# Crystal structure of the GINS complex and functional insights into its role in DNA replication

Y. Paul Chang*†, Ganggang Wang†, Vladimir Bermudez‡, Jerard Hurwitz‡§, and Xiaojiang S. Chen*†§

*Graduate Program in Genetic, Molecular, and Cell Biology, and †Section of Molecular and Computational Biology, University of Southern California, Los Angeles, CA 90089; and ‡Molecular Biology Program, Memorial Sloan–Kettering Cancer Center, 1275 York Avenue, Box 97, New York, NY 10021

Contributed by Jerard Hurwitz, June 13, 2007 (sent for review May 21, 2007)

**The GINS complex, which contains the four subunits Sld5, Psf1, Psf2, and Psf3, is essential for both the initiation and progression of DNA replication in eukaryotes. GINS associates with the MCM2-7 complex and Cdc45 to activate the eukaryotic minichromosome maintenance helicase. It also appears to interact with and stimulate the polymerase activities of DNA polymerase $\varepsilon$ and the DNA polymerase $\alpha$-primase complex. To further understand the functional role of GINS, we determined the crystal structure of the full-length human GINS heterotetramer. Each of the four subunits has a major domain composed of an $\alpha$-helical bundle-like structure. With the exception of Psf1, each of the other subunits has a small domain containing a three-stranded $\beta$-sheet core. Each full-length protein in the crystal has unstructured regions that are all located on the surface of GINS and are probably involved in its interaction with other replication factors. The four subunits contact each other mainly through $\alpha$-helices to form a ring-like tetramer with a central pore. This pore is partially plugged by a 16-residue peptide from the Psf3 N terminus, which is unique to some eukaryotic Psf3 proteins and is not required for tetramer formation. Removal of these N-terminal 16 residues of Psf3 from the GINS tetramer increases the opening of the pore by 80%, suggesting a mechanism by which accessibility to the pore may be regulated. The structural data presented here indicate that the GINS tetramer is a highly stable complex with multiple flexible surface regions.**

Cdc45 | DNA helicase | minichromosome maintenance complex | DNA polymerase

Eukaryotic DNA replication is controlled by a series of ordered and regulated steps (1–3) that commence with the binding of the six-subunit origin recognition complex (ORC) to replication origins. During the G1 phase of the cell cycle, Cdc6, and Cdt1 are recruited to the origin, and together with ORC, support the loading of the heterohexameric MCM2-7 complex (minichromosome maintenance, MCM) to form the prereplication complex (pre-RC). Although a substantial amount of data suggest that MCM acts as the replicative helicase, MCM present in the pre-RC (as well as isolated MCM) is devoid of helicase activity (summarized in ref. 4). At the $G_1$/S transition of the cell cycle, it appears that the MCM helicase activity is activated by a complex and an as yet poorly understood series of modifications that require the action of two protein kinases, DDK (Cdc7-Dbf4) and CDK (cyclin-dependent), as well as the participation of at least eight additional factors, including Mcm10, Cdc45, Dpb11, GINS, synthetic lethal with *dpb11* mutant-2 (Sld2), and Sld3 (4). Two of these components, Cdc45 and GINS, appear critical for helicase activation because DNA unwinding is observed (3, 5), concomitant with their loading at origins. In accord with these findings, a complex containing near-stoichiometric levels of MCM, Cdc45, and GINS was isolated from *Drosophila* and shown to possess DNA helicase activity (6). Studies with *Xenopus* extracts revealed that a complex that included MCM, Cdc45, and GINS was found at sites at which replication forks were halted artificially by a streptavidin–biotin complex (7). In *Saccharomyces cerevisiae*, GINS was shown to play a critical role in supporting interactions between MCM and Cdc45, as well as a number of key regulatory proteins. They together formed a large replisome progression complex that migrated with the replication fork. Upon selective degradation of the Psf2 subunit of GINS, replication was halted and Cdc45 was no longer associated with MCM. These findings suggest that interactions between MCM and other key replication factors might be mediated by GINS. Collectively, they indicate that GINS is an essential component of the replicative machinery that moves with the replication fork.

GINS is a heterotetrameric complex consisting of Sld5, Psf1 (partner of Sld5-1), Psf2, and Psf3 and was first discovered by using a variety of genetic screens in *S. cerevisiae* (8). The four GINS subunits are paralogs, among which the specific subunit pairs Psf1–Sld5 and Psf2–Psf3 are more closely related (9). Each of the subunits is relatively small ($\approx$200 aa) and highly conserved in all eukaryotes. In archaea, only two homologues, Gins15 and Gins23, have been identified that appear to interact and form a dimer of the heterodimer, suggesting that, like its eukaryotic counterpart, it is a tetramer (10). Direct interactions between the archaeal GINS complex and the archaeal MCM, as well as primase, have been reported (10). Other reports have recently appeared suggesting that GINS may serve as an accessory factor for eukaryotic DNA polymerases, including DNA polymerase (Pol) $\varepsilon$ (11) and the DNA Pol $\alpha$-primase complex (12).

