BIOLOGICAL CRYSTALLOGRAPHY

ISSN 1399-0047

The quaternary structure of the eukaryotic DNA replication proteins Sld7 and Sld3

Hiroshi Itou,[a]* Yasuo Shirakihara[a] and Hiroyuki Araki[b]

[a]Structural Biology Center, National Institute of Genetics, Yata 1111, Mishima, Shizuoka 411-8540, Japan, and [b]Division of Microbial Genetics, National Institute of Genetics, Yata 1111, Mishima, Shizuoka 411-8540, Japan. *Correspondence e-mail: hitou@nig.ac.jp

Received 26 March 2015
Accepted 31 May 2015

Edited by K. Miki, Kyoto University, Japan

Keywords: DNA replication; initiator protein; Sld3; Sld7.

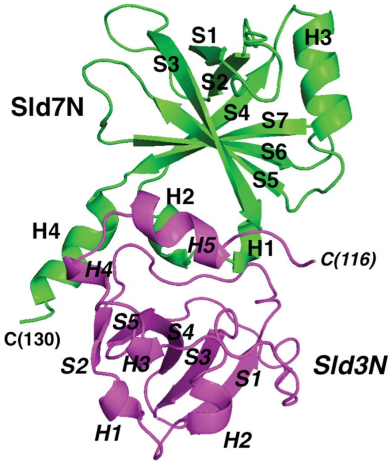
PDB references: C-terminal domain of Sld7, 3x38; N-terminal domain of Sld7 in complex with Sld3, 3x37

Supporting information: this article has supporting information at journals.iucr.org/d

The initiation of eukaryotic chromosomal DNA replication requires the formation of an active replicative helicase at the replication origins of chromosomes. Yeast Sld3 and its metazoan counterpart treslin are the hub proteins mediating protein associations critical for formation of the helicase. The Sld7 protein interacts with Sld3, and the complex formed is thought to regulate the function of Sld3. Although Sld7 is a non-essential DNA replication protein that is found in only a limited range of yeasts, its depletion slowed the growth of cells and caused a delay in the S phase. Recently, the Mdm2-binding protein was found to bind to treslin in humans, and its depletion causes defects in cells similar to the depletion of Sld7 in yeast, suggesting their functional relatedness and importance during the initiation step of DNA replication. Here, the crystal structure of Sld7 in complex with Sld3 is presented. Sld7 comprises two structural domains. The N-terminal domain of Sld7 binds to Sld3, and the C-terminal domains connect two Sld7 molecules in an antiparallel manner. The quaternary structure of the Sld3–Sld7 complex shown from the crystal structures appears to be suitable to activate two helicase molecules loaded onto replication origins in a head-to-head manner.

1. Introduction

Eukaryotic chromosomes must be duplicated exactly once in each cell cycle to ensure the integrity of the genetic information. The formation of an active replicative helicase at the replication origins of chromosomal DNA is a key regulatory step. The formation of an active replicative helicase requires loading of the essential replication proteins Cdc45 and GINS onto the Mcm2–7 helicase core complex at the origins (Moyer *et al.*, 2006). Sld3 is an essential protein that mediates protein interactions during this process. Cdc45 forms a complex with Sld3 to associate with origins (Kamimura *et al.*, 2001), and the association of Sld3 with origins is regulated by Dbf4-dependent kinase (DDK; Masai *et al.*, 2006; Sheu & Stillman, 2006, 2010). Sld3 is a substrate of cyclin-dependent kinase (CDK; Tanaka *et al.*, 2007), and phosphorylation of Sld3 recruits GINS onto origins *via* interaction with Dpb11 (Muramatsu *et al.*, 2010). Sld3 is conserved in yeast and fungi, and treslin, also known as Ticrr, has been identified as the functional counterpart of Sld3 in metazoans (Kumagai *et al.*, 2010; Sansam *et al.*, 2010). Although Sld3 and treslin differ in their molecular sizes and amino-acid sequences, they both have a homologous region called the Sld3/treslin domain in the middle part of the molecule. It has been suggested that this domain is important for their common function (Sanchez-Pulido *et al.*, 2010). Our previous study demonstrated that the Sld3/treslin domain from yeast Sld3 is responsible for binding to Cdc45 (the Cdc45-binding domain of Sld3; Sld3-CBD), and



the domain has a compact rhombus-shaped structure that seems to be suitable for binding to the Cdc45 molecule. Our previous study provided the structural basis for the essential Sld3–Cdc45 interaction (Itou *et al.*, 2014).

Sld7 was identified as another binding partner of Sld3. This protein forms a complex with Sld3 throughout the cell cycle, and associates with and dissociates from origins in an Sld3-dependent manner (Tanaka, Umemori *et al.*, 2011). Although Sld7 is not essential for cell growth, depletion of *sld7* slows cell growth and delays the S phase (Tanaka, Umemori *et al.*, 2011). Sld7 binds to a domain in the N-terminal side of Sld3-CBD, and the same study suggested that Sld7 plays a regulatory role during the initiation of DNA replication by stabilizing Sld3 molecules and reducing their binding affinity to Cdc45 through the formation of an Sld3–Sld7 complex. Sld7 proteins are found in a limited range of yeasts related to *Saccharomyces*, such as *Zygosaccharomyces*, *Kluyveromyces* and *Pichia* (Tanaka, Umemori *et al.*, 2011). However, a recent study showed that the Mdm2-binding protein (MTBP) has a similar role to Sld7 in humans (Boos *et al.*, 2013). MTBP regulates the cellular level of the p53 protein by binding to Mdm2, an E3 ubiquitin ligase of p53 (Brady *et al.*, 2005). Boos *et al.* (2013) demonstrated the distinct function of MTBP in the DNA replication process, irrespective of the p53 status of a cell. MTBP interacts with treslin throughout the cell cycle, and its depletion resulted in slower replication and a longer S phase, which was similar to the depletion of Sld7 in yeast. Furthermore, the MTBP-binding site is located immediately to the N-terminal side of the Sld3/treslin domain in treslin, which is the same as the Sld7-binding site in Sld3 (Boos *et al.*, 2013). The functions of Sld7 and MTBP in DNA replication initiation are still enigmatic; however, evidence suggests their functional relatedness during the process and the importance of their steric configuration in the protein complex.

