The structure and polymerase-recognition mechanism of the crucial adaptor protein AND-1 in the human replisome

Chengcheng Guan†1,†‡, Jun Li†1,†‡, Dapeng Sun‡, Yingfang Liu‡1,2, and Huanhuan Liang‡3

From the †National Laboratory of Biomacromolecules, Institute of Biophysics, Chinese Academy of Sciences, 15 Datun Road, Chaoyang District, Beijing 100101 and the ‡College of Life Sciences, University of Chinese Academy of Sciences, Beijing 100049, China

Edited by Patrick Sung

DNA replication in eukaryotic cells is performed by a multienzymatic complex called the replisome, which consists of helicases, polymerases, and adaptor molecules. Human acidic nucleoplasmic DNA-binding protein 1 (AND-1), also known as WD repeat and high mobility group (HMG)-box DNA-binding protein 1 (WDHD1), is an adaptor molecule crucial for DNA replication. Although structural information for the AND-1 yeast ortholog is available, the mechanistic details for how human AND-1 protein anchors the lagging-strand DNA polymerase α (pol α) to the DNA helicase complex (Cdc45-MCM2–7-GINS, CMG) await elucidation. Here, we report the structures of the N-terminal WD40 and SepB domains of human AND-1, as well as a biochemical analysis of the C-terminal HMG domain. We show that AND-1 exists as a homotrimer mediated by the SepB domain. Mutant study results suggested that a positively charged groove within the SepB domain provides binding sites for pol α. Different from its ortholog protein in budding yeast, human AND-1 is recruited to the CMG complex, mediated by unknown participants other than Go Ichi Ni San. In addition, we show that AND-1 binds to DNA in vitro, using its C-terminal HMG domain. In conclusion, our findings provide important insights into the mechanistic details of human AND-1 function, advancing our understanding of replisome formation during eukaryotic replication.

In eukaryotic cells, chromosomal DNA is replicated by virtue of a large multiprotein complex known as the replisome (1, 2). The main enzymatic units of the replisome include a helicase complex, mini-chromosome maintenance 2–7 (MCM2–7), and three DNA polymerases, polymerase α, ε, and δ. The double-strand DNA template is unwound by the hetero-hexameric MCM2–7 assisted by the cell division cycle protein 45 (Cdc45) and Go Ichi Ni San complex (GINS) (3, 4), together forming the Cdc45-MCM2–7-GINS (CMG) complex. Subsequently, new DNA strands are synthesized with bases complementary to the separate leading and lagging strands. The DNA replication process based on the leading strand is performed continuously by polymerase ε (pol ε), which directly binds to the CMG complex (1). In contrast, a multiple-step replication process based on the lagging strand is performed requiring primase, polymerase α (pol α), and polymerase δ (pol δ), producing discontinuous Okazaki fragments. During this process, a short (10 nucleotides, nt) RNA primer is first synthesized by primase (5). The primer is further extended to about 20 nt of DNA by pol α (6) and finally elongated by pol δ (7). pol α contains an accessory B subunit and a catalytic subunit p180, termed POL1 in yeast and POLA in human. In contrast to pol ε, neither pol α nor pol δ directly attaches to the CMG complex. Instead, specific adaptor molecules are required to ensure the integrity of the replisome complex and the efficiency of the replication process.

Specific replication factors have been identified as adaptor molecules between CMG complex and pol α, including a class of orthologous genes termed adaptor chromosome transmission fidelity 4 (Ctf4) in budding yeast (8–10) and acidic nucleoplasmic DNA-binding protein 1 (AND-1, also known as WD repeat and HMG-box DNA-binding protein 1, WDHD1) in humans (8, 9, 11, 12). Ctf4 is composed of an N-terminal WD40 domain and a C-terminal SepB domain. It was recently shown that Ctf4 interacts with both POL1 and Sld5, the Go subunit of GINS, in its trimeric form. The SepB domain mediates the Ctf4 trimer formation and provides a binding groove for POL1 and Sld5, both of which possess a Ctf4-binding motif (10, 13). In contrast to Ctf4, AND-1 possesses a high mobility group (HMG) DNA-binding domain in its C terminus (14, 15). The N-terminal WD40 domain, the following SepB domain, and the HMG domain are connected by two flexible loops with 120 and 170 amino acids. Because the sequences of the SepB domains are highly conserved, the CMG-pol α coupling mechanism of Ctf4 is thought to be universal. Nevertheless, given the fact that human cells possess a larger genome with more complex regulatory mechanisms, the precise functions of AND-1 remain to be elucidated.
To investigate the structure-function relationship of AND-1 during CMG-pol α coupling, we studied the structures of its WD40 and SepB domains by X-ray crystallography and electron microscopy (EM). In addition, we used a biochemical approach to study the interaction between AND-1 and its putative partners, including POLA, Sld5, GINS, and various DNA substrates. We found that AND-1 recruits POLA but not Sld5 or GINS, suggesting a different CMG-pol α coupling mechanism. In addition, the HMG domain of AND-1, which is not present in Ctf4, binds long and damaged DNA substrates suggesting possible roles of AND-1 in DNA repair.

Results

Crystal structures of WD40 and SepB domains in AND-1

AND-1 contains three domains linked by two loops (Fig. 1A). The inter-domain loops of AND-1 are highly flexible, rendering crystallization challenging. We therefore individually purified, crystallized, and determined the structures of the WD40 domain (residues 1–330) and the SepB domain (residues 420–850) (Fig. 1). We determined the structure of the WD40 domain at 1.85 Å resolution using single isomorphous replacement with anomalous scattering (SIRAS) method (Table 1). In the structure, seven blades (Blades i–vii) formed a ring-like propeller, similar to those in the canonical WD40 domain structures (16). As shown in Fig. 1B, blades v and vii contain three β-strands; all other blades in the propeller contain four anti-parallel β-strands. The extended N terminus inserts into blade vii and interacts in an anti-parallel manner with the last C-terminal β-strand, stabilizing the ring-like structure.