Despite the essential role of GINS in DNA replication, how GINS interacts with MCM, Cdc45, and other protein factors at the replication fork remains unclear. To understand the structural/functional roles of GINS in replication, we crystallized the human GINS complex and determined its crystal structure. This complex included the full-length proteins of each of the four subunits. During the preparation of this work, the structure of the human GINS complex containing a truncated Psf1 subunit appeared (13). The tetramer structure that we have obtained is basically the same as that reported by Kamada *et al.* (13); the folds of each subunit and the interactions between the four subunits are essentially the same with slight variations found for certain loops and $\beta$-strands. However, our crystal structure revealed certain features not reported for the structure of the truncated complex reported by Kamada *et al.* (13) that may have important functional implications. The structural and muta-

BIOCHEMISTRY

Author contributions: Y.P.C., J.H., and X.S.C. designed research; Y.P.C., G.W., and V.B. performed research; Y.P.C., G.W., and X.S.C. analyzed data; and Y.P.C., G.W., V.B., J.H., and X.S.C. wrote the paper.

The authors declare no conflict of interest.

Abbreviations: MCM, minichromosome maintenance; ts, temperature-sensitive; Pol, polymerase.

Data deposition: The atomic coordinates and structure factors have been deposited in the Protein Data Bank, www.pdb.org (PDB ID code 2Q9Q).

§To whom correspondence may be addressed. E-mail: hurwitz@ski.mskcc.org or xiaojiang.chen@usc.edu.

This article contains supporting information online at www.pnas.org/cgi/content/full/0705558104/DC1.

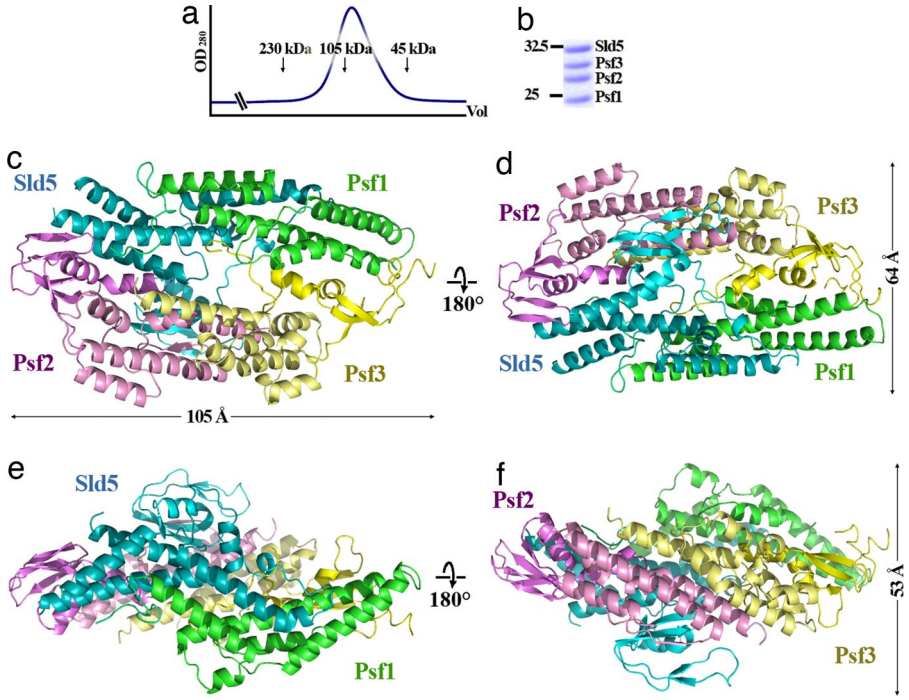


**Fig. 1.** The overall structural features of human GINS complex. (*a*) Superdex-200 gel-filtration profile of the human GINS complex crystallized. (*b*) The SDS/PAGE analysis of the peak fraction from *a*, showing that all four subunits are present as full-length proteins. (*c* and *d*) Two views of the wider faces of the GINS complex structure, showing the spindle-shaped structure and part of the central pore opening. The tetramer is composed of predominantly $\alpha$-helices that are arranged in $\alpha$-helical bundles. (*e* and *f*) Two views of the narrower sides of the GINS complex. The four subunits are shown in different colors as indicated.

tional data we obtained suggest that the dimension of a central pore in GINS appears to be regulated by a short N-terminal peptide of Psf3. The positions of disordered regions in our structure, including the C-terminal 51 residues of Psf1, colocalize on the surface of the GINS complex as patches and likely serve as interaction sites for the binding of GINS to its replication protein partners.

## Results

**Overall Structural Features of the GINS Complex.** The four full-length subunits of GINS were coexpressed in *Escherichia coli* and the complex purified to homogeneity. The isolated complex had an apparent molecular mass of ≈90 kDa as estimated from gel filtration chromatography (Fig. 1*a*) and glycerol gradient sedimentation (data not presented) consistent with a 1:1:1:1 molar ratio of the four different proteins in the tetrameric complex (Fig. 1*b*). We crystallized the GINS complex as described in *Materials and Methods*, and SDS/PAGE and mass spectrometric analyses confirmed that all four proteins present in the crystals were full-length. We determined the x-ray structure of the GINS complex to 2.36 Å resolution [statistics presented in supporting information (SI) Table 1]. Each asymmetric unit in a crystal cell contained two GINS heterotetramers with identical conformations. The overall morphology of the GINS tetramer complex resembles a slightly elongated spindle (Fig. 1 *c–f*) with a visible central hole (Fig. 1 *c* and *d*). The body of the tetramer is composed of $\alpha$-helices with few peripheral short $\beta$-strands. The gross structural features are essentially the same as those recently reported for the structure containing a truncated Psf1 subunit (13); in the reported structure, 47 residues of the C-terminal region of Psf1 had to be deleted for crystallization. Surprisingly, even though we crystallized GINS with all full-length proteins, only the first 145 residues of Psf1 were ordered. These residues were present in the structure containing truncated Psf1 reported by Kamada *et al.* (13). The C-terminal 51 residues of Psf1 are not visible in our structure, indicating that this region is intrinsically disordered.