In this study, we determined the tertiary structure of yeast Sld7 in complex with Sld3 to gain an insight into the function of Sld7. The crystal structures showed that Sld7 forms a heterodimer with Sld3 *via* their N-terminal domains, and two Sld7 molecules form a homodimer using their C-terminal domains. The resulting quaternary structure of the Sld3–Sld7 complex appears to be suitable to attach Cdc45 and GINS onto each of the two Mcm2–7 molecules aligned on a replication origin in a head-to-head orientation.

2. Experimental procedures

2.1. Protein preparation and crystallization

The N-terminal domains of *Z. rouxii* Sld3 (Sld3N; residues 1–115) and Sld7 (Sld7N; residues 1–155) were co-expressed in *Escherichia coli* BL21 (DE3) cells *via* a pETDuet-1 vector (Merk Millipore, Darmstadt, Germany). A hexahistidine tag with a linker comprising Leu and Glu was connected to the C-terminus of Sld3N to allow purification using Ni-affinity chromatography. Selenomethionine-substituted proteins were prepared for crystal structure determination using the single anomalous diffraction (SAD) method. The proteins were overexpressed in cells cultured in M9 medium supplemented with 25 µg ml^{-1} selenomethionine and 0.5 m*M* IPTG at 37°C. The proteins were purified using a HisTrap Crude FF column (GE Healthcare Life Science, Piscataway, New Jersey, USA) equilibrated with a buffer consisting of 20 m*M* Tris–HCl pH 7.5, 1.0 *M* NaCl, 20 m*M* imidazole, 10%(*v*/*v*) glycerol. The bound proteins were eluted with a buffer consisting of 20 m*M* Tris–HCl pH 7.5, 1.0 *M* NaCl, 0.5 *M* imidazole, 10%(*v*/*v*) glycerol. The eluted proteins were purified using a Superdex 200 pg 16/60 column (GE Healthcare Life Sciences) equilibrated with a buffer consisting of 20 m*M* Tris–HCl pH 7.5, 0.3 *M* NaCl, 10%(*v*/*v*) glycerol. Sld7N was co-purified with Sld3N as a complex. The fractions containing the Sld3N–Sld7N complex were collected and dialyzed against a buffer consisting of 10 m*M* Tris–HCl pH 7.5, 0.2 *M* NaCl, 1 m*M* DTT and concentrated to 8.8 mg ml^{-1} using an ultrafiltration device (Vivaspin, Sartorius Stedim Biotech GmbH, Göttingen, Germany) for crystallization. Crystallization experiments were performed using the vapour-diffusion method, and the best crystals were obtained using a solution consisting of 0.1 *M* CHES–NaOH pH 8.8, 0.8 *M* sodium citrate at 20°C. The crystals took one month to grow.

The C-terminal domain of *S. cerevisiae* Sld7 (Sld7C; residues 178–257) was overexpressed in *E. coli* cells from a pET-26b(+) expression vector (Merck Millipore). A hexahistidine tag with a linker comprising Leu and Glu was connected to the C-terminus of Sld7C to allow purification using Ni-affinity chromatography. The protein was overexpressed in *E. coli* BL21 (DE3) cells and purified using a HisTrap Crude FF column equilibrated with a buffer consisting of 20 m*M* Tris–HCl pH 7.5, 1.0 *M* NaCl, 20 m*M* imidazole, 10%(*v*/*v*) glycerol. The bound proteins were eluted with a buffer consisting of 20 m*M* Tris–HCl pH 7.5, 1.0 *M* NaCl, 0.5 *M* imidazole, 10%(*v*/*v*) glycerol. The eluted proteins were purified using a Superdex 75 pg 16/60 column (GE Healthcare Life Sciences) equilibrated with a buffer consisting of 20 m*M* Tris–HCl pH 7.5, 0.3 *M* NaCl, 10%(*v*/*v*) glycerol. The Sld7C-containing fractions were collected, dialyzed against a buffer consisting of 10 m*M* Tris–HCl pH 7.5, 0.3 *M* NaCl, 1 m*M* DTT and concentrated to 85 mg ml^{-1} for crystallization. Crystallization experiments were performed using the vapour-diffusion method, and the best crystals were obtained in a solution consisting of 0.1 *M* acetate–NaOH pH 4.5, 0.1 *M* $LiSO_4$, 12.5%(*w*/*v*) PEG 8K at 20°C. The crystals took several weeks to grow.

2.2. Data collection and structure determination

X-ray intensity data for Sld3N–Sld7N complex structure determination were collected from a cryocooled crystal [supplemented with 20%(*v*/*v*) glycerol as a cryoprotectant] on the BL-17A structural biology beamline at Photon Factory (KEK), Tsukuba, Japan. SAD data were collected at a wavelength of 0.97865 Å and processed using *XDS* (Kabsch, 2010). Structure determination and atomic model building were performed using the SAD method with *CRANK* (Ness *et al.*, 2004) in the *CCP*4 suite (Winn *et al.*, 2011). The program

Table 1
Summary of data-collection and refinement statistics.

Values in parentheses are for the outermost resolution shell.

	Sld37–Sld3N complex	Sld7C	
	Se-SAD (peak)	Se-SAD (peak)	Native
Data collection			
Wavelength (Å)	0.97865	0.97921	1.00000
Resolution range (Å)	40.0–2.35 (2.48–2.35)	50.0–1.90 (1.93–1.90)	50.0–1.80 (1.83–1.80)
a, b, c (Å)	125.7, 37.9, 98.4	34.1, 60.5, 93.1	
α, β, γ (°)	90.0, 126.8, 90.0	90.0, 94.1, 90.0	
Space group	$C2$	$P2_1$	
No. of unique reflections	15781	29943	35136
Completeness (%)	99.7 (99.7)	99.0 (100)	98.5 (99.8)
Multiplicity	5.0 (4.6)	7.5 (7.6)	5.6 (5.5)
$\langle I/\sigma(I)\rangle$	14.7 (3.3)	40.9 (6.0)	36.8 (4.2)
$R_{\rm merge}$† (%)	5.9 (37.8)	4.4 (33.0)	4.4 (40.1)
Refinement			
Resolution range (Å)	20.0–2.35	—	34.0–1.80
No. of reflections, working set	15754	—	34602
R factor‡ (%)	20.1	—	21.1
$R_{\rm free}$§ (%)	24.4	—	23.4
Total no. of non-H atoms			
Nonsolvent	2020	—	2559
Solvent	81	—	474
R.m.s. deviation from standard values			
Bonds (Å)	0.009	—	0.007
Angles (°)	1.308	—	1.029
Ramachandran plot¶			
Favoured (%)	97.1	—	99.3
Outliers (%)	2.93	—	0.66

