

## Crystal structure of the human GINS complex

Jung Min Choi,<sup>1,2,3</sup> Hye Seong Lim,<sup>1,2,3</sup>  
Jeong Joo Kim,<sup>1,2</sup> Ok-Kyu Song,<sup>2</sup>  
and Yunje Cho<sup>1,2,4</sup>

<sup>1</sup>National Creative Initiatives for Structural Biology, Pohang University of Science and Technology, Pohang, Kyung Book 790-784, South Korea; <sup>2</sup>Department of Life Science, Pohang University of Science and Technology, Pohang, Kyung Book 790-784, South Korea

**The GINS complex mediates the assembly of the MCM2–7 (minichromosome maintenance) complex with proteins in a replisome progression complex. The eukaryotic GINS complex is composed of Sld5, Psf1, Psf2, and Psf3, which must be assembled for cell proliferation. We determined the crystal structure of the human GINS complex: GINS forms an elliptical shape with a small central channel. The structures of Sld5 and Psf2 resemble those of Psf1 and Psf3, respectively. In addition, the N-terminal and C-terminal domains of Sld5/Psf1 are permuted in Psf2/Psf3, which suggests that the four proteins have evolved from a common ancestor. Using a structure-based mutational analysis, we identified the functionally critical surface regions of the GINS complex.**

Supplemental material is available at <http://www.genesdev.org>.

Received March 2, 2007; revised version accepted April 19, 2007.

Initiation of chromosomal DNA replication is a tightly regulated process in the eukaryotic cell cycle that involves a complex network of protein–protein interactions. Eukaryotic DNA replication starts with a licensing process in which the minichromosome maintenance (MCM) complex binds to the chromatin in an origin recognition complex (ORC)-, Cdc6-, and Cdt1-dependent manner during G1 phase (Bell and Dutta 2002). Once a prereplicative complex (pre-RC) is formed, the MCM2–7 helicase must be activated in order to unwind origins of replication. The activation mechanism of the MCM complex is elusive. However, the complex that is formed by MCM2–7, Cdc45, and GINS exhibits helicase activity, which suggests that Cdc45 and GINS play key roles in the activation of the MCM complex (Moyer et al. 2006).

The GINS complex is essential for both the initiation and elongation stages of eukaryotic DNA replication (Kanemaki et al. 2003; Kubota et al. 2003; Takayama et al. 2003; Gambus et al. 2006). Although the precise molecular function of GINS is unclear, accumulating evi-

dence suggests that GINS is a molecular linker that mediates the assembly of replication factors around the MCM helicase during DNA replication. In budding yeast, chromatin immunoprecipitation analysis identified the GINS complex as a part of the replisome that contains MCM2–7, Cdc45, Pols  $\alpha$  and  $\epsilon$ , and other replication regulation factors (Calzada et al. 2005). The vertebrate replisome at the paused fork also contained GINS, along with Pols  $\alpha$  and  $\epsilon$ , MCM2–7, Cdc45, and MCM10 (Pacek et al. 2006). Moreover, a proteomics study in budding yeast showed that the GINS complex allows the MCM helicase to stably associate with Cdc45 in replisome progression complexes (RPCs) during and after the initiation of replication (Gambus et al. 2006; Kanemaki and Labib 2006). GINS facilitates the association with chromatin of Cdc45 and Dpb11, a protein required for Pols  $\alpha$  and  $\epsilon$  to associate with the replisome (Takayama et al. 2003; Kanemaki and Labib 2006). In vitro, GINS stimulates the activity of Pols  $\alpha$  and  $\epsilon$  (Shikata et al. 2006; De Falco et al. 2007).

The eukaryotic GINS complex is composed of four proteins, Sld5, Psf1, Psf2, and Psf3, which are conserved and essential for cell growth (Kanemaki et al. 2003; Kubota et al. 2003; Takayama et al. 2003). Psf1 is required for cell proliferation in yeast and mice, and loss of Psf1 leads to early embryonic death (Takayama et al. 2003; Ueno et al. 2005). Psf2 regulates the chromosome segregation, presumably through its S-phase role in centromere replication (Huang et al. 2005). In addition, Psf2 is up-regulated in certain liver cancer cells (Obama et al. 2005). In yeast, the loss of any one protein of the GINS complex resulted in a significant impairment of cell cycle progression, suggesting the importance of the assembly of four proteins for the GINS complex (Kanemaki et al. 2003; Takayama et al. 2003). Despite its significance in replication and cell cycle progression, no high-resolution structure is available to explain the molecule architecture of the Psf1, Psf2, Psf3, and Sld5 complex.

Recently, electron microscopy (EM) studies revealed that *Xenopus* GINS has a ring-like structure, thus suggesting that it may function like a clamp proliferating cell nuclear antigen (PCNA) (Krishna et al. 1994; Kubota et al. 2003). While PCNA recruits Pol  $\delta$  onto a primer site and stimulates the processivity of Pol  $\delta$ , GINS interacts with Pol  $\epsilon$  and stimulates its activity (Shikata et al. 2006). To elucidate how the four proteins are assembled in the GINS complex and to understand how GINS plays a role as a molecular linker during DNA replication, we determined the structure of the GINS complex. We also identified the conserved binding sites for other replication partners through mutational studies.