**Structures of Individual GINS Subunits.** Each of the four subunits contains a major domain composed of $\alpha$-helices ($\alpha$-domains, Fig. 2 *a–d*). The folds of the $\alpha$-domains of all four subunits are similar; each contains four to five helices arranged more or less in a parallel fashion to form a partial three-helix bundle structure. In three of the four subunits (Sld5, Psf2, and Psf3), there is a small $\beta$-sheet composed of three antiparallel $\beta$-strands ($\beta$1, $\beta$2, and $\beta$3) near one end of the $\alpha$-domain. Around the $\beta$-sheet are two helices and a $\beta$-hairpin or loop, forming a small but definable $\beta$-domain in these three subunits. The monomeric structures overlap well with the reported GINS structure (13), especially within the $\alpha$-domains. However, conformational differences are noted in the $\beta$-domains of each subunit (Fig. 2 *e–g*). The differences in the $\beta$-domains and in a few loops are all located on the surface of the GINS complex, suggesting a certain degree of plasticity in the surface structures.

Psf1 has only an $\alpha$-domain (residues 1–145), whereas all of its C-terminal 51 residues are disordered. There is substantial space to accommodate the C-terminal 51 residues of Psf1 in at least one of the two GINS tetrameric complexes in the asymmetric unit. Nonetheless, no definable electron density can be seen for the C-terminal 51 residues of Psf1 (residues 146–196), despite the fact that sequence alignment suggests a fold similar to the $\beta$-domain that is present in the other three subunits (SI Fig. 7*a*). Kamada *et al.* (13) reported that their GINS complex crystallized only when a Psf1 mutant lacking the C-terminal 47 residues (14) was used, suggesting that the presence of this $\beta$-domain inhibited crystal packing of the GINS complex. Kamada *et al.* (13) proposed that the deleted region of Psf1 folds into a $\beta$-domain structure and that the correct positioning of this domain on the surface of the GINS complex is critical for function. However, we crystallized the GINS complex with all full-length proteins under

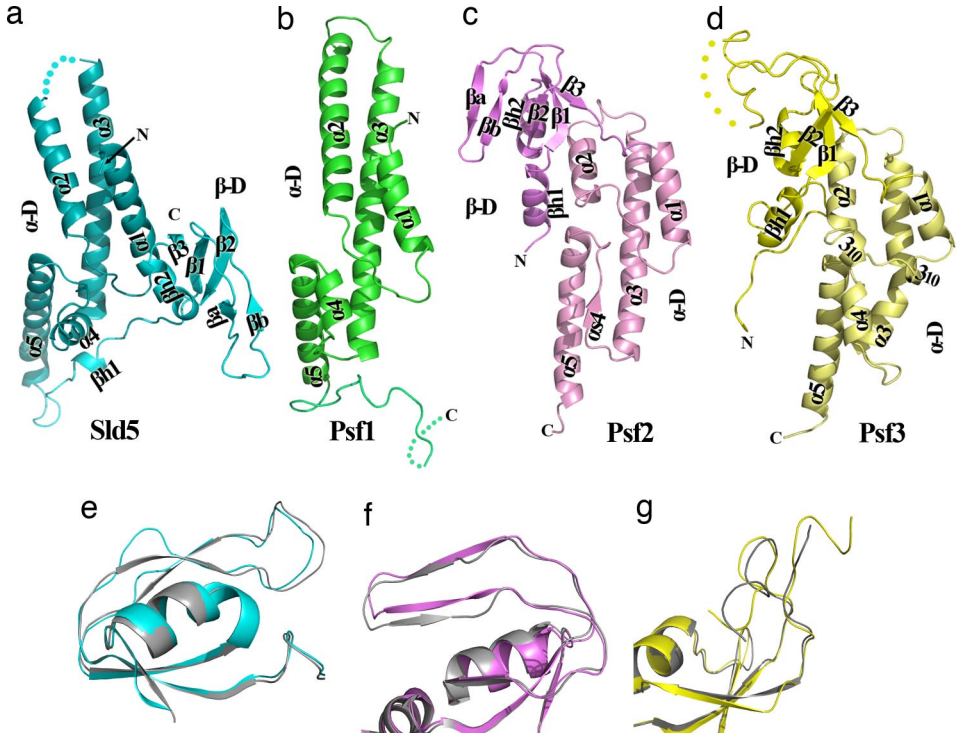


**Fig. 2.** The structural folds of individual subunits. (*a–d*) The detailed structures of each of the four subunits. N and C indicate the position of the N and C termini in each subunit, whereas dashed lines indicate the location of disordered fragments. The α-domains (α-D) and β-domains (β-D) for each subunit are indicated. Labels for the secondary structures in the α-domains start with α and labels for the β-domains start with β. (*e–g*) A comparison of the β-domains of our structure (in color) and that reported for the structure determined by Kamada *et al.* (13), showing conformational differences in Sld5 (e), Psf2 (*f*), and Psf3 (*g*).

different crystallization conditions but in the same space group with similar unit cell dimensions. Our structure revealed that the C-terminal 51 residues of Psf1 are not folded in the heterotetrameric GINS complex and, thus, are not anchored to the GINS surface. These findings suggest that the Psf1 residues proposed to anchor the β-domain on the GINS surface by Kamada *et al.* (13) could, instead, play a role in binding other protein partners.