† $R_{\rm merge} = \sum_{hkl}\sum_i |I_i(hkl) - \langle I(hkl)\rangle| / \sum_{hkl}\sum_i I_i(hkl)$, where $\langle I(hkl)\rangle$ is the mean intensity of symmetry-equivalent reflections. ‡ $R = \sum_{hkl} ||F_{\rm obs}| - |F_{\rm calc}|| / \sum_{hkl} |F_{\rm obs}|$, where $F_{\rm obs}$ and $F_{\rm calc}$ are the observed and calculated structure-factor amplitudes, respectively. § $R_{\rm free}$ was calculated as for the R factor but using only an unrefined subset of reflection data (5%). ¶ The Ramachandran plot was calculated using *RAMPAGE* (Lovell *et al.*, 2003).

was able to build 80% of the atomic model for the Sld3N–Sld7N complex, and the remainder of the model was built manually using *Coot* (Emsley & Cowtan, 2004). The model was refined using the SAD data to 2.35 Å resolution with *LAFIRE* (Yao *et al.*, 2006) and *phenix.refine* (Afonine *et al.*, 2012). Summaries of data collection, phasing and model refinement are given in Table 1.

X-ray intensity data for Sld7C structure determination were also collected from a cryocooled crystal [supplemented with 30%(*v*/*v*) glycerol as a cryoprotectant] on BL-17A at Photon Factory. As the size of the crystal was large enough, the SAD data for substructure analysis and the native data for structure refinement were collected from a single crystal at wavelengths of 0.97921 and 1.00000 Å, respectively. The data were processed using *HKL*-2000 (Otwinowski & Minor, 1997). Structure determination and atomic model building were performed using the SAD method with *AutoSol* (Terwilliger *et al.*, 2009) in the *PHENIX* suite (Adams *et al.*, 2010). The model was refined using the native data to 1.8 Å resolution using *LAFIRE* (Yao *et al.*, 2006) and *phenix.refine* (Afonine *et al.*, 2012). Summaries of data collection, phasing and model refinement are given in Table 1.

3. Results

Our initial attempts to crystallize full-length Sld7 in complex with Sld3 failed. Amino-acid sequence analyses and secondary-structure prediction using the *Phyre*2 server (Kelley & Sternberg, 2009) suggested that Sld7 comprises two distinct structural domains, an N-terminal $\alpha\beta$ domain and a C-terminal α-helical domain, connected *via* an unstructured linker region containing a presumed helix (Fig. 1). Thus, in this study each domain was crystallized separately. We obtained crystals of the N-terminal domain of *Z. rouxii* Sld7 (Sld7N; residues 1–155) in complex with the N-terminal domain of *Z. rouxii* Sld3 (Sld3N; residues 1–115), and of the C-terminal domain of *S. cerevisiae* Sld7 (Sld7C; residues 178–257).

3.1. The N-terminal domain of Sld7 forms a heterodimer with Sld3

The N-terminal domain of Sld7 binds to the N-terminal region of Sld3 (Tanaka, Umemori *et al.*, 2011). These proteins only crystallized when co-expressed in cells and purified as a complex (see §2). The crystal structure of Sld7N in complex with Sld3N was determined at 2.35 Å resolution, and atomic models of Sld7N (Met1–Ile130) and Sld3N (Trp5–Asp115) were obtained. The Sld7N molecule comprises four α-helices and seven β-sheets, forming a compact β-barrel structure (the secondary structures were analyzed by *DSSP*; Joosten *et al.*, 2011; Fig. 2*a*). The Sld3N molecule comprises five α-helices and β-sheets. The crystal structure showed that the structured region occupies only one third of the Sld3N molecule. The Sld3 proteins of some species contain long insertions or deletions. Such alterations occur in the remaining unstructured region, and most of the amino acids involved in the hydrophobic interaction stabilizing the tertiary structure of the Sld3N molecule are conserved among the Sld3 proteins (Fig. 1*b*), suggesting that the core structure is shared among these Sld3 proteins.

The crystal structure showed that the Sld7N and Sld3N molecules form a heterodimer. The amino acids in the S1–S2 loop, S2, S3, S4, S5 and H5 of Sld3N are involved in binding to Sld7 (Fig. 1*b*). This result is consistent with a previous study which showed that *S. cerevisiae* Sld3 has at least two distinct binding sites for Sld7 at amino-acid positions 1–90 and 91–120 (Tanaka, Umemori *et al.*, 2011). The crystal structure showed that although these binding sites were placed close to each other in the tertiary structure (Fig. 2*b*), they are connected by a long unstructured loop that shows the most obvious variation in length among the Sld3 proteins (Fig. 1*b*). Thus, binding to Sld7N appeared to be necessary to stabilize the conformation of Sld3N. The Sld7-binding surface of Sld3N is significantly negatively charged, and Sld7N binds to the acidic

