C for adequate binding. The beads were then collected into a column and extensively washed with lysis buffer to remove unwanted proteins. Target proteins were then eluted down with elution buffer (50 mM HEPES, pH 7.8, 80 mM KCl, 20 mM sodium glutamate, 5% glycerol and 300 mM imidazole). The proteins were further purified by HiTrap Heparin HP 5ml Column (GE Healthcare) and size exclusion column (Superose 6 Increase 10/300 GL, GE Healthcare) in buffer (20 mM HEPES, pH 7.8, 150 mM KCl, 1 mM DTT and 5% glycerol).

For expression of HROB, 10 mL *E. coli* codon plus strain containing the recombinant pET28a-HROB plasmid was inoculated into 1 L LB medium. The bacteria solution was firstly cultured at 37°C until the $OD_{600} \approx 0.8$ and then cooled down to 16°C. The expression of recombinant protein was induced by IPTG (isopropyl-β-d-thiogalactoside) at a concentration of 0.2 mmol/L and kept expressing for another 20 hours.

The strategies of purification of MCM8/9 mutants and HROB proteins were as the same way as that of core hMCM8/9 complex.

Crystal optimization and data collection

hMCM8/9 was concentrated to about 30 mg/mL and incubated with 1 mM ATPγS. The crystallization experiments were performed by hanging-drop vapor diffusion method at 16°C. 1 μL proteins was mixed with 1 μL reservoir solution. The crystals were produced in 100 mM Bis-Tris propone, pH 6.0-7.5, 600 mM sodium citrate in one

week. The crystals were cryoprotected by 30% glycerol before flash-cooled in liquid nitrogen.

The diffraction data was collected at 100 K on BL17U beam line at Shanghai Synchrotron Radiation Facility (SSRF), China (40). The data were processed by HKL3000 (41).

Crystal structure determination

Obtaining the phase information of human core MCM8/9 diffraction data is rather difficult due to the low resolution. The MCM8/9 NTD dimer structure models were extracted from the human MCM8/9 NTD hexamer model reported before (21) and used as search models for phaser_MR (42). A promising MR solution that forms a very similar NTD hexameric ring was found after adjusting the packing criterion to 15%. The TFZ value was increased to 9.7 from initial 7.5 after a round of rigid_body refine by Phenix.refine (43), indicating a right solution.

A series of predicted models of 8C and 9C were produced by the Rosetta (44). We used the top 5 predicted models for further MR trails. Two of the 9C models gave the single MR solution with TFZ = 8.1 and 9.1, respectively. After another round of refinement, the TFZ was further increased to 12.2. The 8C models were used as the third search model and two of them produced correct single solution, with TFZ = 10.4 and 10.5, respectively. But further attempt to find the MCM8 WHD (PDB: 6L0O) failed without no reasonable solution. The model containing 8N, 9N, 8C (without WHD) and 9C was further refined to R_{work}/R_{free} = 0.4130/0.4476. The flexible regions with very weak electron density were trimmed manually before the final refinement. The final R_{work}/R_{free}=0.3216/0.3981. The accession code for the structure model is 7WI7 in Protein Data Bank.

Cryo-EM sample preparation

The fresh gMCM8/9 proteins were used to prepare cryo-EM sample. Briefly, 0.8 μL crosslinker $BS(PEG)_9$ (Thermo scientific) and 10 μM protein were incubated for 1 hour at 4°C and filtered by Superose 6 in buffer (20 mM HEPES, pH 7.8, 150 mM KCl) after

centrifugation. The fractions were identified by SDS-PAGE and the peak fraction was diluted to about 0.3 mg/mL to prepare cryo-EM samples. 4 μL fresh sample was applied on glow-discharged Cu holey carbon grids (Quantifoil R 1.2/1.3) and incubated 1 min at room temperature. Grids were then blotted for 3 s in 100% humidity at 4°C and plunged frozen in liquid ethane cooled by liquid nitrogen using Vitrobot (Thermo).

Cryo-EM data acquisition

The grids were loaded onto a FEI Arctica transmission electron microscope operated at 200 kV in liquid nitrogen temperatures. Images were recorded on a Gatan K2 camera. Images with a total dose of 50 e^-/ $Å^2$. The total dose was fractionalized to 32 frames equally, corresponding to a physical pixel size of 1.0 Å and a super-resolution pixel size of 0.5 Å. The defocus range was set from -1.7 to -2.2 μm. 2,891 raw movie micrographs were collected using the serialEM software (45).

Single particle data processing

The data was processed by Cryosparc-3.2 (46). All the movies were processed using patch motion correction and patch CTF estimation. A total of 2,134,191 particles were automatically picked by template picker and 1,574,821 particles were produced after the inspect particle picks treatment. Three parallel runs of 2D classification (K = 100) were performed using the data (binned 4) with a box size at 256 pixels. A small subset of the "good" particles with clear secondary-structure features was selected from 2D classification results and were used to generate 4 initial models. The 434,320 particles in the best initial model class were re-extracted with original pixel size and further refined to a 6.7 Å low-resolution reconstruction using non-uniform refinement (New) (47).Four resolution gradient templates generated from non-uniform refinement were used for the multi-reference heterogeneous refinement, resulting in a 4.9 Å global reconstruction (190,119 particles). The NTD map was improved to 3.71 Å with local refinement following by applying a tight N-terminal mask with *C3* symmetry. Local CTF refinement further improved the resolution to 3.67 Å.

The EMDB accession code is EMD-32346 and the PDB entry ID is 7W7P. The image processing and 3D reconstruction steps were illustrated in Figure S1.

