nature structural & molecular biology

Structure of the heterodimeric core primase

Si-Houy Lao-Sirieix[1], Ravi K Nookala[2], Pietro Roversi[3], Stephen D Bell[1] & Luca Pellegrini[2]

Primases are DNA-dependent RNA polymerases that synthesize the oligoribonucleotide primers essential to DNA replication. In archaeal and eukaryotic organisms, the core primase is a heterodimeric enzyme composed of a small and a large subunit. Here we report a crystallographic and biochemical analysis of the core primase from the archaeon *Sulfolobus solfataricus*. The structure provides the first three-dimensional description of the large subunit and its interaction with the small subunit. The evolutionary conservation of amino acids at the protein-protein interface implies that the observed mode of subunit association is conserved among archaeal and eukaryotic primases. The orientation of the large subunit in the core primase probably excludes its direct involvement in catalysis. Modeling of a DNA-RNA helix together with structure-based site-directed mutagenesis provides insight into the mechanism of template DNA binding and RNA primer synthesis.

Replication of the genome relies on the function of specialized polymerases called primases, which synthesize *de novo* the RNA primers required by the DNA polymerases[1,2]. The activity of DNA primases at the replication fork is coordinated through direct interaction with other replication proteins such as DNA helicases, DNA polymerases and single-stranded DNA-binding proteins. Primases can be broadly grouped into two classes. One contains those from bacteria and bacteriophages and the second encompasses the eukaryotic primases. Distant orthologs of the eukaryotic primases have also been found in several viruses and in some bacteria, as part of a rudimentary end-joining mechanism of DNA repair[3]. The proteins in the two classes differ in structure and in their relationship with the other components of the replication apparatus.

In eukaryotic organisms, a heterotetrameric enzyme formed by DNA polymerase α, the B subunit and two primase subunits has primase activity[4,5] (**Fig. 1a**). Here we refer to the heterodimeric complex of the primase subunits as the core primase. The RNA polymerase activity resides within the smaller subunit of the core primase[6–8]. Disruption of the gene for the large subunit is lethal in yeast cells[9]. However, the molecular details of the involvement of the large subunit in primase function are still obscure. Proposed roles include stabilization of the small subunit, involvement in synthesis initiation, improvement of primase processivity, determination of product size and transfer of the product to DNA polymerase α[6–8,10,11].

The DNA replication machinery of archaeal organisms recapitulates in a simpler form the complexities of eukaryotic replication[12–14], and archaeal orthologs for the two subunits of the eukaryotic core primase have been identified[15–19]. The structural data available for primases of the archaeal and eukaryotic classes are currently limited to the highly similar structures of the small subunit of the *Pyrococcus furiosus* (*Pfu*) and *Pyrococcus horikoshii* (*Pho*) core primases[20,21]. These structures show that the archaeal and eukaryotic primases have a polymerase fold unrelated to that of other known polymerases. However, the conservation and three-dimensional arrangement of the catalytic aspartate residues make it likely that archaeal and eukaryotic primases adopt the general two-metal-ion mechanism of oligonucleotide synthesis[22]. Recently, the structure of a bifunctional primase-polymerase protein encoded by the archaeal plasmid pRN1 from *Sulfolobus islandicus* (*Sis*) has been solved[23]; it shows limited three-dimensional relationship with the small subunit of the *Pfu* and *Pho* core primases.

Here we present a structural and functional study of the core primase from the archaeon *S. solfataricus* (*Sso*). The large subunit folds into a multi-domain protein of novel configuration that binds an evolutionarily conserved region on the surface of the small subunit. Primase mutations that cause a DNA repair defect in yeast cells[24] and disrupt *in vitro* subunit association in the mouse primase[25] map to the protein interface between small and large subunit. Furthermore, the structure of the *Sso* core primase provides insight into the function of the large subunit and represents a useful model for understanding the mechanism of template DNA binding and RNA synthesis.