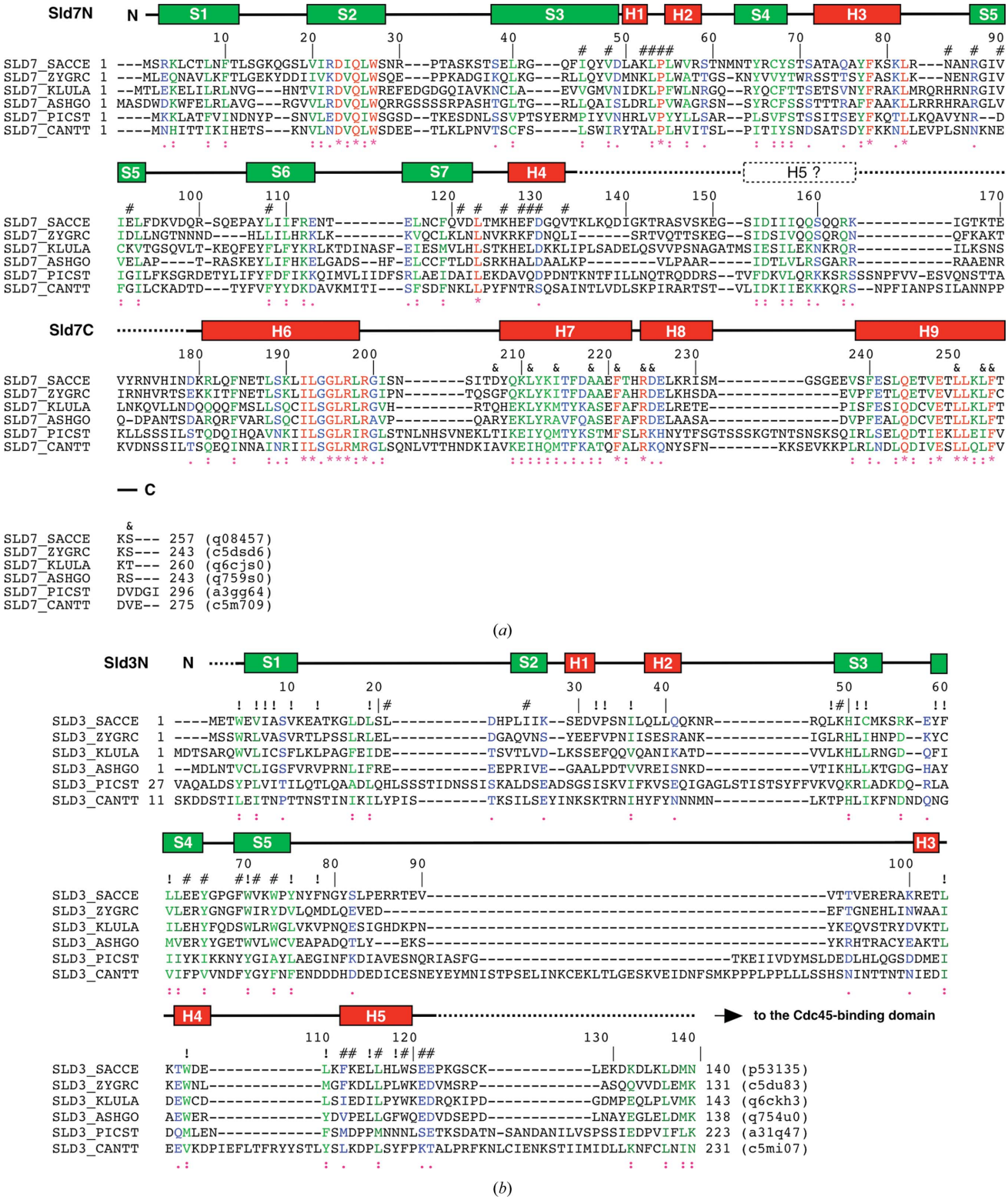


Figure 1
Amino-acid sequence alignments of Sld7 (*a*) and the N-terminal region of Sld3 (*b*) from yeasts (SACCE, *Saccharomyces cerevisiae*; ZYGRC, *Zygosaccharomyces rouxii*; KLULA, *Kluyveromyces lactis*; ASHGO, *Ashbya gossypii*; PICST, *Pichia stipitis*; CANTT, *Candida tropicalis*). The alignments were generated using *ClustalW* (Larkin *et al.*, 2007). Hyphens indicate deletions. The numbers beside the sequences indicate the residue numbers of each protein, and the accession numbers of the sequences are shown in parentheses after the sequences. The residue numbers of *S. cerevisiae* Sld7 are shown above the sequence. The conserved residues (*) and those with conservative changes (: and .) are indicated below the sequences. The secondary structures of Sld7N and Sld3N from *Z. rouxii* and Sld7C from *S. cerevisiae* are indicated above the sequence. The hydrophobic residues involved in intramolecular hydrophobic core formation of Sld3N are indicated by '!' above the sequence. Residues whose side chains are involved in the interaction between Sld7N and Sld3N are indicated by '#' above the sequence. Residues whose side chains are involved in the interaction between Sld7C subunits to form a homodimer are indicated by '&' above the sequence.