The structure of the SepB domain was determined at 2.75 Å resolution using a Se-Met-based single wavelength anomalous dispersion method (Table 1). As shown in Fig. 1C, the structure of the SepB domain covers residues 420–850 of AND-1. The crystal adopts an F432 space group with one monomer in each asymmetry unit. Our analysis showed that SepB domain contains a WD40 subdomain and a helical subdomain, which is connected by a 10-residue loop. The WD40 structure includes six tandem blades (blades I–VI), forming a ring-shaped structure with a gap between blades I and VI. Similar to those in the reported WD40 domain structures, the blades consist of three or four β-strands. The helical subdomain locates at the lateral bottom of the WD40 subdomain, under blades V and VI. It is composed of five α-helixes, α1–5, arranged as a helical bundle, which extend out in an arm-like manner.

To explore the functional role of the SepB domain of AND-1, we compared our structure of SepB domain with that of Ctf4 (PDB code 4C8S) using PDBeFold (17). Our comparative analysis showed that the overall root mean square deviation of the two structures is 2.25 Å. As shown in Fig. 1D and E, when superposing the two structures, we found that the major differences originate from the flexible loops. Their secondary structural elements exhibit an identity of 88%. The high structural similarity indicates the functions of AND-1 are comparable with those of Ctf4.

Assembly of AND-1 homotrimer is mediated by SepB domain

Trimerization of Ctf4 is currently thought to be essential for cellular functioning (10). Similarly, AND-1 has previously been reported to assemble into an oligomeric structure (8, 9). To identify the oligomeric state of AND-1, we performed size-exclusion chromatography. As shown in Fig. 2A, full-length AND-1(1–1129), WD40 with SepB domains (WD40 + SepB(1–850)), and SepB domain(420–850) are predominantly present in their trimeric form in vitro. In contrast, both the
WD40 domain and the HMG region are exclusively present as monomers. These results suggest the AND-1 trimer is mediated by its SepB domain.

In our crystal structures, three SepB domain molecules in the cubic cell interact closely with each other, sharing a 3-fold symmetric axis of the $F_{432}$ space group. As shown in Fig. 2B, the WD40 subdomains of these three SepB molecules stack side-by-side, generating a large clover-like platform. The helical subdomains locate on the bottom of the platform, constructing an equilateral triangle.

Table 1
Data collection and refinement statistics

|                        | Se-Met-SepB | Native WD40 | Pt-derivative WD40 |
|------------------------|------------|-------------|---------------------|
| **Data collection**    |            |             |                     |
| Beamline               | SSRF BL17U | SSRF BL17U  | SSRF BL17U          |
| Wavelength (Å)         | 0.9792     | 0.9177      | 0.9177              |
| Space group            | $F_{432}$  | $P_{4,2,2}$ | $P_{4,2,2}$         |
| **Cell parameters**    |            |             |                     |
| $a$, $b$, $c$ (Å)      | 249.65, 249.65, 249.65 | 79.90, 79.90, 163.74 | 80.20, 80.20, 165.15 |
| $\alpha$, $\beta$, $\gamma$ (°) | 90.00, 90.00, 90.00 | 90.00, 90.00, 90.00 | 90.00, 90.00, 90.00 |
| Resolution (Å)         | 50.00–2.75 (2.85–2.75)* | 50.00–1.85 (1.92–1.85) | 50.00–2.90 (2.95–2.90) |
| $R_{crys}$             | 0.104 (0.693) | 0.061 (0.644) | 0.097 (0.720)       |
| $I/\sigma(I)$          | 28.25 (5.31) | 30.48 (2.51) | 22.64 (4.97)        |
| Completeness (%)       | 99.9 (100.0) | 99.5 (94.6) | 99.8 (100.0)        |
| Redundancy             | 19.7 (18.7) | 7.7 (5.9)   | 15.3 (16.1)         |
| **Refinement**         |            |             |                     |
| Resolution (Å)         | 50.00–2.75 | 50.00–1.85  |                     |
| Total/free reflections | 32,164/3182 | 45,822/2000 |                     |
| $R_{work}$             | 0.209 (0.316) | 0.174 (0.247) |                     |
| $R_{free}$             | 0.254 (0.419) | 0.192 (0.305) |                     |
| $N_O$ atoms            | 3346       | 2881        |                     |
| Protein                | 3256       | 2501        |                     |
| Ligand/ion             | 0          | 0           |                     |
| Water                  | 90         | 380         |                     |
| $B$-factor             | 40.90      | 34.94       |                     |
| Protein                | 41.09      | 32.34       |                     |
| Ligand/ion             | 0          | 0           |                     |
| Water                  | 33.97      | 52.03       |                     |
| Root mean square deviations |        |             |                     |
| Bond length (Å)        | 0.004      | 0.008       |                     |
| Bond angles (°)        | 0.71       | 0.93        |                     |
| Ramachandran plot      |            |             |                     |
| Favored (%)            | 97.3       | 96.6        |                     |
| Allowed (%)            | 2.7        | 3.4         |                     |
| Outlier (%)            | 0.00       | 0.00        |                     |

Values in parentheses are for the highest-resolution shell.

Figure 2. AND-1 is a homotrimer mediated by SepB domain. A, FL AND-1 (amino acids 1–1129, black), WD40 + SepB domain (amino acids 1–850, green), SepB domain (amino acids 420–850, blue), WD40 repeats domain (amino acids 1–330, magenta), and HMG-box region (amino acids 1017–1129, brown) show distinct elution volumes in size-exclusion chromatography (HiLoad 16/600 Superdex 200 pg, GE Healthcare). The standard molecular mass marker is labeled on the top. B, double-layer triangular structures of trimeric SepB. The 3-fold symmetric axis is perpendicular (left) to or parallel (right) with paper. C, negative staining EM images of FL AND-1. Three typical images after reference-free class averages are displayed on the left, and the corresponding models are shown on right.
involved in the inter-molecular interaction. These residues mainly originate from the WD40 subdomain and participate in the interaction between the H9252-propellers. However, a small number of these residues contribute to the hydrogen bonds between helical subdomains and adjacent H9252-propellers. Together, our structure provides the structural basis for AND-1 trimerization.