### Results and Discussion

#### *Overall structure of GINS*

We crystallized the core of the human GINS complex, which is formed with Sld5 (residues 11–213), Psf1 (residues 1–151), full-length Psf2, and full-length Psf3 (see the Supplemental Material). The statistics from the crystallographic analysis are summarized in Supplemental Table S1. The crystals contain three GINS complexes in an asymmetric unit. The C $\alpha$  atoms of one complex can be superimposed on the C $\alpha$  atoms of another complex

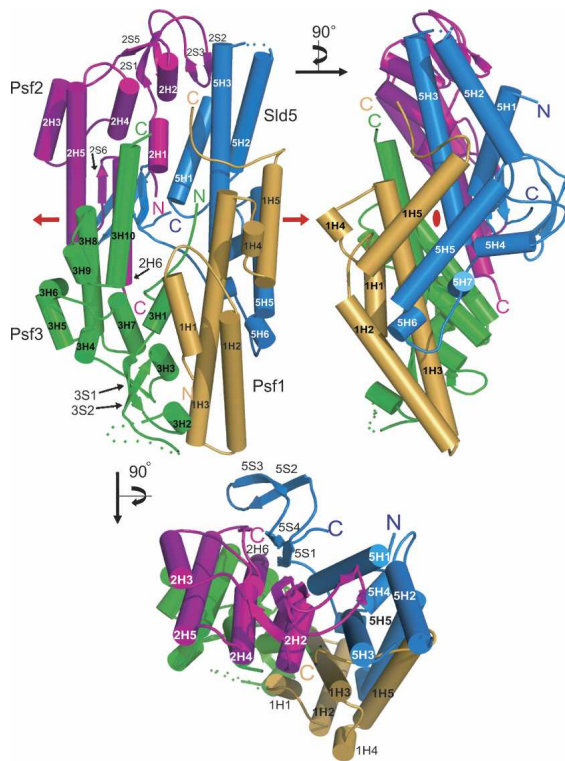
[*Keywords*: GINS; molecular assembly; protein evolution; replication forks]

<sup>3</sup>These authors contributed equally to this work.

<sup>4</sup>Corresponding author.

EMAIL [yunje@postech.ac.kr](mailto:yunje@postech.ac.kr); FAX 82-54-279-8111.

Article is online at <http://www.genesdev.org/cgi/doi/10.1101/gad.1548107>.



**Figure 1.** Overall structure of the human GINS complex. Three different views are shown. On *top left*, Sld5 (blue), Psf1 (orange), Psf3 (green), and Psf2 (magenta) are arranged in a clockwise direction. The pseudo-twofold symmetry axis is shown in red arrows or a red circle between the Psf2–Sld5 and Psf3–Psf1 complexes. In front of the secondary structure, labels (H or S) 1, 2, 3, and 5 are added for Psf1, Psf2, Psf3, and Sld5, respectively, in all figures.

with root mean square deviations (RMSDs) of 0.86–1.23 Å. The GINS complex forms an elliptical shape with overall dimensions of 55 Å × 106 Å × 61 Å (Fig. 1). At the center of the GINS complex is a channel with a diameter of ~5 Å. However, this channel is not large enough to accommodate single- or double-stranded DNA, and it is unlikely that the GINS complex interacts with DNA through this channel.

Figure 1 shows that the four components are packed in a clockwise order—Sld5, Psf1, Psf3, and Psf2—and interact with each other extensively. The most striking observation is that the structures of Sld5 and Psf2 resemble those of Psf1 and Psf3, respectively, and this structural similarity results in the presence of a pseudo-twofold symmetry within the whole molecular architecture of the GINS complex (Fig. 1).

#### Structures of Psf1, Psf2, Psf3, and Sld5

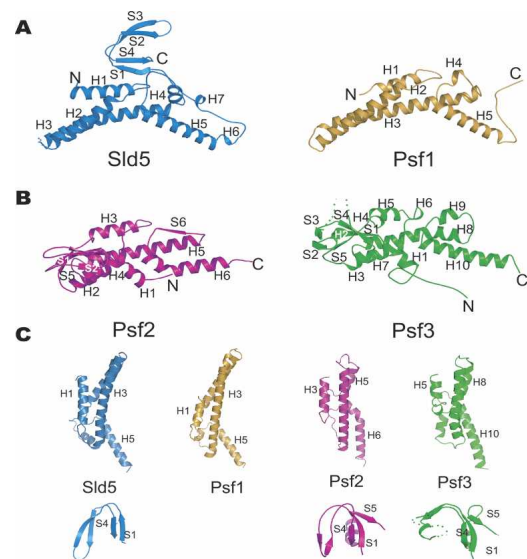
Despite their limited sequence identity (~15%), Psf1 and Sld5 can be superimposed with an RMSD of 2.4 Å for 116 C $\alpha$  atoms (Fig. 2A; Supplementary Fig. S1). Sld5 is formed with seven helices and four  $\beta$ -strands. The arch-shaped N-terminal domain is formed with a distinctive three-helical bundle (H1 to H3) and three additional helices (Fig. 2A,C). The C-terminal domain consists of an  $\alpha/\beta$  mixture with four  $\beta$ -strands, which form a jelly-roll structure. The C-terminal domain is on top of the arch, and strand S1 of Sld5 interacts with helix H1 through hydrophobic interaction.