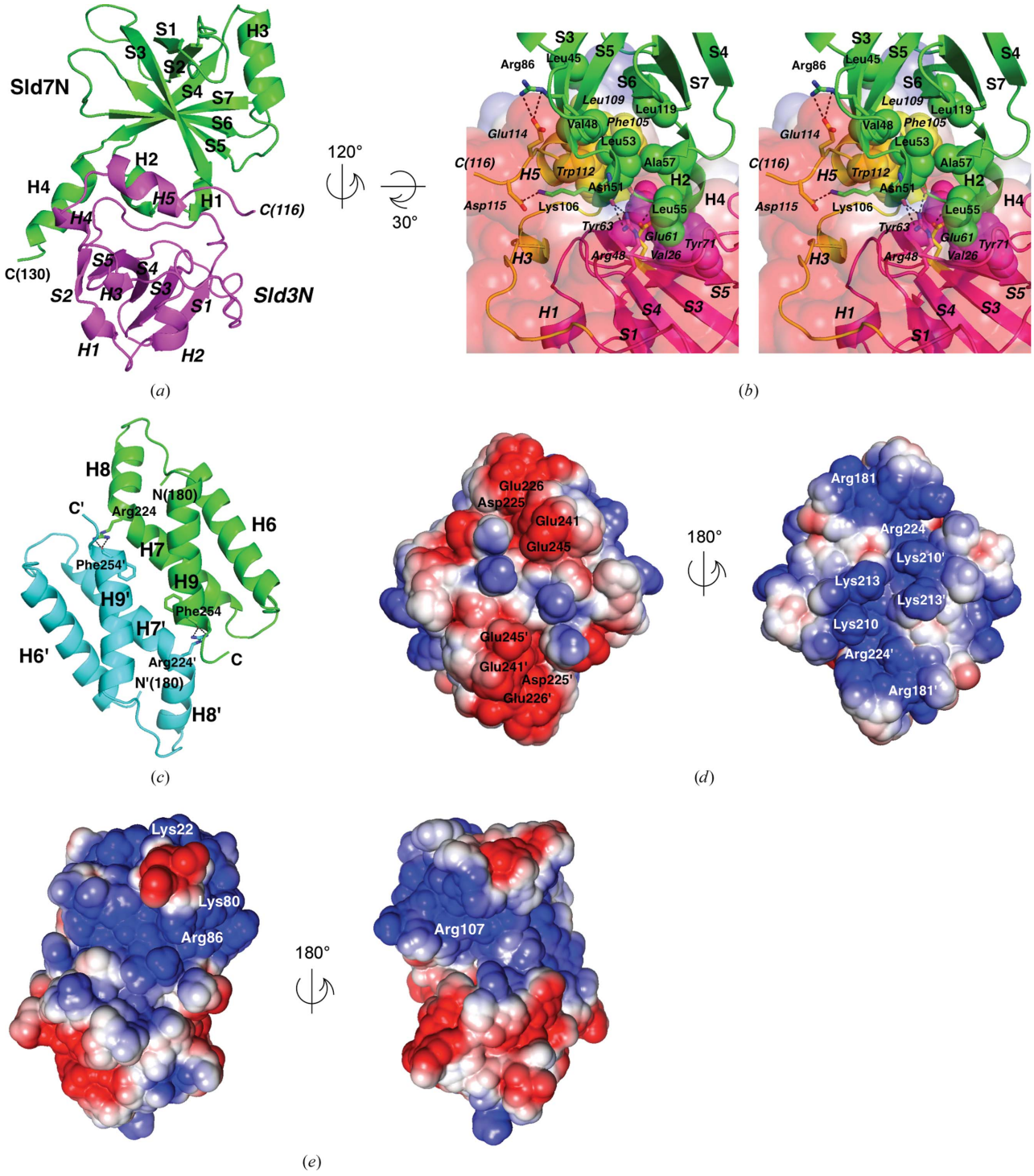


Figure 2
Crystal structures of the Sld7N–Sld3N complex from *Z. rouxii* and of Sld7C from *S. cerevisiae*. (*a*) Overall view of the Sld7N–Sld3N complex. The green and magenta ribbons show the Sld7N and Sld3N proteins, respectively. All of the figures representing tertiary structures in this paper were generated using *PyMOL* (Schrödinger, New York, USA). (*b*) Close-up stereoview of the interaction between Sld7N and Sld3N after the given rotations from (*a*). The colour keys are the same as in (*a*). In this panel, the region in Sld3N corresponding to that of residues 90–120 in *S. cerevisiae* is shown in yellow. The surface charge of Sld3N is also indicated. The surface electric potential representation was calculated using *APBS* (Baker *et al.*, 2001). Blue and red show positively and negatively charged areas, respectively. The conserved side chains involved in the hydrogen-bond network between the Sld7 and Sld3 molecules are shown as stick models. Glutamate Glu61, which corresponds to the Glu63 that is substituted in the *sld3–6* mutant, is shown as a yellow stick model. Broken lines indicate important hydrogen bonds. The side chains involved in the hydrophobic interaction between Sld7 and Sld3 are shown as space-filling models. (*c*) Overall view of the Sld7C homodimer. The two Sld7C subunits are shown in cyan and green ribbons, respectively. The side chains of the perfectly conserved Arg224 and Phe254 are shown as stick models. Dashed lines indicate the hydrogen bonds formed between amino acids. (*d*) Surface-charge representation of the Sld7C homodimer in the same orientation as in (*c*) (left) and after the indicated rotation (right). The conserved acidic and basic residues are indicated. (*e*) Surface-charge representation of the Sld7N–Sld3N complex in the same orientation as in (*a*) (left) and after the indicated rotation (right). The conserved basic residues in the N-terminal domain of Sld7 are indicated.

hollow formed by the two Sld7-binding sites of Sld3N (Fig. 2*b*). Five salt bridges and 13 hydrogen bonds (11 of the hydrogen bonds were observed among the side-chain and the main-chain atoms) were observed between the molecules (analyzed by *PISA*; Krissinel & Henrick, 2007). The amino-acid alignments showed that many of the residues involved in the interactions, including Arg48, Glu114 and Asp115 in Sld3N and Arg86 and Lys106 in Sld7N (shown in Fig. 2*b*), are conserved or changed conservatively among the Sld3 and Sld7 proteins (Fig. 1). The amino acid mutated in *sld3–6*, which shows decreased interaction with Sld7 in *S. cerevisiae* (Glu63Gly; Kamimura *et al.*, 2001; Tanaka, Umemori *et al.*, 2011), corresponds to Glu61 of Sld3N (Fig. 1*b*), and this amino acid is in the binding interface with Sld7N and is involved in the conserved hydrogen-bond network (Fig. 2*b*). Thus, this mutation may hamper hydrogen bonding to Sld7 and disturb the secondary structure of Sld3 at the centre part of the binding interface. On the other hand, many of the hydrophobic amino acids involved in forming the heterodimer are also conserved or changed conservatively among yeasts (Leu45, Val48, Leu53, Leu55, Leu57 and Leu119 of Sld7N and Val26, Tyr63, Ile69, Tyr71, Phe105, Leu109 and Trp112 of Sld3N; Figs. 1 and 2*b*). These results suggest that all Sld3 and Sld7 proteins form a complex similar to the crystal structure obtained in this study.