RELION-3.1.1 (48) was also used to process the same data, resulting in a 6.6 Å global resolution map and a 4.3 Å N-terminal resolution map. All the movie frames were motion-corrected by MotionCorr2 (49) with a patch alignment of 5×5. CTF parameters were estimated using Gctf (50). 965,175 particles were auto-picked and extracted to 4×binned particles, which rescales the initial box size at 256 pixels to the final box size at 64 pixels. The particles were subjected to 2 rounds of 2D classifications (K = 100). 388,378 particles with clear secondary-structure features were selected for 3D initial model generation. Three parallel runs multi-reference 3D classification (K = 9) were performed and the best class were merged. 144,202 particles were re-extracted at original pixel size and were submitted to the final 3D model auto-refinement. Finally, we yielded a 6.6 Å 3D density map estimated by the gold-standard Fourier shell correlation at a correlation cutoff value of 0.143.

As the density of the gMCM8/9 CTD was blurred, we then used the multi-body refinement program (51) to process the dataset. gMCM8/9 NTD and CTD were subtracted as the first and the second subtracted body respectively using Chimera (52). The corresponding masks were generated using mask_creation in RELION. The multi-body refinement resulted in a 4.1 Å 3D density map of NTD and a 6.7 Å 3D density map of the CTD. It generated 12 components, with each including a discrete number of independently moving bodies. The first three components explained 61.5% of the variance in the data.

Model building and refinement

The gMCM8 NTD and gMCM9 NTD monomer models were predicted through SWISS-MODEL by providing the hMCM8 NTD (PDB: 7DP3) and hMCM9 NTD (PDB: 7DPD) structural template, respectively. Firstly, the predicted structures were used as the initial model directly docking into the EM map. As the gMCM8/9 NTD has a C3 symmetry axis, another two copies were generated by simply horizontal rotating for subsequent rigid body-fitting into the 3D density map with program Chimera (52).

Secondly, use the automatic refinement plugin function in COOT (53). The final model was refined against the corresponding map using PHENIX (43)in real space with secondary structure and geometry restraints. The structures of the gMCM8/9 NTD complex were validated through examination of the Molprobity scores and statistics of the Ramachandran plots. Molprobity scores were calculated as described.

Helicase assays

To prepare the substrate, the oligonucleotide (5'-$(dT)_{40}$GTTTTCCCAGTCACGACG -TTGTAAAACGACGGCCAGTGCC-3') containing a 40 nt region complementary to the M13mp18(+) stand and a 40 nt oligo-dT at the 5′ end was labeled at the 3′ terminus with [α-^{32}P] dCTP (Perkin Elmer) and annealed to the single-stranded DNA M13mp18(20). The DNA substrate were mixed with recombinant proteins as indicated within each 15 μl volume reaction in the helicase buffer (25 mM HEPES, pH 7.5, 1 mM magnesium acetate, 25 mM sodium acetate, pH 5.2, 4 mM ATP, 0.1 mg/ml BSA, 1 mM DTT). To avoid re-annealing, the reaction was supplemented with a 100-fold unlabeled oligonucleotide. The reactions were than incubated at 37 °C for 60 min and stopped adding 1 μl of stop buffer (0.4% SDS, 30 mM EDTA, and 6% glycerol) and 1μl of proteinase K (20 mg/ml, Sigma) into the reaction for another 10 min incubation at 37 °C. The products were separated by 15% polyacrylamide gel electrophoresis in 1× TBE buffer and analyzed by the Amersham typhoon (Cytiva).

Cell survival assay

The CCK8 assay of DT40 cells were performed as described previously (54). Briefly, the MCM8 or MCM9 KO DT40 cell lines were transfected with plasmids containing wildtype MCM8, MCM9 or the indicated mutants by electroporation. 1.0×10^3 cells were seeded into each well of a 96-well plate and treated with a range of concentrations of cisplatin. After 48 h of incubation, the cell viability was measured by the Cell counting kit 8 (CCK8) according to the standard procedures. All experiments were performed at least 3 times.

Declaration of interests

The authors declare no competing interests.

Acknowledgements

We thank the Shanghai Synchrotron Radiation Facility (SSRF) BL17U for X-ray beam time and the Cryo-EM Centre, Southern University of Science and Technology for assistance in cryo-EM data collection. We are grateful to Hongjie Zhang (Institute of Biophysics, CAS) for technical assistance in helicase assay. This work was supported by grants from Shenzhen Science and technology planning project (project No.JCYJ20180307155005435 to Y.L., project No.RCBS20200714114922284 to Z.W.), Pearl River Scholar Team Project of Guangdong Province (project No.2019ZT08Y464 to Y.L.) and National Natural Science Foundation of China (project No.32000860 to Z.W.)

Author contributions

Z.W., J.L. and J.Z. designed and performed the experiments. J.L. expressed, purified and crystallized human MCM8/9 complex, performed X-ray data collection and structure determination. J.Z. and H.L. performed protein purification, cryo-EM sample preparation, data collection and structure model building. Z.W and Y.Z. performed the helicase assay. Z.L performed cell survival assay. Y.L., H.L., Z.W. and J.L. analyzed the data. Z.W. and H.L prepared the figures. Z.W., J.L., H.L. and Y. L. wrote the paper. All authors reviewed the results and approved the final version of the manuscript.

Reference

1. Brosh, R.M., Jr. and Matson, S.W. (2020) History of DNA Helicases. *Genes (Basel)*, **11**.
2. Heyer, W.D., Ehmsen, K.T. and Liu, J. (2010) Regulation of homologous recombination in eukaryotes. *Annu Rev Genet*, **44**, 113-139.