Even though the GINS subunits have similar α- and β-domains, the relative arrangements of the two domains differ among the subunits, as predicted in ref. 9. In Sld5 and Psf1, the larger N-terminal portion forms the α-domain, and the smaller C-terminal fragment corresponds to the β-domain (SI Fig. 7*a*). In contrast, the order of α- and β-domains of Psf2 and Psf3 is reversed (i.e., the β-domain is at the N terminus and the α-domain is at the C terminus; SI Fig. 7*b*). In addition, the space between the α- and β-domains for Psf2/Psf3 is only 6 residues, but is 21 residues in Sld5 and possibly about the same length in Psf1, based on sequence alignment (SI Fig. 7 *a* and *b*). Despite the differences in spacer length, the β-domains present in Sld5, Psf2, and Psf3 appear anchored to their respective α-domains through direct contacts (Fig. 2 *a*, *c*, and *d*).

**Tetramer Formation.** The structure of the GINS tetramer was reported to be formed by an arbitrarily assigned "vertical" interface formed through the α-domains and a "horizontal" interface mediated through the β-domains (13). Despite the similarities of our GINS tetramer structure to that reported, we interpret the interactions that support tetramerization in our crystal structure somewhat differently, as described below. In our structure, each of the four subunits interacts with two other molecules to form a ring-like structure with a central hole (Fig. 3*a*). The four subunits are arranged around the ring in the order Sld5, Psf1, Psf3, and Psf2, such that Psf2 contacts Sld5 to complete the circle. Within the pore created by this ring, there are no direct bonding contacts between Sld5 and Psf3 or between Psf1 and Psf2. The intersubunit interactions are mainly through the sides of the α-helical bundles to generate extensive contacts between subunits with buried interface areas between two neighboring subunits ranging from 2,900 to 4,000 Å$^2$. These large interface areas presumably provide strong bonding forces at the subunit interfaces, which explains why the tetrameric complex is stable in solution. Major bonding forces are provided through helix–helix interactions between adjacent α-domains of different subunits (Sld5–Psf1, Psf1–Psf3, Psf3–Psf2, and Psf2–Sld5), mediated mostly by hydrophobic residues. However, the β-domains at the interface between Psf2–Sld5 also play a role in stabilizing the tetramer. The β2 of the Sld5 β-domain interacts with a small β-strand (αs4) from the Psf2 α-domain, expanding the Sld5 β-domain to a four-stranded β-sheet (Fig. 3*b*). Additionally, a β-hairpin from the Psf2 β-domain contacts an α-helix of the Sld5 α-domain (Fig. 3*c*).

**Possible Roles of the Unstructured Regions of Sld5, Psf1, and Psf3.** Although our crystal structure of the GINS complex included full-length proteins of all four subunits, portions of Sld5, Psf1, and Psf3 have no visible electron density (represented by dashed lines in Fig. 2 *a–d* and SI Fig. 7 *a* and *b*), likely due to the flexibility of these regions. Comparisons between our structure and that reported by Kamada *et al.* (13) reveal that the disordered regions in Sld5 and Psf3 are basically the same in both structures. Because our GINS constructs and crystallization conditions differed from those of Kamada *et al.* (13), the fact that similar regions are missing from the electron density maps of both structures suggests that these regions are highly flexible and unstructured in the GINS complex. The last visible C-terminal residue of Psf1 (S145) before the disordered C-terminal domain is adjacent to the disordered fragment of Sld5 (residue 65–71) and to the disordered C-terminal residues of Psf3 (residues 194–216); these disordered regions are shown as spheres in Fig.