Psf1 is composed of five helices. To obtain crystals, we removed the C-terminal portion of Psf1. However, considering the overall structural similarity between Sld5 and Psf1, it is possible that the truncated portion of Psf1 is formed with the  $\beta$ -strands. Psf1 also forms an arch shape with a three-helical bundle (H1, H2, and H3) and an additional helix, H5 (Fig. 2A,C).

Both Psf2 and Psf3 consist of an N-terminal domain containing an  $\alpha/\beta$  fold with a jelly-roll structure and a C-terminal domain with an entirely  $\alpha$ -helical structure (Fig. 2B). These two proteins can be superimposed with an RMSD of 2.7 Å for 149 aligned C $\alpha$  atoms. Psf2 is formed with six helices and six strands, whereas Psf3 is formed with 10 helices and five strands. Although the lengths of the helices are shorter than those of Sld5 and Psf1, helices H3, H4, H5, and H6 of Psf2 and helices H5, H7, H8, and H10 of Psf3 form arches equivalent to the four helices in Sld5 and Psf1 (Fig. 2C).

#### Domain exchange between Sld5/Psf1 and Psf2/Psf3

Despite the structural similarities between Sld5 and Psf1, and between Psf2 and Psf3, the overall structures of Psf2/Psf3 and Sld5/Psf1 are significantly different (Fig. 2A,B). Nevertheless, all of the proteins of GINS share four helices that form an arch and a jelly-roll fold. Sld5 and Psf1 possess four helices that form an arch at their N-terminal domains, whereas Psf2 and Psf3 have these helices in their C-terminal domains (Fig. 2C). All of the arch-forming helices from the four proteins can be superimposed with RMSDs of 2.1–4.5 Å. The jelly-roll motif of the C-terminal domain of Sld5 can also be superimposed with the N-terminal domains of Psf2 and Psf3, with RMSDs of 2.1–3.1 Å. Thus, our structure indicates that the N-terminal and C-terminal domains were permuted between Sld5/Psf1 and Psf2/Psf3 during the pro-



**Figure 2.** Structures of the four proteins of the GINS complex. (A) Structural similarities between Sld5 (blue) and Psf1 (orange). The N-terminal domain forms an arch shape, and the C-terminal domain is on top of the arch. The structure of Psf1 resembles that of Sld5. (B) Structural similarities between Psf2 (magenta) and Psf3 (green). (C) The four proteins of GINS share four arch-forming helices (*top*) and a jelly-roll fold (*bottom*). The N-terminal and C-terminal domains of Sld5/Psf1 are permuted in Psf2/Psf3.

cess of evolution, and the four proteins may have evolved from a common ancestor, which confirms the prediction proposed by recent bioinformatics studies (Makarova and Koonin 2003; Makarova et al. 2005).

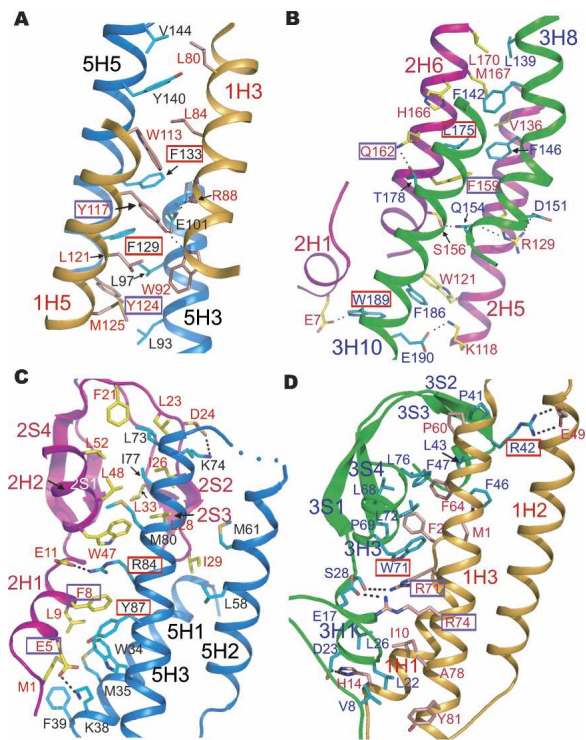
Crystal structure of the human GINS complex explains the two important features of archaeal GINS. First, in some archaea, including *Sulfolobus sulfataricus*, only two GINS components (GINS15 and GINS23) are found, and they form a tetramer (Makarova et al. 2005; Marinsek et al. 2006). GINS15 resembles eukaryotic Psf1 and Sld5, and GINS23 resembles eukaryotic Psf2 and Psf3. Second, most archaea encode a single GINS protein, and sequence comparison analysis shows that these proteins are similar to *S. sulfataricus* GINS15. Although we do not exclude the possibility that GINS23 cannot be detected by currently available methods, it is likely that many archaeal species may indeed have only one GINS component. It is possible that archaeal RPCs may be formed with relatively fewer components compared with eukaryotic RPCs. Since GINS mediates interactions of various replication factors in an RPC, archaeal GINS might be formed with a single protein, GINS15, in an environment where replication is controlled in a rather simple manner. However, as regulation of the initiation and elongation of DNA replication becomes more complicated in eukaryotes, archaeal GINS15 diverged to Sld5, Psf1, Psf2, and Psf3, which interact with various replication factors.