3. Maiorano, D., Lutzmann, M. and Méchali, M. (2006) MCM proteins and DNA replication. *Current opinion in cell biology*, **18**, 130-136.
4. Yuan, Z., Riera, A., Bai, L., Sun, J., Nandi, S., Spanos, C., Chen, Z.A., Barbon, M., Rappsilber, J., Stillman, B. *et al.* (2017) Structural basis of Mcm2-7 replicative helicase loading by ORC-Cdc6 and Cdt1. *Nature structural & molecular biology*, **24**, 316-324.
5. Li, N., Zhai, Y., Zhang, Y., Li, W., Yang, M., Lei, J., Tye, B.K. and Gao, N. (2015) Structure of the eukaryotic MCM complex at 3.8 Å. *Nature*, **524**, 186-191.
6. Abid Ali, F., Renault, L., Gannon, J., Gahlon, H.L., Kotecha, A., Zhou, J.C., Rueda, D. and Costa, A. (2016) Cryo-EM structures of the eukaryotic replicative helicase bound to a translocation substrate. *Nat Commun*, **7**, 10708.
7. Samel, S.A., Fernández-Cid, A., Sun, J., Riera, A., Tognetti, S., Herrera, M.C., Li, H. and Speck, C. (2014) A unique DNA entry gate serves for regulated loading of the eukaryotic replicative helicase MCM2-7 onto DNA. *Genes Dev*, **28**, 1653-1666.
8. Slaymaker, I.M. and Chen, X.S. (2012) MCM structure and mechanics: what we have learned from archaeal MCM. *Sub-cellular biochemistry*, **62**, 89-111.
9. Barry, E.R., Lovett, J.E., Costa, A., Lea, S.M. and Bell, S.D. (2009) Intersubunit allosteric communication mediated by a conserved loop in the MCM helicase. *Proc Natl Acad Sci U S A*, **106**, 1051-1056.
10. Sakakibara, N., Kasiviswanathan, R., Melamud, E., Han, M., Schwarz, F.P. and Kelman, Z. (2008) Coupling of DNA binding and helicase activity is mediated by a conserved loop in the MCM protein. *Nucleic Acids Res*, **36**, 1309-1320.
11. Nishimura, K., Ishiai, M., Horikawa, K., Fukagawa, T., Takata, M., Takisawa, H. and Kanemaki, M.T. (2012) Mcm8 and Mcm9 form a complex that functions in homologous recombination repair induced by DNA interstrand crosslinks. *Mol Cell*, **47**, 511-522.
12. Brewster, A.S., Wang, G., Yu, X., Greenleaf, W.B., Carazo, J.M., Tjajadi, M., Klein, M.G. and Chen, X.S. (2008) Crystal structure of a near-full-length archaeal MCM: functional insights for an AAA+ hexameric helicase. *Proc Natl Acad Sci U S A*, **105**, 20191-20196.
13. Griffin, W.C. and Trakselis, M.A. (2019) The MCM8/9 complex: A recent recruit to the roster of helicases involved in genome maintenance. *DNA Repair (Amst)*, **76**, 1-10.
14. Lutzmann, M., Grey, C., Traver, S., Ganier, O., Maya-Mendoza, A., Ranisavljevic, N., Bernex, F., Nishiyama, A., Montel, N., Gavois, E. *et al.* (2012) MCM8- and MCM9-deficient mice reveal gametogenesis defects and genome instability due to impaired homologous recombination. *Mol Cell*, **47**, 523-534.
15. Morii, I., Iwabuchi, Y., Mori, S., Suekuni, M., Natsume, T., Yoshida, K., Sugimoto, N., Kanemaki, M.T. and Fujita, M. (2019) Inhibiting the MCM8-9 complex selectively sensitizes cancer cells to cisplatin and olaparib. *Cancer Sci*, **110**, 1044-1053.
16. McKinzey, D.R., Gomathinayagam, S., Griffin, W.C., Klinzing, K.N., Jeffries,

E.P., Rajkovic, A. and Trakselis, M.A. (2021) Motifs of the C-terminal domain of MCM9 direct localization to sites of mitomycin-C damage for RAD51 recruitment. *J Biol Chem*, **296**, 100355.

17. Park, J., Long, D.T., Lee, K.Y., Abbas, T., Shibata, E., Negishi, M., Luo, Y., Schimenti, J.C., Gambus, A., Walter, J.C. *et al.* (2013) The MCM8-MCM9 complex promotes RAD51 recruitment at DNA damage sites to facilitate homologous recombination. *Mol Cell Biol*, **33**, 1632-1644.
18. Lee, K.Y., Im, J.S., Shibata, E., Park, J., Handa, N., Kowalczykowski, S.C. and Dutta, A. (2015) MCM8-9 complex promotes resection of double-strand break ends by MRE11-RAD50-NBS1 complex. *Nat Commun*, **6**, 7744.
19. Hustedt, N., Saito, Y., Zimmermann, M., Alvarez-Quilon, A., Setiaputra, D., Adam, S., McEwan, A., Yuan, J.Y., Olivieri, M., Zhao, Y. *et al.* (2019) Control of homologous recombination by the HROB-MCM8-MCM9 pathway. *Genes Dev*, **33**, 1397-1415.
20. Huang, J.W., Acharya, A., Taglialatela, A., Nambiar, T.S., Cuella-Martin, R., Leuzzi, G., Hayward, S.B., Joseph, S.A., Brunette, G.J., Anand, R. *et al.* (2020) MCM8IP activates the MCM8-9 helicase to promote DNA synthesis and homologous recombination upon DNA damage. *Nat Commun*, **11**, 2948.
21. Li, J., Yu, D., Liu, L., Liang, H., Ouyang, Q. and Liu, Y. (2021) Structural study of the N-terminal domain of human MCM8/9 complex. *Structure*.
22. Gambus, A. and Blow, J.J. (2013) Mcm8 and Mcm9 form a dimeric complex in Xenopus laevis egg extract that is not essential for DNA replication initiation. *Cell Cycle*, **12**, 1225-1232.
23. Yuan, Z., Georgescu, R., Bai, L., Zhang, D., Li, H. and O'Donnell, M.E. (2020) DNA unwinding mechanism of a eukaryotic replicative CMG helicase. *Nat Commun*, **11**, 688.
24. Lee, J.K. and Hurwitz, J. (2001) Processive DNA helicase activity of the minichromosome maintenance proteins 4, 6, and 7 complex requires forked DNA structures. *Proc Natl Acad Sci U S A*, **98**, 54-59.
25. Traver, S., Coulombe, P., Peiffer, I., Hutchins, J.R., Kitzmann, M., Latreille, D. and Mechali, M. (2015) MCM9 Is Required for Mammalian DNA Mismatch Repair. *Mol Cell*, **59**, 831-839.
26. Nakane, T. and Scheres, S.H.W. (2021) Multi-body Refinement of Cryo-EM Images in RELION. *Methods in molecular biology (Clifton, N.J.)*, **2215**, 145-160.
27. Liu, Y., Huynh, D.T. and Yeates, T.O. (2019) A 3.8 Å resolution cryo-EM structure of a small protein bound to an imaging scaffold. *Nat Commun*, **10**, 1864.
28. Singh, S.K., Gui, M., Koh, F., Yip, M.C. and Brown, A. (2020) Structure and activation mechanism of the BBSome membrane protein trafficking complex. *eLife*, **9**.
29. Noguchi, Y., Yuan, Z., Bai, L., Schneider, S., Zhao, G., Stillman, B., Speck, C. and Li, H. (2017) Cryo-EM structure of Mcm2-7 double hexamer on DNA suggests a lagging-strand DNA extrusion model. *Proc Natl Acad Sci U S A*, **114**,