BIOCHEMISTRY

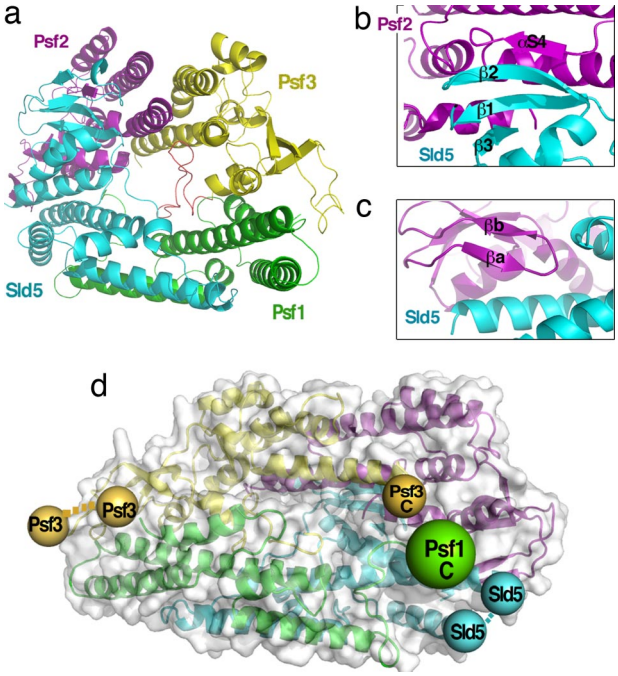


**Fig. 3.** The ring structure of the GINS complex and colocalization of disordered regions on the surface. (*a*) The ring structure of the GINS complex can be visualized along the long axis (the axis of the $\alpha$-helical bundles) of the tetramer. From this view, it is clear that most of the intersubunit interactions are mediated through $\alpha$-helices. The central opening is also apparent. (*b* and *c*) These views show interactions between Sld5 and Psf2, which involve $\beta$-strands at two locations on intersubunit interfaces. (*d*) Colocalization of the disordered regions on the tetramer surface. Each sphere indicates the location of the residue immediately adjacent to the disordered fragment. The subunit (and sphere) color scheme is the same as that used in *a*. The disordered fragments from three subunits colocalize on one end of the tetramer. As shown, the disordered region of Psf3 is on the same face but at the opposite end. These disordered regions may become structured when bound to interacting partners.

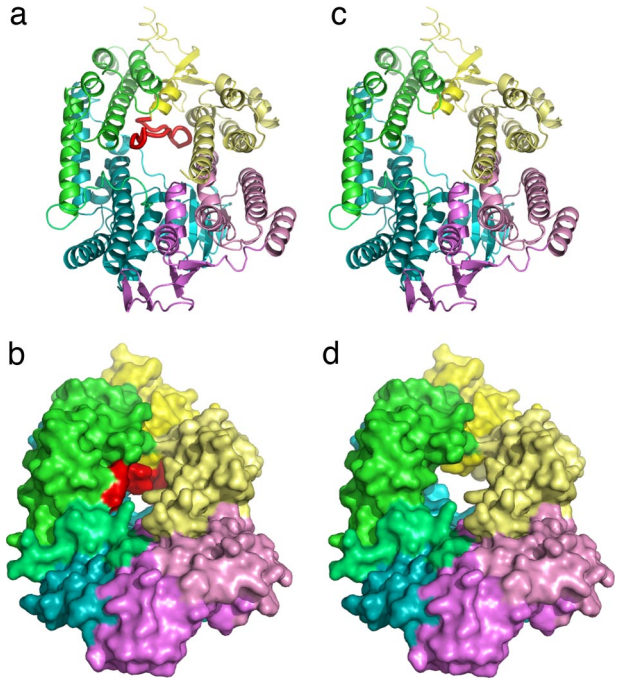


**Fig. 4.** The structure of the central pore of GINS and its accessibility. (*a*) A view of the central pore of the GINS tetramer. The red loop inside the pore is formed by the N-terminal 16 residues of Psf3. The loop is not tightly bonded to the pore surface. (*b*) The surface representation of the view in *a*. (*c*) The same view as in *a* but without the N-terminal residues of Psf3. (*d*) The surface representation of the view shown in *c*.

*3d*. The colocalization of these disordered parts of three different subunits on the GINS surface suggests that this site may bind partner proteins in the replication complex. On the same side, but located at the other end of the tetramer, is the disordered region within the $\beta$-domain of Psf3 (Fig. 3*d*), which may also serve as a protein-binding site. These disordered regions located on the surface of GINS may become structured upon binding to proteins known to interact with GINS, such as MCM, Cdc45, and DNA polymerases. Most of these are large proteins or complexes and the location of disordered sites on opposite sides of the tetramer may allow binding of more than one of these factors at the same time. In keeping with this idea, previous studies on *Xenopus* reported that the chromatin loading of Cdc45 and GINS were mutually dependent (15). Furthermore, Kamada *et al.* (13) observed that the human GINS complex containing the Psf1 subunit with the flexible C-terminal region truncated (labeled Psf1 C in Fig. 3*d*) failed to bind to the *Xenopus* pre-RC and failed to load Cdc45. These findings support our notion that the flexible regions present on the GINS surface are important for its binding to replicative proteins.

**Accessibility of the Central Pore.** When visualized by negatively stained EM, the recombinant *Xenopus* GINS complex has a ring-like structure with a central hole (15). Kamada *et al.* (13) suggested that the ring-like shape observed by EM was an artifact resulting from low-density distribution of electrons at the center of the complex. Examination of our GINS structure, however, revealed that it does have ring-like structure when viewed along the central pore, as shown in Fig. 4*a*. Based on EM measurements, the ring diameter was reported to be ≈9.5 nm (95 Å), with the central pore ≈4 nm (40 Å) in diameter (15); in contrast, the diameter of the ring in our crystal structure is, at maximum, ≈78 Å (measured from edge to edge in the large dimension) and the central pore is 10 Å in diameter (Fig. 4 *a* and *b*). These quantitative differences in dimensions could be due to flattening (deformation) and dehydration effects intrinsic to negative staining methods used to prepare EM samples.