#### Intermolecular interaction between each component of GINS

Intermolecular interfaces are formed by Sld5–Psf1, Psf1–Psf3, Psf2–Psf3, and Sld5–Psf2 complexes (Figs. 1, 3, 4). Overall, similar regions of the four proteins participate in the intermolecular interactions. Helices H3 and H5 of Sld5 and equivalent helices of Psf1, Psf2, and Psf3 contribute to the formation of Sld5–Psf1 and Psf2–Psf3 interfaces. Helices H1 and H2 and strands S2 to S4 of Psf2/Psf3 and helices H1 and H3 of Sld5/Psf1 are involved in the formation of the Psf2–Sld5 and Psf1–Psf3 interfaces (Supplementary Fig. S3). Consequently, the Sld5–Psf1 and Sld5–Psf2 interfaces are very similar to those of Psf2–Psf3 and Psf1–Psf3, respectively (Fig. 3A–D).

**The Sld5–Psf1 and Psf2–Psf3 interfaces** Sld5 and Psf1 dimerize through the surface underneath the arch. The overall interaction in this interface shows a somewhat symmetrical pattern. The hydrophobic packing between Sld5 and Psf1 is extensive, with 24 residues from Sld5 interacting with 24 residues from Psf1. Helices H3 and H5 of Sld5 are packed in an anti-parallel manner against helices H3 and H5 of Psf1 to form an intermolecular four-helical bundle (Fig. 3A; Supplementary Fig. S3A). The ring groups of conserved Phe129 and Phe133 of Sld5 are stacked against the side chains of Trp113 and Tyr117 of Psf1 at the center of the four-helical bundle. Glu101 of Sld5 further stabilizes this interface by forming an ion pair and a H-bond with Arg88 and Tyr117 of Psf1, respectively. Detailed interactions are described in Supplementary Figure S3.

Psf2 and Psf3 interact through each of the two helices at the C-terminal region that are equivalent to helices H3 and H5 of Sld5/Psf1, which form an intermolecular four-helical bundle (Fig. 3B; Supplementary Fig. S3B). The primary interaction involves van der Waals contacts

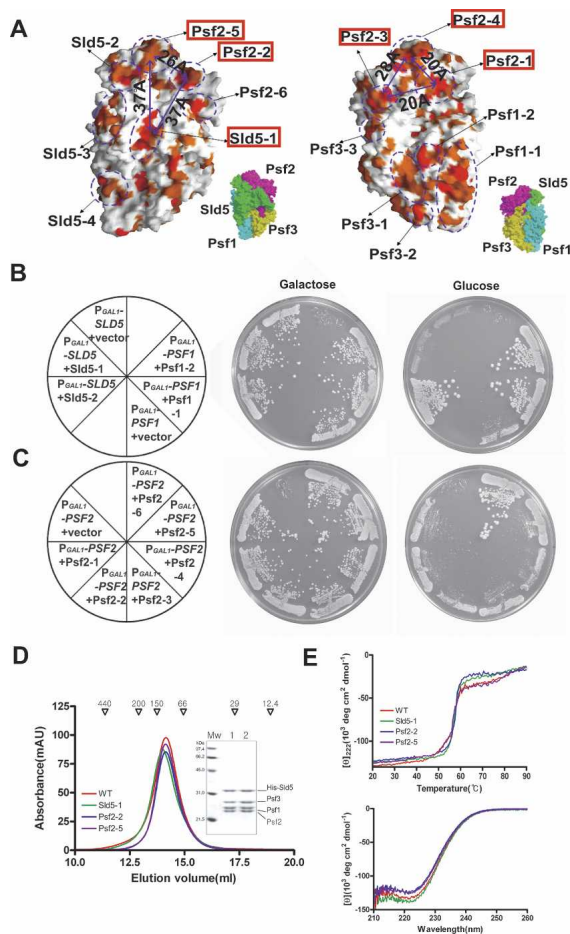


**Figure 3.** Interfaces of the GINS complex. (A) A close-up view of the intermolecular four-helical bundle formed by helices H3 and H5 of Sld5 (blue) and helices H3 and H5 of Psf1 (orange). Residues that are mutated in our analysis are boxed with red or blue. In addition to the extensive hydrophobic interactions, the two ion pairs and the five H-bonds further strengthen this interface (see Supplementary Fig. S3 for details). (B) Psf2 (magenta) and Psf3 (green) dimerize to form an intermolecular four-helical bundle. Helices H5 and H6 of Psf2 interact with helices H8 and H10 of Psf3 primarily through hydrophobic interactions. (C) The atomic interactions between the N-terminal domain of Sld5 (blue) and the N-terminal domain of Psf2 (magenta). Helices H1 and H3 of Sld5 interact with helices H1 and H2 and with strands S2 and S4 of Psf2. (D) The atomic interactions between the N-terminal domain of Psf1 (orange) and the N-terminal domain of Psf3 (green). This interface is very similar to that of Sld5–Psf2 shown in C.