3.2. The C-terminal domain of Sld7 forms a homodimer in an antiparallel manner

The crystal structure of *S. cerevisiae* Sld7C was determined at 1.8 Å resolution and an atomic model from Asp178 to Ser257 was obtained. The Sld7C subunit comprises four α-helices (H6–H9), and two subunits form a homodimer in an antiparallel manner. The helices H7 and H9 are the main interface for dimer formation (Fig. 2*c*). The N^{η} atom of the perfectly conserved Arg224 in helix H8 forms a hydrogen bond to the main-chain O atom of the perfectly conserved Phe254 in helix H9 of another subunit (Figs. 2*c* and 1*a*). The N^{ε} atom of the side chain of the Arg224 side chain also forms a salt bridge with Asp207 from another subunit in Sld7C; however, Asp207 varies among the Sld7 proteins (Fig. 1*a*). The crystal structure showed that the hydrophobic core formed between the subunits is important to stabilize the dimer. The hydrophobic amino acids in the interface helices that stabilize the dimer structure are especially well conserved among the Sld7 proteins (Leu211, Ile215, Ala218, Phe221, Leu250, Leu253 and Phe254; Fig. 1*a*), suggesting that all Sld7 molecules form a similar homodimeric structure through their C-terminal domains.

The crystal structure showed that the Sld7C homodimer exhibits a bipolar surface-charge distribution. The side of helix H9 is negatively charged because of the conserved acidic amino acids in helices H8 and H9, while the opposite side of the molecule is significantly positively charged because of the conserved basic amino acids in helices H6 and H7 (Figs. 2*d* and 1*a*). Thus, we examined the possibility that the surface charges mediate the interaction between the Sld7C homodimer and the complex comprising the N-terminal domains of Sld3 and Sld7, because the crystal structure showed that the N-terminal domain of Sld7 is significantly positively charged (Fig. 2*e*). The purified Sld7C protein was mixed with the purified N-terminal domains of *S. cerevisiae* Sld7 (amino acids 1–141) in complex with the N-terminal domain of Sld3 (amino acids 1–125), and the apparent molecular size of the mixture was analyzed by size-exclusion chromatography. The result showed that there was no obvious interaction between these two molecules under these conditions [20 m*M* Tris–HCl pH 7.5, 0.1 *M* NaCl, 10%(*v*/*v*) glycerol]. We cannot rule out the possibility that direct interaction between the N-terminal and C-terminal domains of Sld7 through the charged areas occurs under physiological conditions. Such a unique charge distribution on the Sld7C surface implied that those charged areas act as the binding interface to form a complex with the other domains or replication proteins involved in pre-initiation complex formation. Sld3 has been suggested to bind to DDK-phosphorylated Mcm proteins (Heller *et al.*, 2011; Tanaka, Nakato *et al.*, 2011; Yabuuchi *et al.*, 2006); thus, it is possible that the positively charged side of the Sld7C homodimer faces toward Mcm and the acidic side may contribute to binding to the other replication proteins.

4. Discussion

Sld7 binds to Sld3 throughout the cell cycle (Tanaka, Umemori *et al.*, 2011). The crystal structure showed that Sld7 comprises two compact domains. The N-terminal domain forms a complex with the N-terminal domain of Sld3, and two Sld7 molecules form a homodimer *via* their C-terminal domains. These results indicated that two Sld3–Sld7 molecules form a dimer in which two Sld3 molecules are connected through the Sld7 dimer (Fig. 3*a*). The molecular mass of the *S. cerevisiae* Sld3–Sld7 complex calculated from the amino-acid sequences is 106 kDa (those of Sld3 and Sld7 monomers are 77 and 29 kDa, respectively), and the result of a glycerol gradient sedimentation analysis showed that the Sld3–Sld7 complex purified from yeast cells accumulates in the fraction corresponding to 180 kDa, while each protein prepared separately accumulated in fractions corresponding to 65 kDa (Tanaka, Umemori *et al.*, 2011). Thus, the quaternary structure indicated by the crystal structures is consistent with the results of the sedimentation analysis. In this model, because two C-terminal domains of Sld7 bind in an antiparallel manner, two Sld3 molecules are arranged symmetrically in the complex (Fig. 3*a*). Such a steric configuration of Sld3 molecules appears to be suitable to ensure attachment of the essential replication proteins Cdc45 and GINS onto both of the Mcm2–7 helicase core complexes on origins in a head-to-head orientation (Fig. 3*b*).

To initiate DNA replication, two active replicative helicases are generated from a pre-initiation complex to progress the replication forks bidirectionally from origins. In eukaryotic cells, because DNA replication initiates from multiple origins scattered on chromosomal DNA, such bidirectional fork progression from origins is thought to be important to