RESULTS

Structure determination

The *Sso* core primase was prepared by coexpression of its small (PriS) and large (PriL) subunits in bacteria. In *Sso* the large subunit has a slightly lower molecular mass (35.7 kDa) than the small subunit (37.6 kDa). However, here we refer to the subunits by the names of the eukaryotic orthologs to which they bear primary-sequence similarity. Initial attempts to crystallize the intact primase proved fruitless. The structure of the enzyme was therefore probed by limited proteolysis with trypsin and chymotrypsin, and the proteolytic products were characterized by N-terminal sequencing and MALDI techniques

[1]MRC Cancer Cell Unit, Hutchison MRC Research Centre, Hills Road, Cambridge CB2 2XZ, UK. [2]Department of Biochemistry, University of Cambridge, Tennis Court Road, Cambridge CB2 1GA, UK. [3]Laboratory of Molecular Biophysics, University of Oxford, South Parks Road, Oxford, OX1 3QU, UK. Correspondence should be addressed to L.P. (luca@cryst.bioc.cam.ac.uk).

Received 7 July; accepted 23 September; published online 6 November 2005; doi:10.1038/nsmb1013

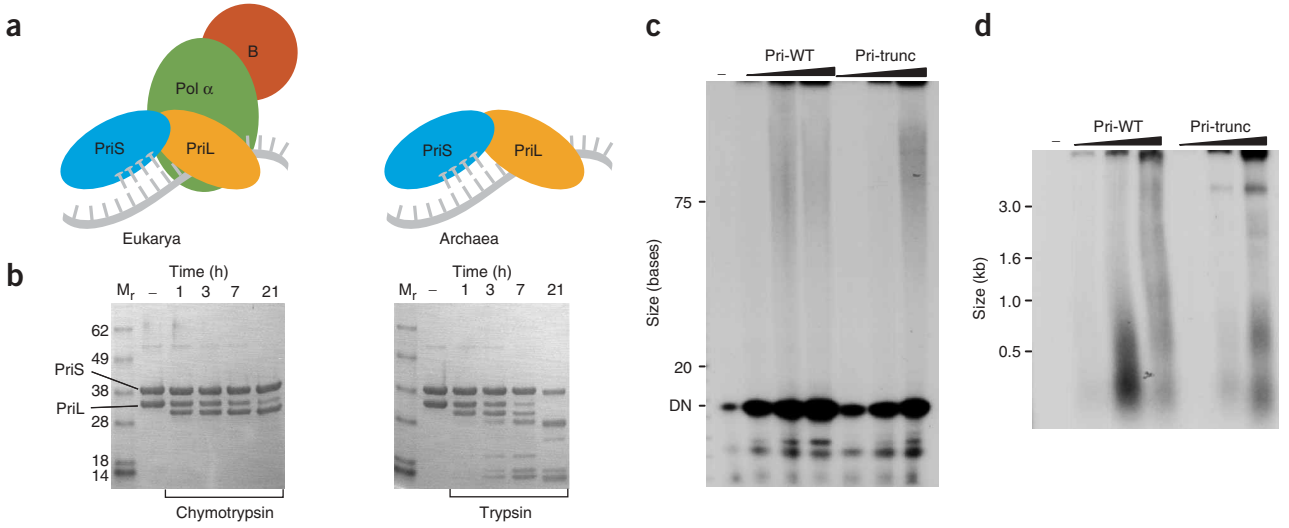

Figure 1 The *Sso* core primase. (**a**) Diagram of the subunit organization of eukaryotic and archaeal primases. Pol α, DNA polymerase α. (**b**) Sensitivity to proteolytic digestion of the *Sso* core primase. The resistance of the PriS–PriL complex to proteolytic cleavage by chymotrypsin and trypsin as a function of time. M_r, relative molecular mass in kDa. (**c**) Comparison of the wild-type (Pri-WT) and proteolytically stable (Pri-trunc) *Sso* core primase activities. Increasing concentrations (0.3, 0.6, 1.2 μM) of enzyme were incubated with 0.1 μM of single-stranded M13 DNA in the presence of nucleoside triphosphates (NTPs) and ^{32}P-labeled ATP at 60 °C. The reaction products were resolved on a denaturing 20% acrylamide gel. DN, dinucleotide product. (**d**) Experiment was performed as in **c**, except that the reaction products were resolved on a formaldehyde 1% agarose gel.