between residues from helices H5 and H6 of Psf2 and residues from helices H8 and H10 of Psf3. In addition, seven H-bonds and the three ion pairs further stabilize the Psf2–Psf3 interface. Helix H6 of Psf2 and helix H8 of Psf3 exhibit notable bending, which allows for tighter interactions in this interface (Fig. 3B; Supplementary Fig. S3B). At the center of the intermolecular four-helix bundle, conserved Phe159 and Leu160 of Psf2 pack against Phe146 and Leu175 of Psf3. Near this center, side chains of Asp151 and Gln154 of Psf3, and Arg129 and Ser156 of Psf2 form a network of H-bonds and an ion pair to further stabilize the Psf2–Psf3 interface.

The Sld5–Psf1 complex contains an additional interface that is not present within the Psf2–Psf3 complex. In this interface, the linker (residues 149–162) between the N-terminal and C-terminal domains of Sld5 interacts with the H2 and H3 helices of a four-helical bundle of Psf1. In addition, the C-terminal loop (residues 128–143) of Psf1 interacts with the H2 and H3 helices of Sld5.

**The Sld5–Psf2 and Psf1–Psf3 interfaces** The N-terminal and C-terminal domains of Sld5 interact with the



**Figure 4.** Mutant complementation analysis of yeast GINS. (A) A surface representation of the human GINS structure with the degree of sequence conservation encoded by the red color. Invariant residues among all five species are colored red, and less conserved residues (80%) are colored orange. Fifteen surface patches for mutations are circled. Surface patches that affect cell viability are marked with a box, and distances between these patches are shown. (B) The lethality of mutant strains ( $P_{GALI}$ - $SLD5$  and  $P_{GALI}$ - $PSF1$ ) transformed with a control vector or with corresponding yeast GINS surface mutant plasmids was tested on SC-Ura media containing different carbon sources, such as galactose (left) and glucose (right). (C) A complementation test of a mutant strain,  $P_{GALI}$ - $PSF2$ , that was transformed with a control vector or with various corresponding yeast Psf2 surface mutant plasmids was performed on SC-Ura + Gal (left) and SC-Ura Glu (right) plates. All plates were incubated at 30°C for 2–3 d. (D) Gel filtration analysis showed that the surface mutant proteins of the human GINS complex, which affects the cell growth, form a tetramer. For clarity, only gel filtration profiles of Sld5-1, Psf2-2, and Psf2-5 are shown. On the right side, an SDS gel is shown for the wild-type (lane 1) and Sld5-1 mutant (lane 2) GINS complexes. Other surface mutants showed the same results and are omitted. (E, top) CD analysis showed that the surface mutations on the Sld5-1 patch (or other Psf2 mutants) did not affect the overall conformation of the GINS complex. For clarity, only Sld5-1, Psf2-2, and Psf2-5 profiles are shown. (Bottom) Temperature scanning analysis showed that the  $T_m$  values of the wild-type GINS complex and the six surface mutant GINS proteins are basically the same.

N-terminal and the C-terminal regions of Psf2, respectively (Figs. 1, 3C; Supplementary Fig. S3C). This interface is the most extensive of the interfaces in the GINS complex. Although two ion pairs and six H-bonds are present, the interface is mainly formed by hydrophobic interactions. In the N-terminal region, helices H1 and

H3 of Sld5 primarily interact with helices H1 and H2 and strands S2 and S4 of Psf2. Conserved Phe8 and Leu9 of Psf2 form a hydrophobic environment with Trp34 and Tyr87 of Sld5. In another hydrophobic patch, conserved Phe21, Leu33, Trp47, and Leu 48 of Psf2 form an extensive van der Waals network with Leu73, Ile77, and Met80 of Sld5.

The Psf1–Psf3 interface is very similar to the N-terminal–N-terminal interface formed by Psf2 and Sld5 (Fig. 3C,D; Supplementary Fig. S3D). Residues from helices H1 and H3 of Psf1 interact with the residues in the N-terminal loop, helices H2 and H3, and strands S2 and S4 of Psf3.

At the C-terminal interface of Sld5–Psf2, residues from four strands (S1 to S4) of Sld5 create an extensive van der Waals network with two helices, H5 and H6, of Psf2. An equivalent interface is absent in the Psf1–Psf3 complex. In addition to the four major interfaces, the two minor interfaces exist between Sld5 and Psf2 and between Psf1 and Psf3 (Supplementary Fig. S4).

#### Conserved surface regions with functional significance

Since the primary function of GINS is to mediate the interactions between replication factors, we wanted to identify the parts of the GINS complex that participate in this function. We searched extensively for highly conserved surface residues in the GINS complex and identified fifteen surface patches in which the invariant and exposed residues are clustered (Fig. 4A). Next, we displaced residues in each surface patch of yeast GINS and constructed 15 single or multiple mutants (Supplementary Table S2). The plasmids containing the mutated genes were transformed to yeast strains in which GINS genes are under the control of a *gal* promoter ( $P_{GALI}$ - $SLD5$ ,  $P_{GALI}$ - $PSF1$ ,  $P_{GALI}$ - $PSF2$  and  $P_{GALI}$ - $PSF3$ ) for mutant complementation assays (Kanemaki et al. 2003; Takayama et al. 2003). These strains grow in a carbon source-dependent manner, and the lethality of GINS mutants is consistent with the previous studies of knockout strains of yeast GINS complex (Kanemaki et al. 2003; Takayama et al. 2003).