Detailed examination of the central pore present in our GINS crystal structure revealed that a 16-residue peptide loop from the N terminus of Psf3 appeared to fit loosely into the pore, effectively restricting the opening (Fig. 4 *a* and *b*, red). Interactions of this peptide loop with the surface of the pore are limited, suggesting that the peptide may enter and leave the pore without a significant energy barrier. Interestingly, sequence alignment of this 16-aa sequence at the N terminus of Psf3 revealed that it is present in human and some higher eukaryotes but not in many other organisms (SI Fig. 8*a*). This sequence data suggest that the first 16 N-terminal residues of Psf3 may not be needed for the structural integrity of the GINS tetramer. A speculative role of this 16-residue peptide may be to regulate the dimensions of the central pore, and thus its accessibility, by plugging and unplugging this cavity.

To test the notion that this N-terminal 16-residue peptide is not required for tetramer formation and stability of the GINS complex, we generated Psf3 constructs lacking either 10 or 18 residues of the N terminus and examined whether such constructs supported tetramer formation. Both Psf3 constructs formed stable tetramers under high and low salt conditions and behaved similarly to the GINS complex containing full-length Psf3 in gel filtration analyses (SI Fig. 8*b*). This finding indicates

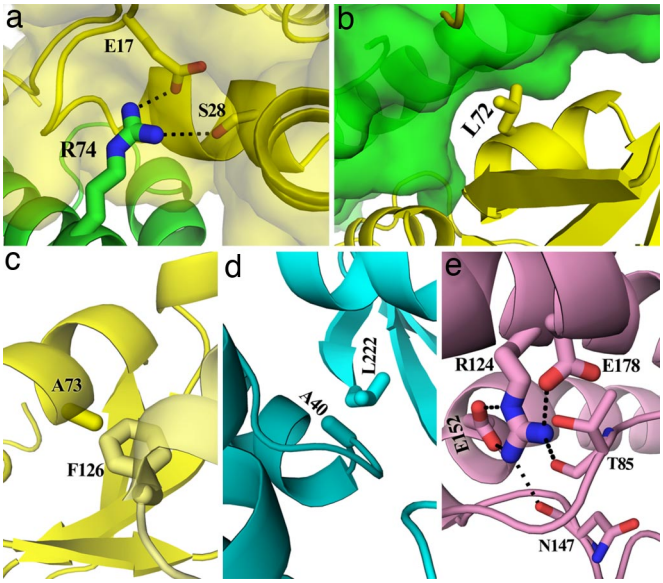


**Fig. 5.** The mapping of the yeast ts mutants on the equivalent sites present in human GINS structure. (*a* and *b*) The class I ts mutant residues involved in the intersubunit interactions between Psf1 R74 and Psf3 (*a*) and between Psf3 L72 and Psf1 (*b*). (*c*–*e*) The class II ts mutant residues Psf3 A73, Sld5 L222, and Psf2 R124 are involved in intrasubunit interactions. Mutation of these residues, as indicated, would be expected to affect the folding and stability of individual subunits within the GINS complex.

that the N-terminal 16-residue loop of Psf3 is not needed for folding and stability of the GINS complex.

The diameter of the central pore is ≈10 Å at its narrowest point, but the opening is increased to 18 Å upon removal of the N-terminal 16-residue loop of Psf3 (Fig. 4 *c* and *d*). This finding may also partially explain the larger pore size observed in EM (15) because the acidic negative staining solution used for EM may dislodge the N-terminal peptide of Psf3 from the pore. Thus, the 16-residue peptide at the N terminus of human Psf3 might regulate pore opening and closing. Although a functional role for the central pore is as yet unknown, we speculate that one possible role is to bind and hold a fragment from MCM, Cdc45, or DNA polymerases during DNA synthesis. It is also possible that the pore could bind ssDNA because its dimension is sufficiently large to accommodate ssDNA even with the N-terminal peptide of Psf3 situated within the pore. In support of this possibility, there are 10 Arg and Lys residues, as well as several Asn and Gln amino acids distributed on the surface of the pore (SI Fig. 9), despite the overall negatively charged outer surface of the tetramer. This charge distribution is similar to that of the proliferating cell nuclear antigen (PCNA) and T7 gp4 helicase, both of which have positively charged and polar residues around the central openings even though they have an overall negatively charged outer surface (14, 16). The GINS complex might require an additional factor(s) to bind ssDNA because DNA-binding activity has not as yet been detected for the GINS complex isolated from human (data not shown) or other organisms (6, 11).

**Understanding Temperature-Sensitive (ts) Mutants of GINS.** A total of nine different ts mutations have been identified in the four yeast GINS proteins (8, 17, 18). These mutated residues are highlighted in the sequence alignments shown in SI Fig. 7 *a* and *b*, eight of which are highly conserved across eukaryotes. Our GINS structure provides a rationale for all of these ts mutations. The nine ts mutations can be divided into three classes according to their possible defect(s) at nonpermissive temperatures. The

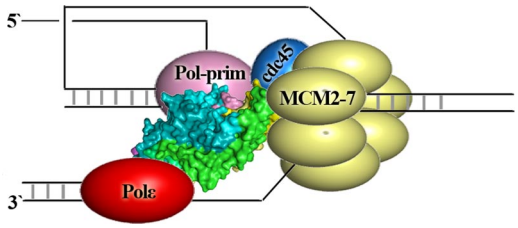


**Fig. 6.** A model for the GINS complex coordinating MCM, Cdc45, Pol ε, and Pol α-primase complex at the replication fork. The GINS complex is shown in surface representation. All other components (drawn as cartoons) are shown as bound directly to the GINS complex.