Furthermore, we verified the presence of the trimeric structure of AND-1 using the negative staining EM method (Fig. 2C). We collected 7852 particles in the images and classified the particles into 16 classes (supplemental Fig. S1 and supplemental Table S1). In our EM images, the WD40 domain and the SepB domain are clearly visible. However, the 10-kDa HMG domain is too small to be identified. In the center of the EM images, a triangle map fits well with the trimeric crystal structure of the SepB domain. Thus we assigned the stable core in the images with trimeric SepB domains. Subsequently, we assigned the maps surrounding the SepB core to the WD40 domains. The particle numbers of the 16 classes are comparable, suggesting that the WD40 domains move around SepB domains without specific orientations. Together, our structural analysis clearly shows that the basis for AND-1 trimerization is the strong interaction between the SepB domains. This result is similar to that reported for Ctf4 in a previous study (10).

**Helical subdomain in SepB domain binds pol α**

In the replisome of budding yeast, POL1 binds to the SepB domain of Ctf4 using a 13-amino acid helix called the Ctf4-binding motif. An earlier structure of the Ctf4-POL1 complex suggested that the Ctf4-binding motif of POL1 binds to the helical region of the Ctf4 SepB domain (10). Because both the Ctf4-binding motif and SepB domain are conserved from yeast to human (Fig. 3, A and B), we hypothesized that AND-1 interacts with pol α using a similar manner. To test this, we employed an in vitro pulldown assay. Using sequence alignment, we identified a putative AND-1-binding motif (151–171 residues) of POLA. The AND-1-binding motif was fused with a glutathione S-transferase (GST) protein at the N terminus. The purified GST-tagged POLA was used as bait to pull down SepB domain. As shown in Fig. 3E, the SepB domain specifically interacted with the GST-tagged AND-1-binding motif. To identify the binding surface for POLA in the SepB domain of AND-1, we compared the structures of the SepB domain in AND-1 and Ctf4 (Fig. 3C). In the structure of Ctf4, a semi-enclosed groove between helices H2 and H4 provides binding sites for POL1. Although residues in the helical subdomain of Ctf4 and AND-1 are not completely conserved, the overall root mean square deviation of these two subdomains is only 1.70 Å (Figs. 1E and 3, A and C), suggesting this region is structurally conserved from yeast to human. Therefore, we propose that the anti-parallel helices α2 and α4 of AND-1, corresponding to the helices H2 and H4 in Ctf4, provide the binding surface for the AND-1-binding motif of POLA. To test this assumption, we designed five mutants in α2 and α4, including Ala-770, Cys-773, and Ala-792, which are conserved in Ctf4 and AND-1, as well as Met-766 and Tyr-799, which are unique in AND-1. As shown in Fig. 3D, all these residues locate in the groove, with
side chains fully or partially exposed on the protein surface. To change the size of the groove and decrease its hydrophobicity, we replaced both alanine residues with positively charged lysines (A770K and A792K) and the hydrophobic residues with small alanines (M766A, C773A, and Y799A). As shown in Fig. 3E, the binding abilities of M766A, A770K, A792K, and Y799A were almost totally abolished, indicating that the residues in the groove are crucial for POLA recognition. In contrast, C773A still interacted with POLA, probably because it is located within the part of the groove distal to the putative POLA motif (Fig. 3D). Together, these results indicate that the interaction between AND-1 and polα is conserved from yeast to human.

Lack of in vitro interaction between AND-1 and GINS complex

In budding yeast, the interaction between the MCM2–7 helicase and Ctf4 is mediated through the N-terminal helix (1–19 amino acids) of Sld5, one of the GINS subunits (10). This helix is structurally analogous to the Ctf4-binding motif in POL1, both of which share the binding grooves in Ctf4. However, this motif is found only in fungi (including *Saccharomyces cerevisiae*) (10) but is lacking in *Schizosaccharomyces pombe* and multicellular organisms, including human (Fig. 4A). To test the interaction between AND-1 and GINS, an in vitro pulldown assay was performed using purified GST-tagged Sld5 protein or maltose-binding protein (MBP)-tagged GINS complex. Using GST-Sld5 as bait, we were unable to pull down more AND-1 derivatives (WD40, SepB, or SepB + HMG domains) than with GST alone (supplemental Fig. S2). We further tested the interaction between GINS complex and AND-1 derivatives. As shown in Fig. 4B, the GINS complex was pulled down by FLAG-Cdc45 but not by AND-1 derivatives. These results suggest that the interaction between AND-1 and Sld5 is much weaker than that of Ctf4-Sld5 in budding yeast (see Fig. 3D in Ref. 10). This finding is consistent with a previous report showing that the assembly of the human AND-1-GINS complex isolated from insect cells is unstable (8). Therefore, we assume that other as yet unknown proteins mediate the interaction between CMG and AND-1 in human cells, thus rendering the CMG-polα coupling mechanism of AND-1 within the replisome different from that of Ctf4.