E9529-e9538.
30. Kanemaki, M.T. (2013) The dimeric Mcm8-9 complex of Xenopus laevis likely has a conserved function for resistance to DNA damage. *Cell Cycle*, **12**, 1338-1339.
31. Maiorano, D., Cuvier, O., Danis, E. and Mechali, M. (2005) MCM8 is an MCM2-7-related protein that functions as a DNA helicase during replication elongation and not initiation. *Cell*, **120**, 315-328.
32. Kumar, C. and Remus, D. (2016) Eukaryotic replication origins: Strength in flexibility. *Nucleus (Austin, Tex.)*, **7**, 292-300.
33. Li, H. and O'Donnell, M.E. (2018) The Eukaryotic CMG Helicase at the Replication Fork: Emerging Architecture Reveals an Unexpected Mechanism. *BioEssays : news and reviews in molecular, cellular and developmental biology*, **40**.
34. Riera, A., Barbon, M., Noguchi, Y., Reuter, L.M., Schneider, S. and Speck, C. (2017) From structure to mechanism-understanding initiation of DNA replication. *Genes Dev*, **31**, 1073-1088.
35. Yuan, Z. and Li, H. (2020) Molecular mechanisms of eukaryotic origin initiation, replication fork progression, and chromatin maintenance. *The Biochemical journal*, **477**, 3499-3525.
36. Baretić, D., Jenkyn-Bedford, M., Aria, V., Cannone, G., Skehel, M. and Yeeles, J.T.P. (2020) Cryo-EM Structure of the Fork Protection Complex Bound to CMG at a Replication Fork. *Mol Cell*, **78**, 926-940.e913.
37. Enemark, E.J. and Joshua-Tor, L. (2006) Mechanism of DNA translocation in a replicative hexameric helicase. *Nature*, **442**, 270-275.
38. Lyubimov, A.Y., Strycharska, M. and Berger, J.M. (2011) The nuts and bolts of ring-translocase structure and mechanism. *Current opinion in structural biology*, **21**, 240-248.
39. Yuan, Z., Bai, L., Sun, J., Georgescu, R., Liu, J., O'Donnell, M.E. and Li, H. (2016) Structure of the eukaryotic replicative CMG helicase suggests a pumpjack motion for translocation. *Nature structural & molecular biology*, **23**, 217-224.
40. Wang, Q.S., Yu, F., Huang, S., Sun, B., Zhang, K.H., Liu, K., Wang, Z.J., Xu, C.Y., Wang, S.S., Yang, L.F. *et al.* (2015) The macromolecular crystallography beamline of SSRF. *Nucl Sci Tech*, **26**, 12-17.
41. Minor, W., Cymborowski, M., Otwinowski, Z. and Chruszcz, M. (2006) HKL-3000: the integration of data reduction and structure solution - from diffraction images to an initial model in minutes. *Acta Crystallogr D*, **62**, 859-866.
42. McCoy, A.J. (2007) Solving structures of protein complexes by molecular replacement with Phaser. *Acta Crystallogr D Biol Crystallogr*, **63**, 32-41.
43. Afonine, P.V., Grosse-Kunstleve, R.W., Echols, N., Headd, J.J., Moriarty, N.W., Mustyakimov, M., Terwilliger, T.C., Urzhumtsev, A., Zwart, P.H. and Adams, P.D. (2012) Towards automated crystallographic structure refinement with phenix.refine. *Acta Crystallogr D Biol Crystallogr*, **68**, 352-367.
44. Raman, S., Vernon, R., Thompson, J., Tyka, M., Sadreyev, R., Pei, J., Kim, D.,

Kellogg, E., DiMaio, F., Lange, O. *et al.* (2009) Structure prediction for CASP8 with all-atom refinement using Rosetta. *Proteins*, **77 Suppl 9**, 89-99.