mutants in class I involve residues important for intersubunit interactions. Psf1-R74G (Fig. 5*a*) and Psf3-L72P (Fig. 5*b*) belong to this class. Psf1 R74 (yeast Psf1–1 R84G) forms hydrogen bonds with Psf3 E17 and S28 (Fig. 5*a*); a Gly substitution at this position should abolish the two hydrogen bonds and significantly weaken interactions between Psf1 and Psf3, which may lead to instability of the complex at nonpermissive temperatures. Similarly, the Psf3 L72P mutation (yeast Psf3-21 L46P) should reduce the hydrophobic packing with Psf1 F64 (Fig. 5*b*). The ts mutants in class II involve residues important for intrasubunit interactions. Psf3 A73 (yeast Psf3-21 A47P) and Sld5 L222 (yeast Sld5-13 L293P) are both involved in β-domain packing, as well as α- and β-domain interactions (Fig. 5 *c* and *d*). Psf2 R124 (yeast Psf2-209 R133K) forms five hydrogen bonds within three helices and two loops in its α-domain (Fig. 5*e*). Class III ts mutants involve residues buried in the hydrophobic core structures of Sld5 and Psf3. Mutations of these buried residues would be expected to reduce the thermodynamic stability of the structure, providing a plausible molecular explanation for the phenotypes of these mutants. Surprisingly, ts mutations have not as yet been identified on the surface of the GINS. Because GINS interacts directly with the MCM complex, Cdc45, and other replicative protein in different species (4), it is likely that the binding sites for these proteins are conserved. We speculate that mutations on the surface of GINS that result in temperature sensitivity may help identify their binding sites.

**Structural Conservation of GINS in Archaea.** The archaeon *Sulfolobus solfataricus* has two GINS homologs, Gins15 and Gins23 (10). Sequence analysis indicates that eukaryotic Psf1/Sld5 and Psf2/Psf3 are close paralogs of Gins15 and Gins23, respectively (9), suggesting that these two GINS proteins can form a tetrameric complex through the dimerization of the heterodimer. The two independently derived crystal structures of human GINS reveal that the residues involved in the interfaces between subunits are conserved among archaeal and eukaryotic GINS proteins (data not shown), providing further support for a structural and functional conservation between archaeal and eukaryotic GINS complexes. In keeping with these considerations, it was shown that the GINS complex formed in the archaeon is a tetramer formed through the dimerization of the Gins15–Gins23 heterodimer (10).

**A Model for GINS Role at the Replication Fork.** A speculative model of how the human GINS tetrameric complex interacts and coordinates the activities of its binding partners is proposed in Fig. 6. In this model, we suggest a direct contact between Psf1 and Pol ε because an interaction between *S. cerevisiae* Psf1 and Dpb2 (the second largest subunit of Pol ε) was detected in a yeast two-hybrid screen (8). Because archaeal Gins23 binds to the N terminus of MCM (10) and an interaction of Psf3 with MCM has been detected in the yeast two-hybrid system (S. Azuma and H. Masukata, unpublished data in ref. 18), our model depicts Psf3

in contact with MCM. We assume that Psf2 contacts the Pol $\alpha$-primase complex based on the report that archaeal Gins23 interacts with primase (10). The model also allows MCM, Cdc45, and GINS to contact each other based on the isolation of this complex from *Drosophila* eggs (6). Finally, we also suggest that GINS must interact with Pol $\varepsilon$ and the Pol $\alpha$-primase complex to coordinate leading- and lagging-strand synthesis, respectively (11, 12).

In summary, the structural and biochemical data presented both here and by Kamada *et al.* (13) suggest that GINS functions as a tight heterotetrameric complex. The molecular interactions between the subunits of GINS are mediated mostly through helix–helix interactions that amplify the helix-bundle-like structure of each individual subunit. In agreement with an EM study of the *Xenopus* GINS complex (15), we find an open pore along the long axis of the ring-like tetramer structure. The pore size appears to be influenced by the position of the N-terminal 16–20 aa residues of Psf3, providing a possible mechanism for regulating the accessibility of the central opening for its binding to a peptide from an interacting factor or to ssDNA. The conservation of the GINS sequence among archaea and eukaryotes suggests a fundamental functional role for this complex in DNA replication; it is likely that GINS serves as a scaffold for the assembly and maintenance of an active helicase and replication complex at the fork. This structure should provide a framework for future studies directed at how the GINS complex interacts with other replication proteins that jointly support both the initiation of DNA synthesis as well as fork progression.