**AND-1 binds DNA through HMG domain**

Compared with Ctf4, AND-1 contains a unique HMG domain at its C terminus, a type of DNA-binding module that recognizes a variety of DNA conformations (15, 19). To test the DNA-binding specificity and affinity of the HMG domain, we designed six types of model DNAs, including long-ssDNA (59 nt), long-dsDNA (59 bp), short-ssDNA (24 nt), short-dsDNA (24 bp), 3’-recessed dsDNA, and Holiday junction (HJ) structure (Table 2). Using an electrophoretic mobility shift assay (EMSA), we found that the HMG domain bound to 3’-recessed dsDNA, long-dsDNA, HJ structure, as well as long-ssDNA fragments. In contrast, the interactions between HMG domain and short DNAs, including single-strand and double-strand DNA fragments were negligible (Fig. 5A). The DNA-binding ability of AND-1 protein was further tested and verified, as shown in Fig. 5B. HMG domain and full-length (FL) AND-1 showed comparable binding ability; in contrast, the WD40 + SepB domain barely interacted with any tested DNAs. The dissociation constant $K_d$ of AND-1 and DNAs was further measured (Fig. 5, C and D, and supplemental Fig. S3). The $K_d$ values of the HMG domain and four types of DNAs, including long-ssDNA, long-
dsDNA, 3’-recessed dsDNA, and HJ, were determined as 1.5 ± 0.8 × 10⁻⁶ M, 1.5 ± 0.5 × 10⁻⁶ M, 1.8 ± 0.3 × 10⁻⁶ M, and 1.6 ± 0.9 × 10⁻⁶ M, respectively; the $K_D$ values of full-length AND-1 protein and these DNAs were 4.5 ± 1.4 × 10⁻⁷ M, 1.0 ± 0.2 × 10⁻⁷ M, 4.5 ± 0.6 × 10⁻⁷ M, and 4.9 ± 0.9 × 10⁻⁷ M, respectively (Fig. 5E). Together, these data suggest that AND-1 is a DNA-binding protein by virtue of possessing an HMG domain.

**Discussion**

In this report, we reveal the crystal structures of the WD40 N-terminal domain and SepB domain from the human AND-1 protein. To elucidate the role of AND-1 during replication, we analyzed the interactions between AND-1 and its potential protein partners in the replisome. Our results reveal the direct interaction between AND-1 and pol α. However, the interaction between AND-1 and CMG complex remains unclear. Therefore, elucidation of the precise mechanism of AND-1 function during DNA replication awaits further study.

The ortholog of AND-1, Ctf4, is a multifunctional adaptor molecule involved in a variety of cellular processes (20–27). Recent reports suggested that the functions of Ctf4 are mainly dependent on its SepB domain and Ctf4-binding motifs, also...
known as Ctf4-interacting peptide (CIP) boxes in Ctf4 partner molecules (28, 29). The cellular functions of AND-1 are similar to those of Ctf4, including DNA replication, sister chromatid cohesion establishment (30, 31), centromere formation (32, 33), epigenetic regulation of chromosome (33–35), checkpoint activation, and DNA repair (31, 36, 37). Here, we show that AND-1 possesses a SepB domain similar to that in Ctf4, suggesting a common structural basis for these biological functions. Thus, we propose that AND-1 also functions as a hub in human cells by linking proteins possessing multiple AND-1 interacting peptides to fulfill the cellular functions.

As a DNA-binding domain, the HMG domain may be critical of the cellular functions of AND-1. The structure of the HMG domain (PDB code 2D7L) is similar to that found in HMGB1 (supplemental Fig. S4), a canonical HMGB subfamily member that binds to bent or distorted DNA (19). Because the abnormal DNA substrates (long-ss, 3′-recessed, and HJ) recognized by AND-1 are frequently encountered in DNA lesions, we hypothesize that the HMG domain may bind to damaged DNA and could affect DNA repair pathways (15, 37, 39–41). Taken together, our results provide a strong structural basis for future studies on AND-1 in replisome formation and DNA repair process.

**Experimental procedures**

**Genes and plasmids**

The cDNA of human AND-1, Cdc45, and GINS (Sld5, Psf1, Psf2, and Psf3) was obtained by reverse transcription of mRNA from HEK293 cells. The DNA fragment encoding pol α (residues 151–171) was synthesized by Sangon. The MBP gene was amplified from pMAL-C2 vector (New England Biolabs).

To construct the expression plasmids of AND-1 derivatives, the DNA fragments of full-length (1–1129) or truncated AND-1 (1–850 for WD40 + SepB; 420–1129 for SepB + HMG; 1–330 for WD40; 420–850 for SepB; and 1017–1129 for HMG) with the H.i5 tag at the C terminal were digested by BamHI and Xhol and ligated into pFastBac1 vector (Invitrogen). The plasmids expressing the GINS complex, including pFastBac1-Sld5, pFastBac1-MBP-Psf1, pFastBac1-Psf2, and pFastBac1-Psf3, were constructed similarly, as mentioned above. The expression plasmids of pGEX-6P-1-pol α and pGEX-6P-1-Sld5 were constructed by inserting pol α (residues 151–171) or full-length Sld5 into pGEX-6P-1 vector (Invitrogen), using Xhol at the 3′-end and BamHI at the 5′-end as digestion sites. The plasmid expressing Cdc45 (pcDNA3.1-FLAG-Cdc45) was constructed by inserting full-length Cdc45 fused with an N-terminal FLAG tag into pcDNA3.1 vector (Invitrogen). All the recombinant plasmids were verified by sequencing (BGI).

**Antibodies**

The antibody for FLAG tag (anti-FLAG) was purchased from TransGen Biotech (HT201). The antibody for His tag (anti-His) was purchased from TIANGEN (AB102). The polyclonal antibody for MBP protein (anti-MBP) was generated using recombinant MBP protein according to a previous report (42).

**Protein expression and purification**

**Full-length and truncated AND-1—**The AND-1 proteins fused with a His6 tag at the C terminus were overexpressed in *Trichoplusia ni* (BT1-Tn5B1–4, Hi5) insect cell line using a Bac-to-Bac expression system (Invitrogen). The recombinant baculoviruses were prepared according to the manufacturer’s instructions. The Hi5 cells were suspended and cultured in ESF921 medium (Expression System) at 27 °C. To express the recombinant protein, the Hi5 cells (about 2.0 × 10⁶ cells/ml) were infected with 1% (v/v) recombinant baculovirus for 48 h. Cells were harvested by centrifugation at 4000 × g. The cell pellet was freshly frozen by liquid nitrogen and stored at −80 °C. For purification, the cell pellets were resuspended with lysis buffer (50 mM HEPES, pH 7.5, 200 mM NaCl, 5% glycerol, and 35 mM imidazole) and lysed by a cell homogenizer. The cell lysate was clarified by centrifugation at 35,000 × g. The supernatant was loaded onto a nickel-NTA (Novagen) column pre-equilibrated with the lysis buffer. After extensive washing with lysis buffer, the target protein was released with elution buffer (30 mM HEPES, pH 7.5, 5% glycerol, and 300 mM imidazole). The eluted sample was further purified using a Mono Q column (GE Healthcare) and a size-exclusion column (HiLoad 16/600 Superdex 200 pg, GE Healthcare) using 25 mM HEPES, pH 7.5, 200 mM NaCl, and 5% glycerol.