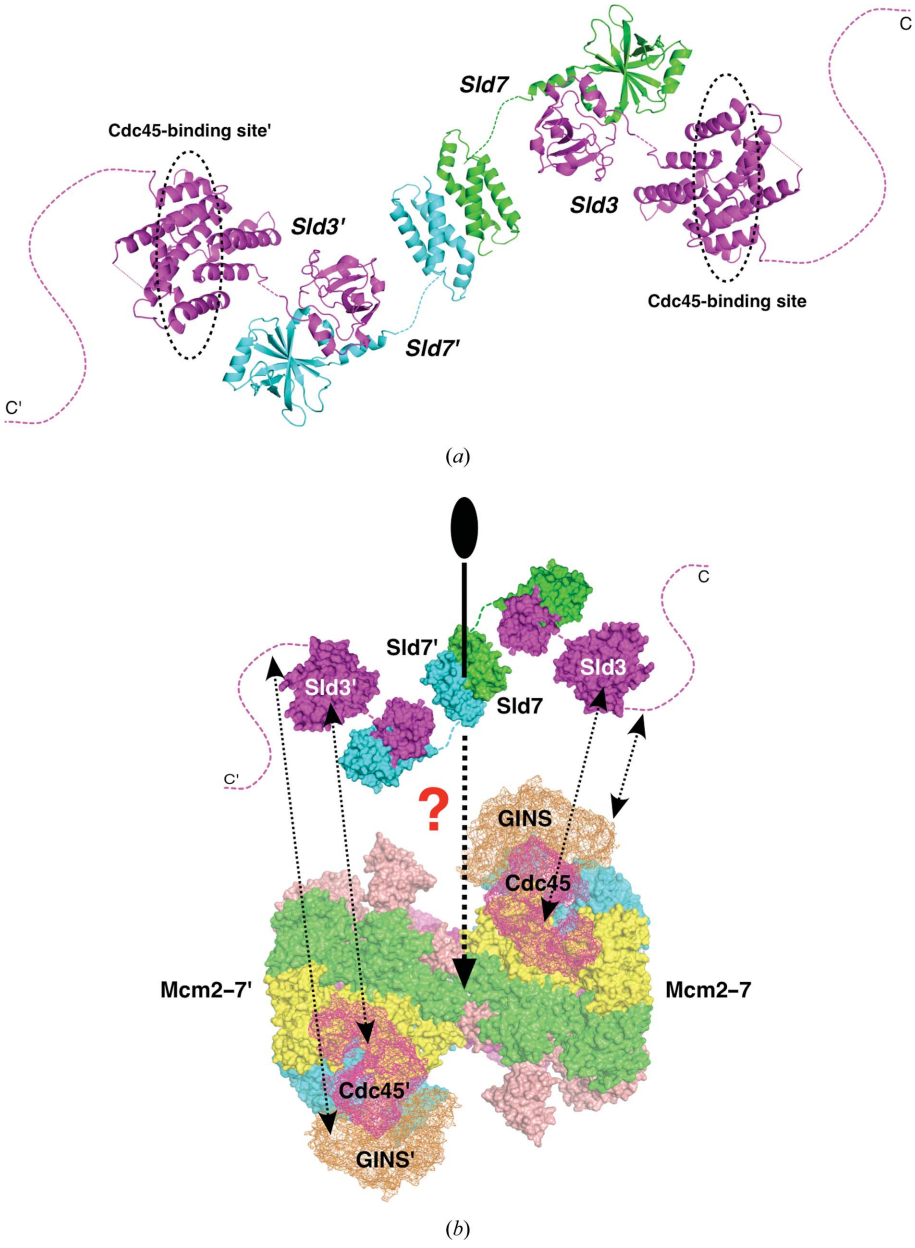


Figure 3
The suggested model of the quaternary structure of the Sld3–Sld7 complex. (*a*) Ribbon representation of the Sld3–Sld7 complex. Green and cyan ribbons represent the two Sld7 molecules in the homodimer. The magenta ribbon represents the Sld3 molecule. Coordinates for the Cdc45-binding domain of Sld3 were obtained from the Protein Data Bank (PDB entry 3wi3; Itou *et al.*, 2014). Dashed lines show the linker regions connecting each structured domain and the C-terminal unstructured region of Sld3. The dashed circle indicates the Cdc45-binding site suggested by a previous study (Itou *et al.*, 2014). (*b*) Schematic representation of the Sld3–Sld7 complex and Mcm2–7 complex loaded on origins in a head-to-head orientation. The model of the Mcm2–7 complex bound by Cdc45 and GINS was obtained by fitting the atomic models of an archaeal Mcm hexamer, RecJ and the human GINS complex (PDB entries 4r7y, 2e9x and 1ir6, respectively; Miller *et al.*, 2014; Kamada *et al.*, 2007; Yamagata *et al.*, 2002) into the high-resolution cryo-electron microscopy map of the CMG complex, an active helicase molecule (EMDB code EMD-277; Costa *et al.*, 2014). Another set of fitted models was placed to generate a head-to-head double hexamer of the Mcm complex. Cdc45 (RecJ) and GINS are shown as magenta and orange mesh models, respectively, to show the spatial arrangement of the proteins on the Mcm complexes. Dashed lines show the linker regions connecting each structured domain of Sld3 and Sld7 and the C-terminal unstructured region of Sld3 [the same as in (*a*)]. The twofold symmetries expected to appear in the quaternary structure of the Sld3–Sld7 complexes and the Mcm double hexamers on the origins are indicated by a bold line. Protein interactions that are important to form the pre-initiation complex, as shown by previous studies (Kamimura *et al.*, 2001; Takayama *et al.*, 2003, Itou *et al.*, 2014), are indicated by dashed arrows.

complete DNA replication and faithful chromosome segregation within a limited time. If only one of the two helicases is activated, the region on the other side of the origin would not be duplicated, jeopardizing the maintenance of genome integrity. To avoid such a risk, it is reasonable to assume that the DNA replication initiation machinery has mechanisms to ensure bidirectional fork progression from the origins. Sld3 is the replication protein that mediates critical protein interactions to attach Cdc45 and GINS onto the Mcm2–7 complex at origins. To generate two replicative helicases properly, it is reasonable that two Sld3 molecules bind to a pair of Mcm2–7 complexes aligned at origins. However, a previous study reported that the copy numbers of the Sld3 and Sld7 molecules are less than half of those of origins (Tanaka, Nakato *et al.*, 2011). This stoichiometry in yeast cells implies that if Sld3 binds stochastically to Mcm2–7 as a monomer, its binding to each Mcm2–7 molecule is not ensured, even though the two Mcm2–7 molecules are on the same origin. Our study showed that two Sld3 molecules are connected through Sld7 molecules (Fig. 3*a*). The quaternary structure of the Sld3–Sld7 complex would be suitable to ensure the binding of two Sld3 molecules to both of the Mcm2–7 complexes loaded onto an origin in a head-to-head orientation (Fig. 3*b*). A previous study showed that Sld7 reduces the binding affinity of Sld3 to Cdc45 (Tanaka, Umemori *et al.*, 2011). The Sld7-binding domain is located on the N-terminal side of the Cdc45-binding domain of Sld3 (Figs. 3 and 1*b*). At present, although we cannot observe direct interaction between Sld7 and Cdc45, their steric configuration in the complex may weaken the interaction between Sld3 and Cdc45, which would allow Cdc45 to bind Mcm2–7 in association with origins. Recently, MTBP was identified as a treslin-binding protein in humans, and its depletion causes defects in cells similar to those caused by the depletion of Sld7 in yeast. Although MTBP and Sld7, and treslin and Sld3, respectively, differ in their molecular sizes and amino-acid sequences, MTBP binds to the N-

terminal side of the Sld3/treslin domain in treslin molecules, similarly to Sld7 and Sld3 (Boos *et al.*, 2013), implying the functional importance of their steric configuration in the complex. Although the replication protein homologous to Sld7 has not been identified in fission yeast to date, the evidence suggests the generality and importance of the roles played by Sld7 during the process of initiation of DNA replication.