45. Mastronarde, D.N. (2005) Automated electron microscope tomography using robust prediction of specimen movements. *J Struct Biol*, **152**, 36-51.
46. Punjani, A., Rubinstein, J.L., Fleet, D.J. and Brubaker, M.A. (2017) cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. *Nat Methods*, **14**, 290-296.
47. Punjani, A., Zhang, H. and Fleet, D.J. (2020) Non-uniform refinement: adaptive regularization improves single-particle cryo-EM reconstruction. *Nat Methods*, **17**, 1214-1221.
48. Scheres, S.H. (2012) RELION: implementation of a Bayesian approach to cryo-EM structure determination. *J Struct Biol*, **180**, 519-530.
49. Zheng, S.Q., Palovcak, E., Armache, J.P., Verba, K.A., Cheng, Y. and Agard, D.A. (2017) MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat Methods*, **14**, 331-332.
50. Zhang, K. (2016) Gctf: Real-time CTF determination and correction. *J Struct Biol*, **193**, 1-12.
51. Nakane, T., Kimanius, D., Lindahl, E. and Scheres, S.H. (2018) Characterisation of molecular motions in cryo-EM single-particle data by multi-body refinement in RELION. *eLife*, **7**.
52. Pettersen, E.F., Goddard, T.D., Huang, C.C., Couch, G.S., Greenblatt, D.M., Meng, E.C. and Ferrin, T.E. (2004) UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem*, **25**, 1605-1612.
53. Emsley, P., Lohkamp, B., Scott, W.G. and Cowtan, K. (2010) Features and development of Coot. *Acta Crystallogr D Biol Crystallogr*, **66**, 486-501.
54. Gao, S., Feng, S., Ning, S., Liu, J., Zhao, H., Xu, Y., Shang, J., Li, K., Li, Q., Guo, R. *et al.* (2018) An OB-fold complex controls the repair pathways for DNA double-strand breaks. *Nat Commun*, **9**, 3925.

Figure 1

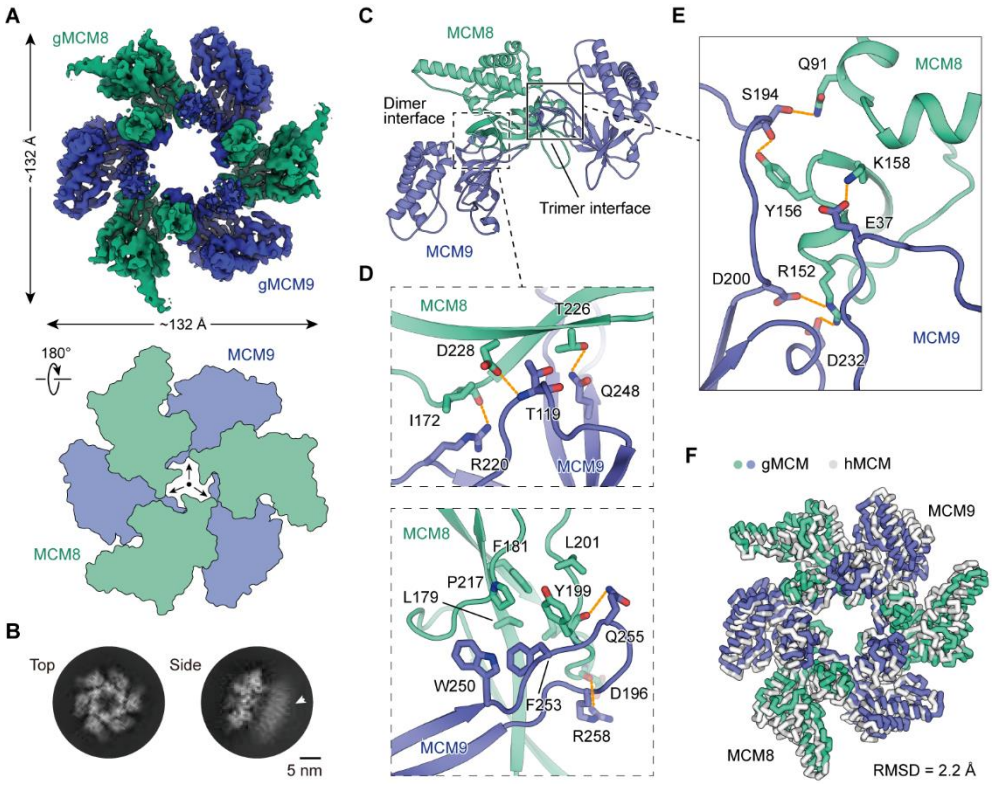

Figure 1. The cryo-EM structure of gMCM8/9 NTD ring. (**A**) Reconstructed cryo-EM map of gMCM8/9 N-tier ring. The diameter of the N-tier ring was measured at ~132 Å. The cartoon presents the 3-fold symmetry axis of the gMCM8/9 hexameric structure. (**B**) Selected 2D classes of gMCM8/9 showing its typical top and side views. The while triangle marks the blurry C-tier ring of gMCM8/9. (**C**) The structure of gMCM8/9 presents in cartoon. The dimer interface and trimer interface were indicated by boxes with dotted line and solid line, respectively. (**D, E**) The dimer interface (D) and trimer interface (E) were mediated by hydrophobic interaction and polar interactions. The interaction details between gMCM8 and gMCM9 in two interfaces are shown in stereo view. (**F**) Structural superposition of NTDs of gMCM8/9 and hMCM8/9. The overall root mean square deviation (RMSD) (2.2 Å) was listed below the structures. The structures were shown as ribbons. gMCM8/9, greencyan and slate; hMCM8/9, grey.