The selenomethionine (Se-Met) derivative of the SepB domain(420–850) was expressed in Hi5 insect cell line grown in methionine-free ESF921 medium (Expression Systems) with an additional 0.1 g/liter Se-Met and purified according to the protocol mentioned above.

**GST-pol α and GST-Sld5—**The pol α (residues 151–171) or full-length Sld5 fused with glutathione S-transferase (GST) at the N terminus was expressed in the *Escherichia coli*. The expression plasmid, pGEX-6P-1-Polα or pGEX-6P-1-Sld5, was transformed into *E. coli* strain BL21 (DE3) using a heat shock method. Cells were cultured in Luria-Bertani (LB) medium with 100 mg/liter ampicillin at 37 °C until the A₆₀₀ of the culture reached 0.8–1.0. The expression of the target protein was induced by 0.5 mM isopropyl thio-β-galactosidase (Sigma) for 20 h at 16 °C. Cells were harvested by centrifugation at 4000 × g. The pellet was resuspended in lysis buffer (50 mM Tris-HCl, pH 7.0, 500 mM NaCl, 10% glycerol, and 1 mM DTT) and lysed by sonication. The lysate was clarified by centrifugation at 35,000 × g. The supernatant was mixed with glutathione-Sepharose beads (GE Healthcare) pre-equilibrated in the lysis buffer and incubated for 1 h at 4 °C. The contaminating proteins were removed by washing three times with lysis buffer, followed by washing three times with 1 ml of pulldown buffer (20 mM HEPES, pH 7.2, 150 mM NaCl, 5% glycerol, 0.1% Nonidet P-40, and 1% BSA).

**GINS complex—**The human GINS complex was expressed using a Bac-to-Bac expression system (Invitrogen). The baculoviruses encoding Sld5, MBP-Psf1, Psf2, and Psf3 were prepared using pFastBac1-Sld5, pFastBac1-MBP-Psf1, pFastBac1-Psf2, and pFastBac1-Psf3, respectively. Hi5 cells infected with the four viruses were cultured in ESF921 medium (Expression System) for 60 h at 27 °C. After centrifugation, the cell pellet was resuspended in lysis buffer (50 mM HEPES, pH 7.5, 200 mM
**Structure of human acidic nucleoplasmic DNA-binding protein 1**

NaCl, and 5% glycerol) and lysed using a cell homogenizer. The insoluble component of the cell lysate was removed by centrifugation at 35,000 × g. The supernatants were loaded into an MBP column (Novagen) pre-equilibrated with lysis buffer. After extensive washing with lysis buffer, the target protein was eluted with elution buffer (30 mM HEPES, pH 7.5, 200 mM NaCl, 30 mM maltose, and 5% glycerol).

**FLAG-Cdc45—HEK293 cells transfected with pcDNA3.1-FLAG-Cdc45 were harvested 24 h post-transfection. The cell pellet was incubated in lysis buffer (20 mM Tris, pH 7.5, 150 mM NaCl, 5 mM EDTA, 0.5% Nonidet P-40, 10% glycerol, protease inhibitor mixture (Roche Applied Science)) at 4 °C for 1 h and lysed by sonication. The insoluble component of the cell lysate was removed by centrifugation. The FLAG-Cdc45 protein in the supernatant was purified using anti-FLAG M2 magnetic beads (Sigma).**

**Crystallization and structure determination**

To set up trials for crystallization, the proteins of WD40 (100 mg/ml) or SepB (10 mg/ml) were mixed with precipitant (Hampton Research), according to a protein/precipitant ratio of 1:1, in hanging-drop vapor diffusion experiment at 16 °C. The crystals of WD40 domain were generated in 1.8 M ammonium citrate tribasic, pH 7.5. To prepare the heavy atom derivatives, the WD40 crystals were soaked in mother liquor with potassium tetracyanoplatinate(II) hydrate for 4 h. The crystals of the Se-Met derivative SepB domain were obtained in 2.5M sodium formate, pH 7.0. All the crystals were cryo-protected using 30% (v/v) glycerol in crystallization precipitant before flash-freezing in liquid nitrogen.

The diffraction data were collected at a wavelength of 0.9792 Å (SepB) and 0.9177 Å (WD40) on beamline BL17U at Shanghai Synchrotron Radiation Facility (SSRF), China (43). Data were processed using HKL2000 (44). The initial phases of WD40 were obtained by SIRAS method using data from native and platinum-derivatized crystals. The initial phases of Se-Met derivative SepB were obtained by single wavelength anomalous dispersion method. The heavy atoms in both WD40 and SepB were located with SHELX C/D (45–48). After phase improvement and solvent flattening using Phenix.Autosol, the initial models were built up with Phenix.Autobuild (49). The models were modified and refined using alternate rounds of COOT (50) and Phenix.Refini (51). The statistics for data collection and refinement are shown in Table 1. The structure figures were prepared by PyMOL and Chimera (38).

**Negative-stain EM specimen preparation**

The recombinant full-length AND-1 protein samples were diluted to a final concentration of ~80 nM in buffer containing 25 mM HEPES, pH 7.5, 200 mM NaCl, and 5% glycerol. Subsequently, a 3.5-μl sample was applied to glow-discharged EM copper grids covered by a thin layer of continuous carbon film (Life Trust). The specimen was stained with 2% (w/v) uranyl acetate for 1 min. After blotting off the residual stain solution, the stained EM grids were subsequently placed in a Tecnai 20 electron microscope (FEI) operated at 200-kV acceleration voltage for microscopy. The images were recorded at a nominal magnification of 50,000 using a 2048 × 2048 charge-coupled device camera (Gatan US1000 894), corresponding to a pixel size of 2.2 Å on the specimen. About 60 images were recorded manually with a dose of 35 e/Å² and the defocus ranging from ~1.5 to ~2.5 μm.