Acknowledgements

We would like to thank for the beamline scientists at BL-17A of Photon Factory, Tsukuba, Japan for their valuable help during data collection.

References

Adams, P. D. *et al.* (2010). *Acta Cryst.* D**66**, 213–221.
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. & Adams, P. D. (2012). *Acta Cryst.* D**68**, 352–367.
Baker, N. A., Sept, D., Joseph, S., Holst, M. J. & McCammon, J. A. (2001). *Proc. Natl Acad. Sci. USA*, **98**, 10037–10041.
Boos, D., Yekezare, M. & Diffley, J. F. (2013). *Science*, **340**, 981–984.
Brady, M., Vlatkovic, N. & Boyd, M. T. (2005). *Mol. Cell. Biol.* **25**, 545–553.
Costa, A., Renault, L., Swuec, P., Petojevic, T., Pesavento, J. J., Ilves, I., MacLellan-Gibson, K., Fleck, R. A., Botchan, M. R. & Berger, J. M. (2014). *eLife*, **3**, e03273.
Emsley, P. & Cowtan, K. (2004). *Acta Cryst.* D**60**, 2126–2132.
Heller, R. C., Kang, S., Lam, W. M., Chen, S., Chan, C. S. & Bell, S. P. (2011). *Cell*, **146**, 80–91.
Itou, H., Muramatsu, S., Shirakihara, Y. & Araki, H. (2014). *Structure*, **22**, 1341–1347.
Joosten, R. P., te Beek, T. A., Krieger, E., Hekkelman, M. L., Hooft, R. W. W., Schneider, R., Sander, C. & Vriend, G. (2011). *Nucleic Acids Res.* **39**, D411–D419.
Kabsch, W. (2010). *Acta Cryst.* D**66**, 133–144.
Kamada, K., Kubota, Y., Arata, T., Shindo, Y. & Hanaoka, F. (2007). *Nature Struct. Mol. Biol.* **14**, 388–396.
Kamimura, Y., Tak, Y.-S., Sugino, A. & Araki, H. (2001). *EMBO J.* **20**, 2097–2107.
Kelley, L. A. & Sternberg, M. J. (2009). *Nature Protoc.* **4**, 363–371.
Krissinel, E. & Henrick, K. (2007). *J. Mol. Biol.* **372**, 774–797.
Kumagai, A., Shevchenko, A., Shevchenko, A. & Dunphy, W. G. (2010). *Cell*, **140**, 349–359.
Larkin, M. A., Blackshields, G., Brown, N. P., Chenna, R., McGettigan, P. A., McWilliam, H., Valentin, F., Wallace, I. M., Wilm, A., Lopez, R., Thompson, J. D., Gibson, T. J. & Higgins, D. G. (2007). *Bioinformatics*, **23**, 2947–2948.
Lovell, S. C., Davis, I. W., Arendall, W. B., de Bakker, P. I., Word, J. M., Prisant, M. G., Richardson, J. S. & Richardson, D. C. (2003). *Proteins*, **50**, 437–450.
Masai, H., Taniyama, C., Ogino, K., Matsui, E., Kakusho, N., Matsumoto, S., Kim, J.-M., Ishii, A., Tanaka, T., Kobayashi, T., Tamai, K., Ohtani, K. & Arai, K. (2006). *J. Biol. Chem.* **281**, 39249–39261.
Miller, J. M., Arachea, B. T., Epling, L. B. & Enemark, E. J. (2014). *eLife*, **3**, e03433.
Moyer, S. E., Lewis, P. W. & Botchan, M. R. (2006). *Proc. Natl Acad. Sci. USA*, **103**, 10236–10241.
Muramatsu, S., Hirai, K., Tak, Y.-S., Kamimura, Y. & Araki, H. (2010). *Genes Dev.* **24**, 602–612.
Ness, S. R., de Graaff, R. A., Abrahams, J. P. & Pannu, N. S. (2004). *Structure*, **12**, 1753–1761.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Sanchez-Pulido, L., Diffley, J. F. & Ponting, C. P. (2010). *Curr. Biol.* **20**, R509–R510.
Sansam, C. L., Cruz, N. M., Danielian, P. S., Amsterdam, A., Lau, M. L., Hopkins, N. & Lees, J. A. (2010). *Genes Dev.* **24**, 183–194.
Sheu, Y.-J. & Stillman, B. (2006). *Mol. Cell*, **24**, 101–113.
Sheu, Y.-J. & Stillman, B. (2010). *Nature (London)*, **463**, 113–117.
Takayama, Y., Kamimura, Y., Okawa, M., Muramatsu, S., Sugino, A. & Araki, H. (2003). *Genes Dev.* **17**, 1153–1165.
Tanaka, S., Nakato, R., Katou, Y., Shirahige, K. & Araki, H. (2011). *Curr. Biol.* **21**, 2055–2063.
Tanaka, S., Umemori, T., Hirai, K., Muramatsu, S., Kamimura, Y. & Araki, H. (2007). *Nature (London)*, **445**, 328–332.
Tanaka, T., Umemori, T., Endo, S., Muramatsu, S., Kanemaki, M., Kamimura, Y., Obuse, C. & Araki, H. (2011). *EMBO J.* **30**, 2019–2030.
Terwilliger, T. C., Adams, P. D., Read, R. J., McCoy, A. J., Moriarty, N. W., Grosse-Kunstleve, R. W., Afonine, P. V., Zwart, P. H. & Hung, L.-W. (2009). *Acta Cryst.* D**65**, 582–601.
Winn, M. D. *et al.* (2011). *Acta Cryst.* D**67**, 235–242.
Yabuuchi, H., Yamada, Y., Uchida, T., Sunathvanichkul, T., Nakagawa, T. & Masukata, H. (2006). *EMBO J.* **25**, 4663–4674.
Yamagata, A., Kakuta, Y., Masui, R. & Fukuyama, K. (2002). *Proc. Natl Acad. Sci. USA*, **99**, 5908–5912.
Yao, M., Zhou, Y. & Tanaka, I. (2006). *Acta Cryst.* D**62**, 189–196.