Figure 2

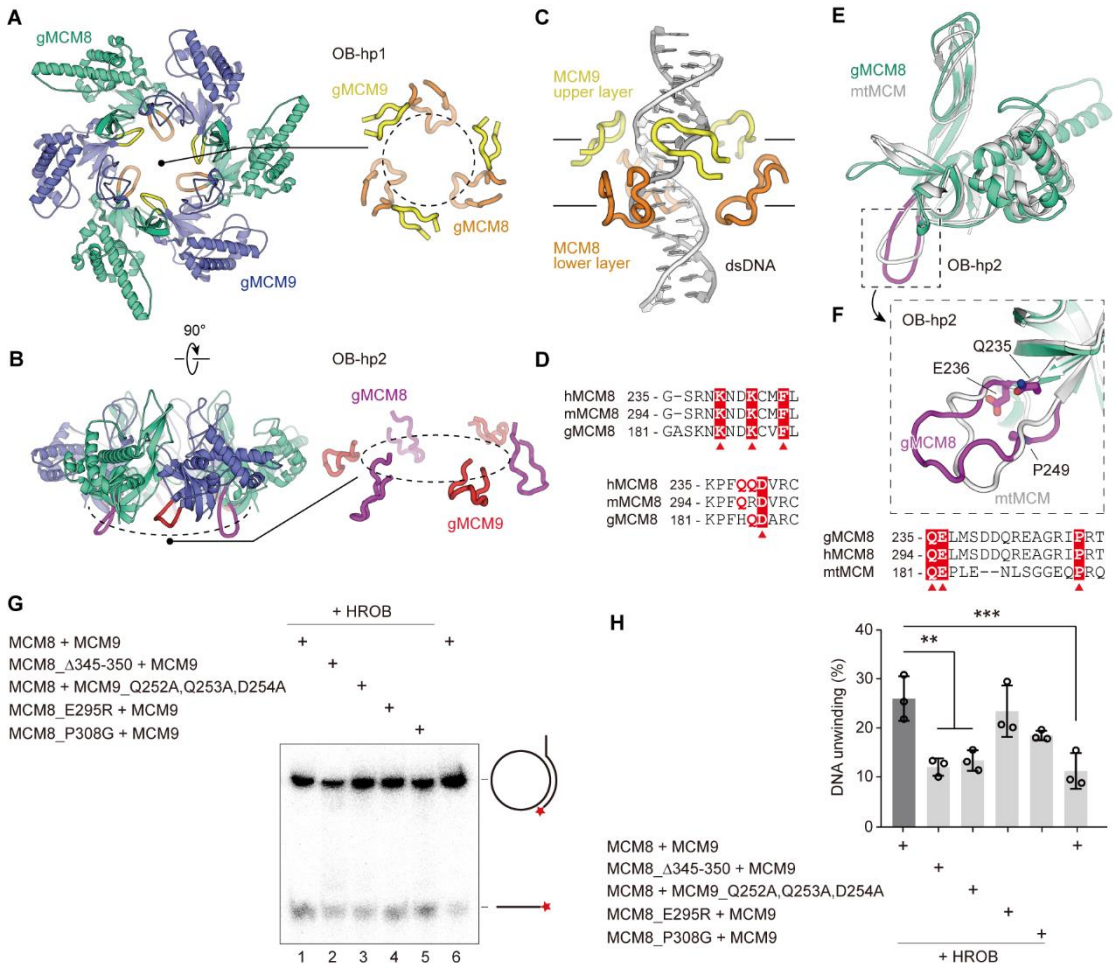

Figure 2. The spatial distribution of OB fold hairpins and their functional roles in DNA unwinding. (**A, B**) Top and side views of the structure of the gMCM8/9 NTD ring. OB-hp1s of gMCM8 and gMCM9 are highlighted in color orange and yellow, respectively (A); while OB-hp2s are highlighted in color purple and red as indicated (B). (**C**) Contacts between OB-hp1s and dsDNA illustrated through fitting a dsDNA fragment into the hexameric gMCM8/9 central channel. While encircling the dsDNA, the OB-hp1s of the gMCM9 form the upper layer and that of gMCM8 constitute the lower layer. (**D**) sequence alignment of the OB-hp1s of MCM8 and MCM9 from different species. h, human; m, mouse; g, chicken (*Gallus gallus*). The highly conserved residues are labeled with red triangle. (**E, F**) Structure superposition of the gMCM8 NTD/ mtMCM NTD (gMCM8, greencyan; mtMCM, PDB: 1LTL, grey). The OB-hp2 was highlighted by colored purple (gMCM8) (E). Structure-based alignment of MCM from different

species are shown below the structure and key residues are shown in stereo view and labeled with red triangles (F). (**G**) Representative autoradiograph of the DNA unwinding reaction conducted using a ^{32}P-labeled ssDNA oligo annealed to the M13mp18 as DNA substrate in the presence HROB with purified MCM8/9 and mutants as indicated. (**H**) Graphical representation of the percentage of DNA unwinding in reactions conducted as in (**G**). The mean±SD of three independent experiments is presented. Statistical analysis was conducted using one-way ANOVA (**$p<0.01$, ***$p<0.001$)

Figure 3

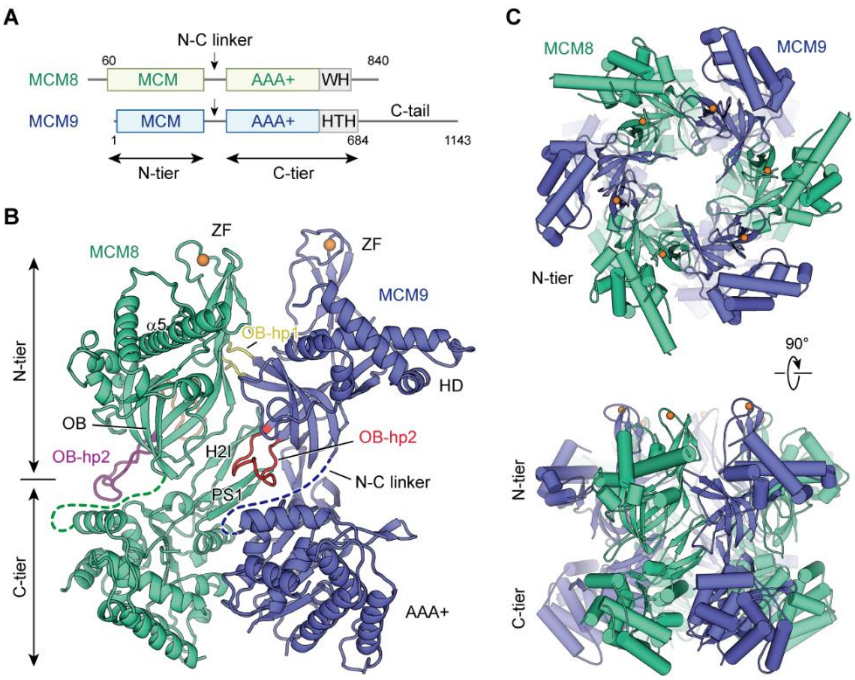

Figure 3. Overall structural features of hMCM8/9 hexamer. (**A**) Domain organization of human MCM8 and MCM9. (MCM8, greencyan; MCM9, slate), N-C linker, linkers connected the NTD and CTD; WH, winged helix domain; HTH, helix-turn-helix; the C-tail domain is not included in this study. (**B**) Fold of the MCM8/9 dimer. Domains are colored as in A. The N-C linkers are shown as dotted lines and the Zinc atoms are shown as orange spheres. The OB hairpins OB-hp1 are highlighted in color orange (MCM8) and yellow (MCM9), while the OB-hp2 are highlighted in color purple (MCM8) and red (MCM9). (**C**) Ribbon diagram showing the top and side views of MCM8/9 hexamer arranged by 3-fold of the heterodimer.