**Negative-stain EM image processing**

Semi-automatic particle selection was performed using e2boxer.py to obtain the particle coordinates (EMAN2) (52). 2D multireference alignment and classification based on the hierarchical clustering approach were accomplished by cl2d command in XMIPP (53). 7852 particles were picked out from 60 micrographs and classified into 16 classes with 10 iterations (supplemental Fig. S1). The number of particles changing classes decreases over iterations and reaches a stable level (about 4% of total particles) (supplemental Table S1, column 2). The number of particles changing classes increased dramatically (supplemental Table S1, column 3) when the particles were further classified into 32 classes, indicating the classification is converged after 10 iterations into 16 classes.

**Pulldown assay**

As mentioned above, the GST-pol α and GST-Sld5 proteins were expressed in E. coli and purified using glutathione-Sepharose beads (GE Healthcare). The MBP-GINS complex was expressed in insect cells and purified using MBP affinity column (Novagen). The FLAG-Cdc45 was purified using anti-FLAG M2 magnetic beads (Sigma).

For pulldown assays shown in Fig. 3E and supplemental Fig. S2, 10 μg of bait proteins (GST-pol α, GST-Sld5, and GST) was bound to glutathione-Sepharose beads (GE Healthcare). The beads were then incubated with 1 mg of purified AND-1 derivative protein (WD40, SepB, or SepB + HMG) for 1 h at 4 °C. Glutathione-Sepharose beads bound with GST protein were used as control. The excessive AND-1 proteins were removed by consecutive washing using pulldown buffer (20 mM HEPES, pH 7.2, 150 mM NaCl, 5% glycerol, 0.1% Nonidet P-40, and 1% BSA) and pulldown buffer without BSA. The proteins on the beads were detected using SDS-PAGE and Coomassie Blue staining.

For pulldown assay shown in Fig. 4B, the purified FLAG-Cdc45 or AND-1 derivative proteins (WD40 and SepB + HMG) were bound to anti-FLAG M2 magnetic beads (Sigma) or nickel-NTA beads (Novagen), respectively. Then the beads were incubated with purified MBP-GINS complex at 4 °C for 4 h and washed five times with lysis buffer (20 mM Tris, pH 7.5, 150 mM NaCl, 5 mM EDTA, 0.5% Nonidet P-40, 10% glycerol, protease inhibitor mixture (Roche Applied Science)). The proteins bound to the beads were separated by SDS-PAGE and immunoblotted with anti-His, anti-MBP, and anti-FLAG, respectively.

**EMSA**

All the oligonucleotide DNAs were purchased from Takara, Inc. The sequences of these nucleotides are listed in Table 2. The nucleotides were annealed by heating to 96 °C and slow cooling to 4 °C (1 °C/min). The HMG, WD40 + SepB, and FL AND-1 proteins were purified as described above. In the reaction mixtures (20 μl), 5 pmol of 5'-FAM-labeled DNAs were
incubated with varying amounts of HMG (5, 10, 25, 50, 100, and 200 pmol) in DNA binding buffer (50 mM Tris-HCl, pH 7.5, 0.5 mM EDTA, 1 mM DTT, 10% glycerol, and 50 mM NaCl) for 20 min at 37 °C. The HMG-DNA complexes were separated from free DNA by electrophoresis through a 12% native PAGE in 0.5 × TBE buffer at 10 V/cm at 4 °C. To examine the DNA-binding ability of WD40 + SepB or FL protein, 5 pmol of 5′-FAM-labeled DNAs and 200 pmol of proteins were used in the 20 μl of EMSA. The $K_D$ values of HMG and FL were detected using 5 pmol of 5′-FAM-labeled DNAs incubated with varying amounts of proteins (5, 10, 25, 50, 100, and 200 pmol) in DNA binding buffer. The HMG or FL bound DNA and free DNA were separated and quantified by Typhoon FLA7000 (GE Healthcare). The concentration of the proteins was quantified using Nanodrop 2000 (Thermo Fisher Scientific), and the intensity of 5′-FAM-labeled DNA was quantified by ImageJ (54). The data were analyzed and shown using the GraphPad Prism software. Using the concentrations of protein ([protein]) and DNA bound to protein ([bound]) as $x$ and $y$, respectively, the equilibrium dissociation constant $K_D$ was calculated by fitting the following formula: $y = B_{max} x/(K_D + x)$.

Accession numbers

The coordinates for WD40 domain (1–330 residues) and SepB domain (420–850 residues, Se-Met derivative) are available from the Protein Data Bank with accession codes 5GVA and 5GVB.

Author contributions—C. G. performed most of the experiments. J. L. determined the structure. D. S. performed the EM assay. J. L., H. L. and Y. L. designed the experiment, analyzed the results, and wrote the manuscript. All authors reviewed the results and approved the final version of the manuscript.

Acknowledgments—We thank the staff at the BL17U beamline of the Shanghai Synchrotron Radiation Facility (SSRF) for assistance during data collection. We thank Dr. Torsten Juelich for linguistic assistance.

Note added in proof—There was an error in the calculation of the $K_d$ values in the EMSAs shown in Fig. 5 and supplemental Fig. S3 in the version of this article that was published as a Paper in Press on April 5, 2017. These errors have now been corrected and do not affect the results or conclusions of this work.