Mutations on Asn69, Glu75, and Leu77 of Sld5 (Sld5-1) did not complement the cell lethality in glucose medium (Fig. 4B). Several of the Psf2 mutations also showed defects in growth in glucose media to levels as severe as that of the  $P_{GALI}$ - $PSF2$  strain containing a negative control (Fig. 4B,C; Supplementary Table S2). These residues include Gln17, Glu21 (Psf2-1), Gln91 (Psf2-2), Asp99, Glu140, Arg142, Gln146, Asp147, Glu150 (Psf2-3), Arg30, Gln81, Lys84, Thr32, Ser83, Ser86 (Psf2-4), Lys36, Asp56, and Asn61 (Psf2-5). The human GINS mutant proteins that are equivalent to these mutants form a tetramer (Fig. 4D). Moreover, circular dichroism (CD) analysis of the equivalent human GINS mutant proteins confirms that the mutation did not affect the overall structure and stability of the GINS complex (Fig. 4E). Thus, it is likely that these mutations on Sld5 and Psf2 affected cell viability by inhibiting the interaction between GINS and other key regulatory proteins in replication.

We then examined the locations of the mutated residues on the surface of the GINS complex. A surface patch of mutant Sld5-1 is located at loop H1–H2 and helix H1 of Sld5 (Figs. 1, 3A). The two surface patches of the Psf2-2 and Psf2-5 mutants are located on the same

face of the surface, and are located ~26–37 Å away from the Sld5-1 patch (Fig. 4A). Three additional patches are located on the opposite face of the surface. A patch of mutant Psf2-1 is located on a small pocket that is formed by helices H1, H2, and H4 of Psf2 and helix H3 of Sld5. About 20 Å from this region, the Psf2-3 patch is located on helices H3 and H5 of Psf2, and the patch of Psf2-4 is located on loops S1–S2, H2–S5, helix H2, and strand S5 of Psf2 (Fig. 4A). The two patches of Psf2-3 and Psf2-4 are separated by 28 Å. Overall, these six patches that affected cell growth are clustered together on top of the surface shown in Figure 4A. Based on our analysis, we propose that these surface patches in Sld5 and Psf2 play crucial roles in recruiting replication factors.

Although these patches are clustered on the surface of the GINS complex, the end-to-end distance between each of the six patches can vary up to 50 Å. A single cellular protein might interact with all of the six patches. However, the area spanned by six patches is rather broad to accommodate a single protein, and thus it is likely that multiple proteins bind to these patches. Thus, based on structural and genetic analyses, we propose that this region within GINS might mediate the assembly of proteins in an RPC (Calzada et al. 2005; Gambus et al. 2006; Pacek et al. 2006).

What are the factors that interact with these regions in Psf2? Although the binding partner of eukaryotic Psf2 has not been identified, archaeal GINS23 is known to interact with archaeal MCM (Marinsek et al. 2006). Components of the RPC such as Ctf4, Cdc45, or the checkpoint factors Mrc1 and Tof1–Csm3 complex may bind to these sites in Psf2 (Gambus et al. 2006). Also, components of Pol  $\epsilon$  such as Dpb2 may bind to this region (Takayama et al. 2003). It is possible that Psf2 interacts with proteins that are not part of the RPC. It has been shown that Psf2 is involved in the regulation of chromosome segregation. Interestingly, a database search revealed that the structure of Psf2 most closely resembles that of the MukF protein (Fennell-Fezzie et al. 2005), which is involved in the separation of sister chromatids. Thus, Psf2 may recruit cellular proteins with other biological functions, such as chromosome segregation, and coordinate the interactions between these cellular proteins and replication factors.

In contrast to the surface mutants of Sld5 and Psf2, none of the Psf1 surface mutants were found to affect cell viability (Fig. 4B). Why did mutations on Psf1 and Psf3 fail to affect cell viability? It is possible that Psf1 or Psf3 interacts with its partners more extensively than we expected. Alternatively, Psf1 or Psf3 might interact with its partners through a less conserved region that was not targeted in our study.

The structure of the GINS complex shows that its component shares an extensive interface. If the main function of the GINS complex is to mediate the interactions between the MCM helicase and other replication factors in an RPC, then four proteins of the GINS complex must maintain the complex formation, and it would be difficult to perturb the intermolecular interface within the GINS complex by any simple mutation. We created multiple mutants to perturb the intermolecular interface within the GINS complex and performed yeast complementation analysis. We simultaneously mutated the highly conserved surface residues of each protein in yeast GINS, which are buried upon the formation of the interface. For instance, we mutated the two residues of

Psf1 in the Sld5-binding surface and two additional residues of Psf1 in the Psf3-binding surface. All of the mutated residues in yeast and their equivalents in human GINS are shown in Figure 3 and Supplementary Table S2. All of the mutants were found to form a tetramer and complement the lethality of corresponding mutant strains in glucose (data not shown), which suggests that the interaction of each component of GINS is so extensive that even multiple mutations in the interface cannot perturb the GINS complex.