Figure 4

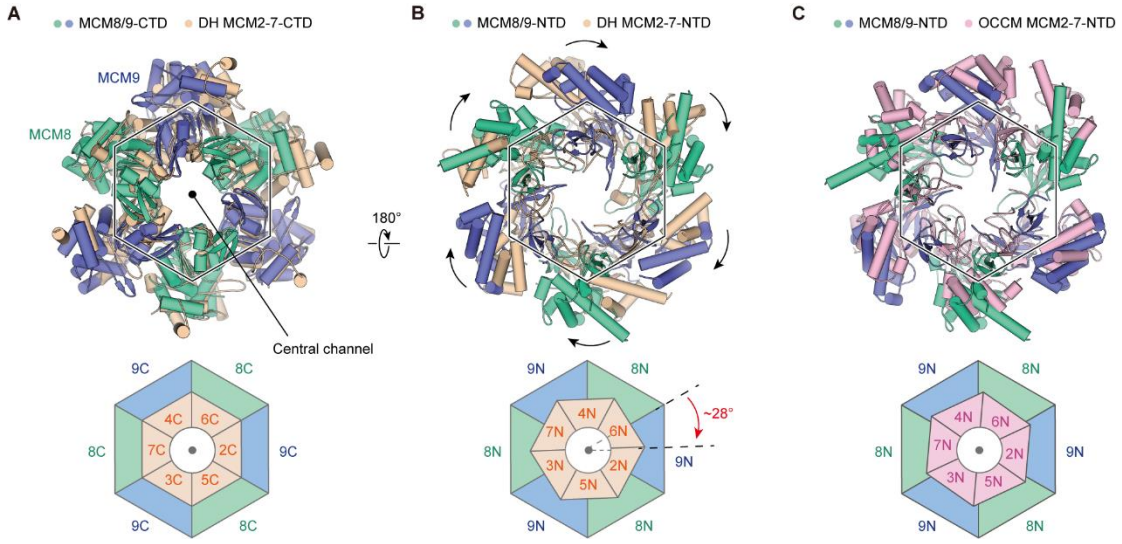

Figure 4. Structural comparison of MCM8/9 with MCM2-7-containing intermediates. (**A, B**) The MCM8/9 hexamer (greencyan and slate) and hexamer from the MCM2-7 double hexamer (wheat) were aligned based on all atoms of their respective CTD. The bottom view (**A**) and top view (**B**) are shown respectively. Note the ~28° clockwise rotation of the MCM8/9 N-tier ring comparing to that of MCM2-7. (**C**) Structural superposition of the MCM8/9 hexamer to the hexamer MCM2-7 (lightpink) from the OCCM by aligning their respective C-tier ring. The N-tier rings comparison are presented here with a slightly rotation beginning at MCM5-MCM2 interface and the following MCM6 and MCM4.