## Materials and Methods

**Cloning and Protein Expression, and Purification GINS.** hSld5 (GenBank accession no. NM_032336), hPsf2 (GenBank accession no. BC062444), and hPsf3 (GenBank accession no. BC005879) cDNAs were amplified from a HeLa cDNA library. hPsf1 (GenBank accession no. BC012542) cDNA was derived from Integrated Molecular Analysis of Genomes and their Expression (IMAGE) clone 4333095. All of these cDNAs were subcloned into pGEX6P-1 (GE Healthcare, Chalfont St. Giles, U.K.) in the order GST-Psf2, Sld5, Psf1, and His-8–Psf3 to produce a polycistronic mRNA. The N-terminal deletion mutants of Psf3 (N-terminal 10 and 18 residues deleted, respectively) were constructed as a C-terminal His-8 fusion (Psf3–His-8). The four subunits of human GINS were coexpressed in *E. coli* cells by isopropyl $\beta$-D-thiogalactoside (Sigma–Aldrich, St. Louis, MO) induction at 18°C. After cells were lysed by passage through a French Press, the GINS complex was purified by nickel-affinity chromatography (Qiagen, Hilden, Germany) and a glutathione resin affinity column (GE Healthcare) in a buffer containing 50 mM Tris·HCl (pH 8.0), 250 mM NaCl (buffer A), and 5 mM 2-mercaptoethanol (Sigma–Aldrich). The GST and His-8 tags were subsequently removed by PreScission protease treatment in buffer A containing 1 mM DTT (Sigma–Aldrich). The GINS complex with 1:1:1:1 molar stoichiometry was obtained by using Resource Q ion-exchange chromatography with a 50–500 mM NaCl gradient elution and gel-filtration chromatography through a Superdex-200 column (GE Healthcare) in buffer containing 50 mM Tris·HCl (pH 8.0) and 50 mM NaCl. The typical yield from 24-liter culture was ≈30 mg.

**Crystallization and Structure Determination.** Crystals were grown by the hanging drop vapor diffusion method with 20 mg/ml GINS complex against a solution containing 60 mM Mes (pH 5.5), 2% (vol/vol) isopropanol, and 34 mM calcium chloride. Multiple anomalous diffraction data from Se-Met crystals were collected using the synchrotron at Argonne National Laboratory (Argonne, IL) from plate crystals $\approx 200 \times 100 \times 20$ $\mu$m in size (SI Table 1). Data were processed with HKL2000. A total of 52 selenium sites were located by the SHELXD program by using multiple anomalous diffraction data between 30 and 3.5 Å resolution range. The SHARP program was used to calculate the experimental phases by using the multiple anomalous diffraction data in the resolution range of 50–2.5 Å. RESOLVE was used for density modification, resulting in a high-quality electron-density map for model building with O refined with CNS to 2.36 Å resolution. The final refinement statistics and geometry as defined by Procheck were in good agreement and are summarized in SI Table 1.

**Note.** As this work was being reviewed, another study reporting the GINS complex structure with truncated Psf1 by Choi *et al.* was published (19).

We thank Dr. R. Zhang at APS 19id in Argonne National Laboratory for assistance with data collection and M. Klein and other X.S.C. laboratory members for their help and discussion. This work was supported in part by a Cellular, Biochemical, and Molecular Sciences training grant (to Y.P.C.) and U.S. Army Medical Research and Materiel Command and National Institutes of Health Grants R01AI055926 (to X.S.C.) and GM 34559 (to J.H.).

1. Forsburg SL (2004) *Microbiol Mol Biol Rev* 68:109–131.
2. Blow JJ, Dutta A (2005) *Nat Rev Mol Cell Biol* 6:476–486.
3. Takahashi TS, Wigley DB, Walter JC (2005) *Trends Biochem Sci* 30:437–444.
4. Labib K, Gambus A (2007) *Trends Cell Biol* 17:271–278.
5. Gambus A, Jones RC, Sanchez-Diaz A, Kanemaki M, van Deursen F, Edmondson RD, Labib K (2006) *Nat Cell Biol* 8:358–366.
6. Moyer SE, Lewis PW, Botchan MR (2006) *Proc Natl Acad Sci USA* 103:10236–10241.
7. Pacek M, Tutter AV, Kubota Y, Takisawa H, Walter JC (2006) *Mol Cell* 21:581–587.
8. Takayama Y, Kamimura Y, Okawa M, Muramatsu S, Sugino A, Araki H (2003) *Genes Dev* 17:1153–1165.
9. Makarova KS, Wolf YI, Mekhedov SL, Mirkin BG, Koonin EV (2005) *Nucleic Acids Res* 33:4626–4638.
10. Marinsek N, Barry ER, Makarova KS, Dionne I, Koonin EV, Bell SD (2006) *EMBO Rep* 7:539–545.
11. Seki T, Akita M, Kamimura Y, Muramatsu S, Araki H, Sugino A (2006) *J Biol Chem* 281:21422–21432.
12. De Falco M, Ferrari E, De Felice M, Rossi M, Hubscher U, Pisani FM (2007) *EMBO Rep* 8:99–103.
13. Kamada K, Kubota Y, Arata T, Shindo Y, Hanaoka F (2007) *Nat Struct Mol Biol* 14:388–396.
14. Gulbis JM, Kelman Z, Hurwitz J, O'Donnell M, Kuriyan J (1996) *Cell* 87:297–306.
15. Kubota Y, Takase Y, Komori Y, Hashimoto Y, Arata T, Kamimura Y, Araki H, Takisawa H (2003) *Genes Dev* 17:1141–1152.
16. Singleton MR, Sawaya MR, Ellenberger T, Wigley DB (2000) *Cell* 101:589–600.
17. Gomez EB, Angeles VT, Forsburg SL (2005) *Genetics* 169:77–89.
18. Yabuuchi H, Yamada Y, Uchida T, Sunathvanichkul T, Nakagawa T, Masukata H (2006) *EMBO J* 25:4663–4674.
19. Choi JM, Lim HS, Kim JJ, Song OK, Cho Y (2007) *Genes Dev* 21:1316–1321.