References

1. Sun, J., Shi, Y., Georgescu, R. E., Yuan, Z., Chait, B. T., Li, H., and O’Donnell, M. E. (2013) The architecture of a eukaryotic replisome. Nat. Struct. Mol. Biol. 20, 976–982
2. O’Donnell, M., Langston, L., and Stillman, B. (2013) Principles and concepts of DNA replication in bacteria, archaea, and eukarya. Cold Spring Harb. Perspect. Biol. 5, a010108
3. 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 A. Nature 524, 186–191
4. Gambus, A., Jones, R. C., Sanchez-Diaz, A., Kanemaki, M., van Deursen, F., Edmondson, R. D., and Labib, K. (2006) GINS maintains association of Cdc45 with MCM in replisome progression complexes at eukaryotic DNA replication forks. Nat. Cell Biol. 8, 358–366
5. Kuchta, R. D., and Stengel, G. (2010) Mechanism and evolution of DNA primases. Biochim. Biophys. Acta 1804, 1180–1189
6. Perera, R. L., Torella, R., Klinge, S., Kilkenny, M. L., Maman, J. D., and Pellegrini, L. (2013) Mechanism for priming DNA synthesis by yeast DNA polymerase α. Elife 2, e00482
7. Smith, D., and Whitehouse, I. (2012) Intrinsic coupling of lagging-strand synthesis to chromatin assembly. Nature 483, 434–438
8. Kang, Y. H., Farina, A., Bermudez, V. P., Tappin, I., Du, F., Galal, W. C., and Hurwitz, J. (2013) Interaction between human Ctf4 and the Cdc45/Mcm2–7/CMG (CMG) replicative helicase. Proc. Natl. Acad. Sci. U.S.A. 110, 19760–19765
9. Bermudez, V. P., Farina, A., Tappin, I., and Hurwitz, J. (2010) Influence of the human cohesion establishment factor Ctf4/And-1 on DNA replication. J. Biol. Chem. 285, 9493–9505
10. Simon, A. C., Zhou, J. C., van Deursen, F., Ehrin, C., Ivanova, M. E., Kilkenny, M. L., Renault, L., Kjaer, S., Matak-Vinkovic, D., Labib, K., Costa, A., and Pellegrini, L. (2014) A Ctf4 trimer couples the CMG helicase to DNA polymerase α in the eukaryotic replisome. Nature 510, 293–297
11. Zhu, W., Ukomadu, C., Jha, S., Senga, T., Dhar, S. K., Wohlschleger, J. A., Nult, L. K., Kornbhhut, S., and Dutta, A. (2007) Mcm10 and And-1/CCtf4 recruit DNA polymerase α to chromatin for initiation of DNA replication. Genes Dev. 21, 2288–2299
12. Im, J. S., Ki, S. H., Farina, A., Jung, D. S., Hurwitz, J., and Lee, J. K. (2009) Assembly of the Cdc45/Mcm2–7–GINS complex in human cells requires the Ctf4/And-1, RecQ4, and Mcm10 proteins. Proc. Natl. Acad. Sci. U.S.A. 106, 15628–15632
13. MacNeil, S. A. (2010) Structure and function of the GINS complex, a key component of the eukaryotic replisome. Biochem. J. 425, 489–500
14. Köhler, A., Schmidt-Zachmann, M. S., and Franke, W. W. (1997) AND-1, a natural chimeric DNA-binding protein, combines an HMG-box with regulatory WD-repeats. J. Cell Sci. 110, 1051–1062
15. Reeves, R., and Adair, J. E. (2005) Role of high mobility group (HMG) chromatin proteins in DNA repair. DNA Repair 4, 926–938
16. Smith, T. F. (2008) Diversity of WD-repeat proteins. Subcell. Biochem. 48, 20–30
17. Krissinel, E., and Henrick, K. (2004) Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. Acta Crystallogr. D Biol. Crystallogr. 60, 2256–2268
18. Krissinel, E., and Henrick, K. (2007) Inference of macromolecular assemblies from crystalline state. J. Mol. Biol. 372, 774–797
19. Bustin, M. (2001) Revised nomenclature for high mobility group (HMG) chromosomal proteins. Trends Biochem. Sci. 26, 152–153
20. Gosnell, J. A., and Christensen, T. W. (2011) Drosophila Ctf4 is essential for efficient DNA replication and normal cell cycle progression. BMC Mol. Biol. 12, 13
21. Xu, H., Boone, C., and Brown, G. W. (2007) Genetic dissection of parallel sister-chromatid cohesion pathways. Genetics 176, 1417–1429
22. Ogawa, H., Ut, A., Lai, M. S., Enomoto, T., and Seki, M. (2007) Chl1 and Ct4 are required for damage-induced recombination. Biochem. Biophys. Res. Commun. 354, 222–226
23. Lengronne, A., McIntyre, J., Katou, Y., Kanoh, Y., Hopfner, K. P., Shiraige, K., and Uhlmann, F. (2006) Establishment of sister chromatid cohesion at the S. cerevisiae replication fork. Mol. Cell 23, 787–799
24. Warren, C. D., Eckley, D. M., Lee, M. S., Hanna, J. S., Hughes, A., Peyser, B., Jie, C., Irizarry, R., and Spencer, F. A. (2004) S-phase checkpoint genes safeguard high-fidelity sister chromatid cohesion. Mol. Biol. Cell 15, 1724–1735
25. Petruczynki, M., Chwalla, B., Siomos, M. F., Yokobayashi, S., Helmhart, W., Deutschbauer, A. M., Davis, R. W., Watanabe, Y., and Nasmyth, K. (2004) Sister-chromatid cohesion mediated by the alternative RF-CCCTf18/Dcc1/Ctf8, the helicase Chl1 and the polymerase-α-associated protein Ct4 is essential for chromatid disjunction during meiosis II. J. Cell Sci. 117, 3547–3559
26. Mayer, M. L., Pot, I., Chang, M., Xu, H., Aneliuunas, V., Kwok, T., Newritt, R., Aebersold, R., Boone, C., Brown, G. W., and Hieter, P. (2004) Identification of protein complexes required for efficient sister chromatid cohesion. Mol. Biol. Cell 15, 1736–1745
Structure of human acidic nucleoplasmic DNA-binding protein 1

27. Hanna, J. S., Kroll, E. S., Lundblad, V., and Spencer, F. A. (2001) Saccharomyces cerevisiae CTF18 and CTF4 are required for sister chromatid cohesion. Mol. Cell. Biol. 21, 3144–3158