In summary, we describe the structures of the four proteins of GINS and provide the basis for the assembly of these proteins into the GINS complex. Our structure, in conjunction with studies of archaeal GINS proteins, clearly demonstrates that the four GINS components diverged from a common ancestor (Makarova and Koonin 2003; Makarova et al. 2005). The presence of GINS in both archaea and eukaryotes suggests that the role of GINS is likely conserved, and is involved in regulating the formation and the progression of the replication forks (Gambus et al. 2006). However, a single component of the archaeal GINS has diverged into the four different proteins in eukaryotic GINS, and these structural variations would allow GINS to interact with various sets of cellular proteins and to perform more complicated functions in eukaryotes.

In contrast to the reported EM analysis in which *Xenopus* GINS forms a ring-like structure in which DNA can pass through its channel, the crystal structure of the human GINS complex exhibits a channel that is too small to accommodate DNA. All three GINS complexes in an asymmetric unit exhibit virtually identical structures, and thus it is unlikely that our observation of a small channel is due to a crystal packing effect. Although the basis for such differences between the two structures is unclear, we expect the structures of *Xenopus* and human GINS to be similar. Although we removed the portions of Sld5 and Psf1 for crystallization, the truncation should increase rather than decrease the size of a channel. One possible reason for the observed differences in the central channels between the two GINS proteins might be the different conditions used in structural determination. For EM images of a rotary-shadowed GINS complex, certain conditions such as high glycerol content, low pH, use of a dried sample, and a vacuum state were employed (Kubota et al. 2003). Under these conditions, surface crevice regions in the central part of the GINS complex might appear as a large cavity in images. In addition, the differences in resolution between the EM structure of *Xenopus* GINS and the crystal structure of human GINS may contribute to the differences observed in the central channels.

Based on the EM studies of *Xenopus* GINS, it has been proposed that (1) the two ring-shaped proteins, GINS and the MCM hexamer, could stack on top of each other in the Cdc45–MCM–GINS complex, which may allow the MCM helicase to interact with one or both strands of the DNA molecule (Moyer et al. 2006); and (2) GINS may be reminiscent of PCNA in its structural and functional aspects to confer high processivity of Pol  $\epsilon$  in leading strand synthesis (Shikata et al. 2006). However, our present studies suggest that the primary role of GINS is to mediate the protein–protein interactions rather than to secure the interactions of proteins and DNA, and GINS might stimulate DNA synthesis catalyzed by Pol  $\epsilon$  through a mechanism other than the sliding clamp

mechanism. The structural and genetic studies of the GINS complex presented here should help facilitate further genetic and biochemical analysis to identify interacting partners and elucidate the functions of GINS.

## Materials and methods

### Crystallization and data collection

Crystals were grown by the hanging drop vapor diffusion method against a crystallization buffer containing 0.1 M Tris-HCl, 2.5 M ammonium sulfate, and 0.2 M lithium sulfate (pH 7.2) at 4°C. The crystals formed in space group C2 with  $a = 164.9 \text{ \AA}$ ,  $b = 177.0 \text{ \AA}$ ,  $c = 126.5 \text{ \AA}$ , and  $\beta = 105.9^\circ$ , and contained three complex molecules in an asymmetric unit. Diffraction data were collected at  $-170^\circ\text{C}$ , with crystals flash-frozen in crystallization buffer containing 30% glycerol.

### Structure determination and refinement

A single-wavelength data set was collected by using a Se-Met crystal on beam line 4A at Pohang Advanced Light Source. Integration, scaling, and merging of the diffraction data were performed by using the HKL2000 program suites (Otwinowski and Minor 1997). Sixty-one initial Se sites were identified, and initial phases were determined by using the CNS program (Brünger et al. 1998). After density modification (solvent flattening and threefold averaging), the electron density map calculated to 3.0-Å resolution was of excellent quality, allowing us to trace most of the chains. The model was refined against a 2.8-Å Se-Met data collected at a wavelength of 1.0 Å. Successive rounds of model building using O (Kleywegt and Jones 1996), refinement using the CNS program (Brünger et al. 1998), and phase combination allowed us to build the complete structure (Supplementary Table S1). The final model contained residues 21–212 (except 69), 1–145, 1–174, and 3–194 (except 49–57) of Sld5, Psf1, Psf2, and Psf3, respectively.

Expression and purification, yeast complementation assay, mutagenesis, and stability measurement of the mutant are described in the Supplemental Material.

## Acknowledgments

We thank anonymous reviewers and Hae-Sung Jeon for their invaluable comments. We also thank Heung Soo Lee, Kyung Jin Kim, and Kyung Hwa Kim for help with data collection (PAL4A and PAL6C). This work was supported by funds from the National Creative Research Initiatives (Ministry of Science and Technology). The coordinates and structure factor amplitudes of GINS have been deposited in the Protein Data Bank (PDB identification code 2EHO).