Figure 5

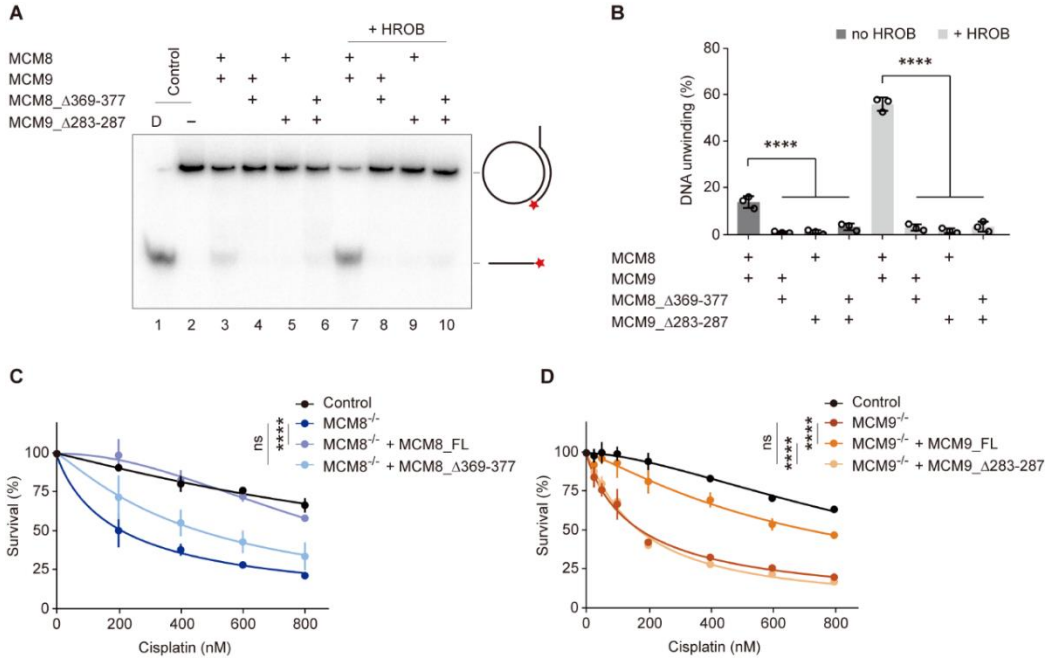

Figure 5. Analysis of the helicase activities and chemoresistance exhibited by the N-C linkers. (**A**) Representative autoradiograph of the DNA unwinding reaction conducted using a ^{32}P-labeled ssDNA oligo annealed to the M13mp18 as DNA substrate in the presence or absence of HROB with purified MCM8/9, MCM8 N-C linker mutant (MCM8Δ369-377) or MCM9 N-C linker mutant (MCM9Δ283-287), either alone or in combination. D—boiled DNA substrate control. (**B**) Graphical representation of the percentage of DNA unwinding in reactions conducted as in (**A**). The mean±SD of three independent experiments is presented. Statistical analysis was conducted using one-way ANOVA (****p<0.0001). (**C**) Survival analysis in DT40 control cell, MCM8 KO cell or cells reconstituted with MCM8 WT or MCM8 N-C linker mutant upon treatment with cisplatin. Cell survival is expressed as a percentage of an untreated control. The mean±SD of three independent experiments is presented. Statistical analysis was conducted on data points at four distinct cisplatin concentrations (200, 400, 600, 800 μM) using Student's t-test (****p<0.0001, at all four concentrations analyzed).

(**D**) Survival analysis in DT40 control cell, MCM9 KO cell or cells reconstituted with MCM9 WT or MCM9 N-C linker mutant upon treatment with cisplatin. Cell survival is represented as in (**C**) and statistical analysis was conducted as in (**C**) (****p<0.0001, at all four concentrations analyzed).

Figure 6

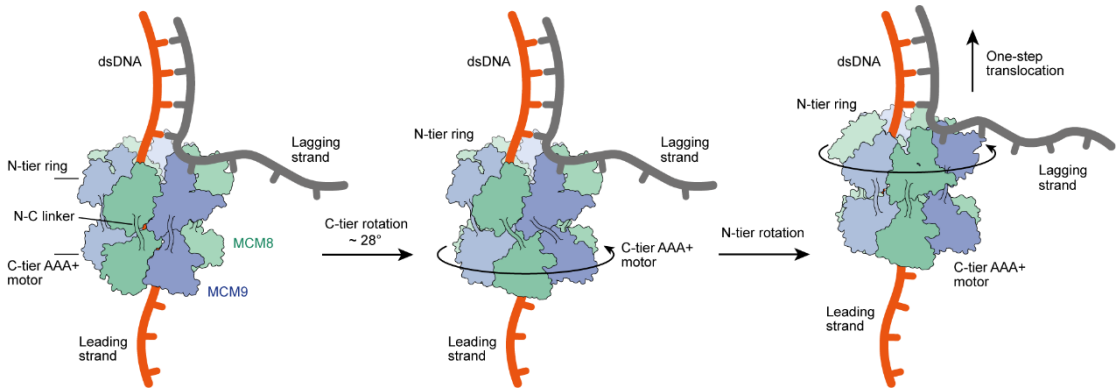

Figure 6. A potential rotation model for MCM8/9 helicase in DNA unwinding: left, the MCM8/9 composed of the N-tier ring and C-tier AAA+ motor ring. The N-tier ring faces to the fork DNA when translocating along the leading strand. Middle, ATP hydrolysis results in twisting and rotating of the CTD ring, which further stretches the N-C linkers. Right. The powered N-C linkers drive N-tier ring upward to separate the duplex DNA. Repeated cycles of ATP hydrolysis result in reciprocal rotation of MCM8/9 to translocate along the leading strand. The black arrows indicate the MCM8/9 helicase movement directions.

Table 1. Cryo-EM 3D reconstruction and refinement of the gMCM8/9 complex.

Table 1 Cryo-EM 3D reconstruction and refinement of the gMCM8/9 complex.	
Data collection and processing	
Magnification	130,000
Voltage (kV)	200
Electron dose (e^{-}/Å^2)	50
Under-focus range (μm)	1.7-2.2
Pixel size (Å)	1.0
Symmetry imposed	C3
FSC threshold	0.143
Relion processing	
Initial particle images (no.)	965k
Final particle images (no.)	144,202
Global map resolution (Å)	6.7
NTD map resolution (Å)	4.1
Cryosparc processing	
Initial particle images (no.)	2,134k
Final particle images (no.)	190,119
Global map resolution (Å)	4.9
NTD map resolution (Å)	3.67
Model composition	
Initial model used (PDB code)	7DP3 and 7DPD
Non-hydrogen atoms	13410
Protein and DNA residues	1698
Ligands	0
R.m.s. deviations	
Bond lengths (Å)	0.002
Bond angels (°)	0.593
Validation	
MolProbity score	1.73
Clashscore	8.36
Rotamers outliers (%)	0.00
Ramachandran plot	
Favored (%)	95.64
Allowed (%)	4.3
Outliers (%)	0.06

Table 2. statistics of data collection and structure refinement of human core MCM8/9

	hMCM8/9
Data collection	
Space group	*$P2_13$*
Cell dimensions	
a, *b*, *c* (Å)	185.62, 185.62, 185.62
α, β, γ (°)	90.00, 90.00, 90.00
Resolution (Å)	50.00-6.60 (6.71-6.60)
R_{sym}	0.061 (0.768)
I / sI	39.8 (4.1)
Completeness (%)	99.7 (100.0)
Redundancy	13.7 (14.3)
CC1/2	0.99
Refinement	
Resolution (Å)	41.50-6.60
No. reflections	
R_{work} / R_{free}	0.3216/0.3981
No. atoms	8839
Protein	8837
Ligand/ion	2
Water	0
B-factors	627.19
Protein	627.13
Ligand/ion	892.24
Water	-
R.m.s. deviations	
Bond lengths (Å)	0.002
Bond angles (°)	0.53
Ramachandran	
Favored (%)	93.20
Allowed (%)	6.44
Outlier (%)	0.36

*Values in parentheses are for highest-resolution shell.