28. Villa, F., Simon, A. C., Ortiz Bazan, M. A., Kilkenny, M. L., Wirthsensohn, D., Wightman, M., Matak-Vinkovic, D., Pellegrini, L., and Labib, K. (2016) Ctf4 Is a hub in the eukaryotic replisome that links multiple CIP-box proteins to the CMG helicase. Mol. Cell 63, 385–396

29. Samora, C. P., Saksouk, J., Goswami, P., Wade, B. O., Singleton, M. R., Bates, P. A., Lengronne, A., Costa, A., and Uhlmann, F. (2016) Ctf4 links DNA replication with sister chromatid cohesion establishment by recruiting the Chl1 helicase to the replisome. Mol. Cell 63, 371–384

30. Tanaka, H., Kubota, Y., Tsujimura, T., Kumanow, M., Masai, H., and Takisawa, H. (2009) Replisome progression complex links DNA replication to sister chromatid cohesion in Xenopus egg extracts. Genes Cells 14, 949–963

31. Yoshizawa-Sugata, N., and Masai, H. (2009) Roles of human AND-1 in chromosome transactions in S phase. J. Biol. Chem. 284, 20718–20728

32. Jaramillo-Lambert, A., Hao, J., Xiao, H., Li, Y., Han, Z., and Zhu, W. (2013) Acidic nucleoplasmic DNA-binding protein (And-1) controls chromosome congression by regulating the assembly of centromere protein A (CENP-A) at centromeres. J. Biol. Chem. 288, 1480–1488

33. Hsieh, C. L., Lin, C. L., Liu, H., Chang, Y. J., Shih, C. J., Zhong, C. Z., Lee, S. C., and Tan, B. C. (2011) WDHD1 modulates the post-transcriptional step of the centromeric silencing pathway. Nucleic Acids Res. 39, 4048–4062

34. Li, Y., Jaramillo-Lambert, A., Hao, J., Yang, Y., and Zhu, W. (2011) The stability of histone acetyltransferase general control non-derepressible (Gcn) 5 is regulated by Cullin4-RING E3 ubiquitin ligase. J. Biol. Chem. 286, 41334–41352

35. Li, Y., Jaramillo-Lambert, A. N., Yang, Y., Williams, R., Lee, N. H., and Zhu, W. (2012) And-1 is required for the stability of histone acetyltransferase Gcn5. Oncogene 31, 643–652

36. Chen, Z., Tran, M., Tang, M., Wang, W., Gong, Z., and Chen, J. (2016) Proteomic analysis reveals a novel mutator 5 (Mut5) partner involved in mismatch repair pathway. Mol. Cell. Proteomics 15, 1299–1308

37. Hao, J., de Renty, C., Li, Y., Xiao, H., Kemp, M. G., Han, Z., DePamphilis, M. L., and Zhu, W. (2015) And-1 coordinates with Claspin for efficient Chk1 activation in response to replication stress. EMBO J. 34, 2096–2110

38. 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

39. Cimprich, K. A., and Cortez, D. (2008) ATR: an essential regulator of genome integrity. Nat. Rev. Mol. Cell Biol. 9, 616–627

40. Luciani, M. G., Oehlmann, M., and Blow, J. I. (2004) Characterization of a novel ATR-dependent, Chk1-independent, intra-S-phase checkpoint that suppresses initiation of replication in Xenopus. J. Cell Sci. 117, 6019–6030

41. Petermann, E., and Caldecott, K. W. (2006) Evidence that the ATR/Chk1 pathway maintains normal replication fork progression during unperturbed S phase. Cell Cycle 5, 2203–2209

42. Leenaars, M., and Hendriksen, C. F. (2005) Critical steps in the production of polyclonal and monoclonal antibodies: evaluation and recommendations. ILAR J. 46, 269–279

43. 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., Pan, Q. Y., Li, L., Zhou, H., Cui, Y., Xu, Q., Earnest, T., and He, J. H. (2015) The macromolecular crystallography beamline of SSRF. Nucl. Sci. Tech. 26, 12–17

44. Otwinowski, Z., and Minor, W. (1997) Processing of X-ray diffraction data collected in oscillation mode. Methods Enzymol. 276, 307–326

45. Sheldrick, G. M. (2008) A short history of SHELX. Acta Crystallogr. A 64, 112–122

46. Schneider, T. R., and Sheldrick, G. M. (2002) Substructure solution with SHELXD. Acta Crystallogr. D Biol. Crystallogr. 58, 1772–1779

47. Sheldrick, G. M. (2002) Macromolecular phasing with SHELXE. Zeitschrift Fur Kristallographie 217, 644–650

48. Sheldrick, G. M. (2010) Experimental phasing with SHELXC/D/E: combining chain tracing with density modification. Acta Crystallogr. D Biol. Crystallogr. 66, 479–485

49. Adams, P. D., Afonine, P. V., Bunkóczi, G., Chen, V. B., Davis, I. W., Echols, N., Headd, J. J., Humphris, G. J., Grosse-Kunstleve, R. W., McCoy, A. J., Moriarty, N. W., Oeffner, R., Read, R. J., Richardson, D. C., et al. (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. Acta Crystallogr. D Biol. Crystallogr. 66, 213–221

50. Emsley, P., Lohkamp, B., Scott, W. G., and Cowtan, K. (2010) Features and development of Coot. Acta Crystallogr. D Biol. Crystallogr. 66, 486–501

51. Afonine, P. V., Grosse-Kunstleve, R. W., Emsley, 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

52. Tang, G., Peng, L., Baldwin, P. R., Mann, D. S., Jiang, W., Rees, I., and Ludtke, S. J. (2007) EMAN2: an extensible image processing suite for electron microscopy. J. Struct. Biol. 157, 38–46

53. Scheres, S. H., Núñez-Ramírez, R., Sorzano, C. O., Carazo, J. M., and Marabini, R. (2008) Image processing for electron microscopy single-particle analysis using XMIPP. Nat. Protoc. 3, 977–990

54. Schneider, C. A., Rasband, W. S., and Eliceiri, K. W. (2012) NIH Image to Image: 25 years of image analysis. Nat. Methods 9, 671–675