## References

- Bell, S.P. and Dutta, A. 2002. DNA replication in eukaryotic cells. *Annu. Rev. Biochem.* **71**: 333–374.
- Brünger, A.T., Adams, P.D., Clore, G.M., DeLano, W.L., Gros, P., Grosse-Kunstleve, R.W., Jiang, J.S., Kuszewski, J., Nilges, M., Pannu, N.S., et al. 1998. Crystallography and NMR system: A new software suite for macromolecular structure determination. *Acta Crystallogr. D Biol. Crystallogr.* **54**: 905–921.
- Calzada, A., Hodgson, B., Kanemaki, M., Bueno, A., and Labib, K. 2005. Molecular anatomy and regulation of a stable replisome at a paused eukaryotic DNA replication fork. *Genes & Dev.* **19**: 1905–1919.
- De Falco, M., Ferrari, E., De Felice, M., Rossi, M., Hubscher, U., and Pisani, F.M. 2007. The human GINS complex binds to and specifically stimulates human DNA polymerase  $\alpha$ -primase. *EMBO Rep.* **8**: 99–103.
- Fennell-Fezzie, R., Gradia, S.D., Akey, D., and Berger, J.M. 2005. The MukF subunit of *Escherichia coli* condensin: Architecture and functional relationship to kleisins. *EMBO J.* **24**: 1921–1930.
- 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.
- Huang, H.K., Bailis, J.M., Levenson, J.D., Gomez, E.B., Forsburg, S.L., and Hunter, T. 2005. Suppressors of Bir1p (Survivin) identify roles for the

- chromosomal passenger protein Pic1p (INCENP) and the replication initiation factor Psf2p in chromosome segregation. *Mol. Cell. Biol.* **25**: 9000–9015.
- Kanemaki, M. and Labib, K. 2006. Distinct roles for Sld3 and GINS during establishment and progression of eukaryotic DNA replication forks. *EMBO J.* **25**: 1753–1763.
- Kanemaki, M., Sanchez-Diaz, A., Gambus, A., and Labib, K. 2003. Functional proteomic identification of DNA replication proteins by induced proteolysis in vivo. *Nature* **423**: 720–724.
- Kleywegt, G.J. and Jones, T.A. 1996. Efficient rebuilding of protein structures. *Acta Crystallogr. D Biol. Crystallogr.* **50**: 829–832.
- Krishna, T.S.R., Kong, X.P., Gary, S., Burgers, P., and Kuriyan, J. 1994. Crystal structure of the eukaryotic DNA polymerase processivity factor PCNA. *Cell* **79**: 1233–1243.
- Kubota, Y., Takase, Y., Komori, Y., Hashimoto, Y., Arata, T., Kamimura, Y., Araki, H., and Takisawa, H. 2003. A novel ring-like complex of *Xenopus* proteins essential for the initiation of DNA replication. *Genes & Dev.* **17**: 1141–1152.
- Makarova, K.S. and Koonin, E.V. 2003. Comparative genomics of archaea: How much have we learned in six years, and what's next? *Genome Biol.* **4**: 115.
- Makarova, K.S., Wolf, Y.I., Mekhedov, S.L., Mirkin, B.G., and Koonin, E.V. 2005. Ancestral paralogs and pseudoparalogs and their role in the emergence of the eukaryotic cell. *Nucleic Acids Res.* **33**: 4626–4638.
- Marinsek, N., Barry, E.R., Makarova, K.S., Dionne, I., Koonin, E.V., and Bell, S.D. 2006. GINS, a central nexus in the archaeal DNA replication fork. *EMBO Rep.* **7**: 539–545.
- Moyer, S.E., Lewis, P.W., and Botchan, M.R. 2006. Isolation of the Cdc45/Mcm2–7/GINS (CMG) complex, a candidate for the eukaryotic DNA replication fork helicase. *Proc. Natl. Acad. Sci.* **103**: 10236–10241.
- Obama, K., Ura, K., Satoh, S., Nakamura, Y., and Furukawa, Y. 2005. Up-regulation of PSF2, a member of the GINS multiprotein complex, in intrahepatic cholangiocarcinoma. *Oncol. Rep.* **14**: 701–706.
- Otwinowski, Z. and Minor, W. 1997. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**: 307–326.
- Pacek, M., Tutter, A.V., Kubota, Y., Takisawa, H., and Walter, J.C. 2006. Localization of MCM2–7, Cdc45, and GINS to the site of DNA unwinding during eukaryotic DNA replication. *Mol. Cell* **21**: 581–587.
- Shikata, T., Sasa-Masuda, T., Okuno, Y., Waga, S., and Sugino, A. 2006. The DNA polymerase activity of Pol  $\epsilon$  holoenzyme is required for rapid and efficient chromosomal DNA replication in *Xenopus* egg extracts. *BMC Biochem.* **7**: 21.
- Takayama, Y., Kamimura, Y., Okawa, M., Muramatsu, S., Sugino, A., and Araki, H. 2003. GINS, a novel multiprotein complex required for chromosomal DNA replication in budding yeast. *Genes & Dev.* **17**: 1153–1165.
- Ueno, M., Itoh, M., Kong, L., Sugihara, K., Asano, M., and Takakura, N. 2005. PSF1 is essential for early embryogenesis in mice. *Mol. Cell. Biol.* **25**: 10528–10532.