(**Fig. 1b**). The limited proteolysis analysis revealed that a region of ~100 residues at the C terminus of PriL is susceptible to proteolytic degradation. On the basis of the results of the proteolysis experiments, we designed and prepared a recombinant version of the core primase, formed by the full-length PriS (residues 1–330) and a cleavage-resistant version of PriL (residues 1–212). The proteolytically resistant primase was purified in the same way as the intact enzyme. A recombinant version of the proteolytically stable primase retained enzymatic activity, albeit with an altered product profile (**Fig. 1c,d**) and somewhat reduced activity (0.3 versus 0.18 fmol NTP min^{-1} pmol $primase^{-1}$ incorporated by the wild-type and truncated versions of the enzyme, respectively), and it produced crystals suitable for

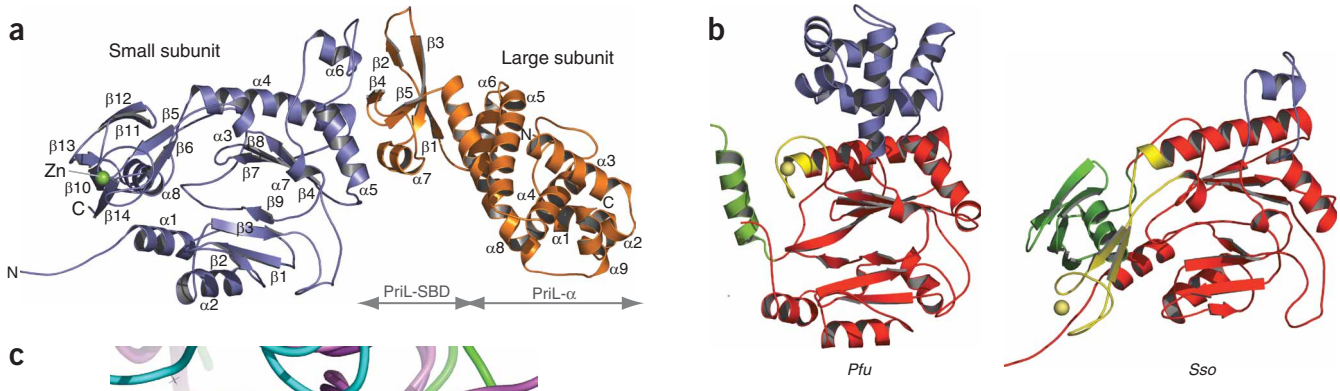

Figure 2 Structure of the *Sso* core primase. (**a**) Crystal structure of the *Sso* core primase. Blue, small subunit; orange, large subunit; green sphere, zinc atom. Secondary structure elements are labeled. (**b**) Structural comparison of the small subunits of *Pfu* (PDB entry 1G71) and *Sso*. Red, PriS-prim; light blue, PriS-coil; yellow, PriS-Zn; green, PriS-CTD. (**c**) Superposition of active site residues for *Sso* (green), *Sis* (cyan), *Pfu* (dark pink) and *Pho* (pink) primases. Tubes, main chain polypeptides; sticks, side chains of the catalytic aspartates; blue, the nucleotide-binding loop of consensus sequence SGXRG. Aspartates that are functionally equivalent to *Sso* Asp103 are grouped by an oval. A UTP nucleotide bound to *Pho* PriS (PDB entry 1V34) is also shown.

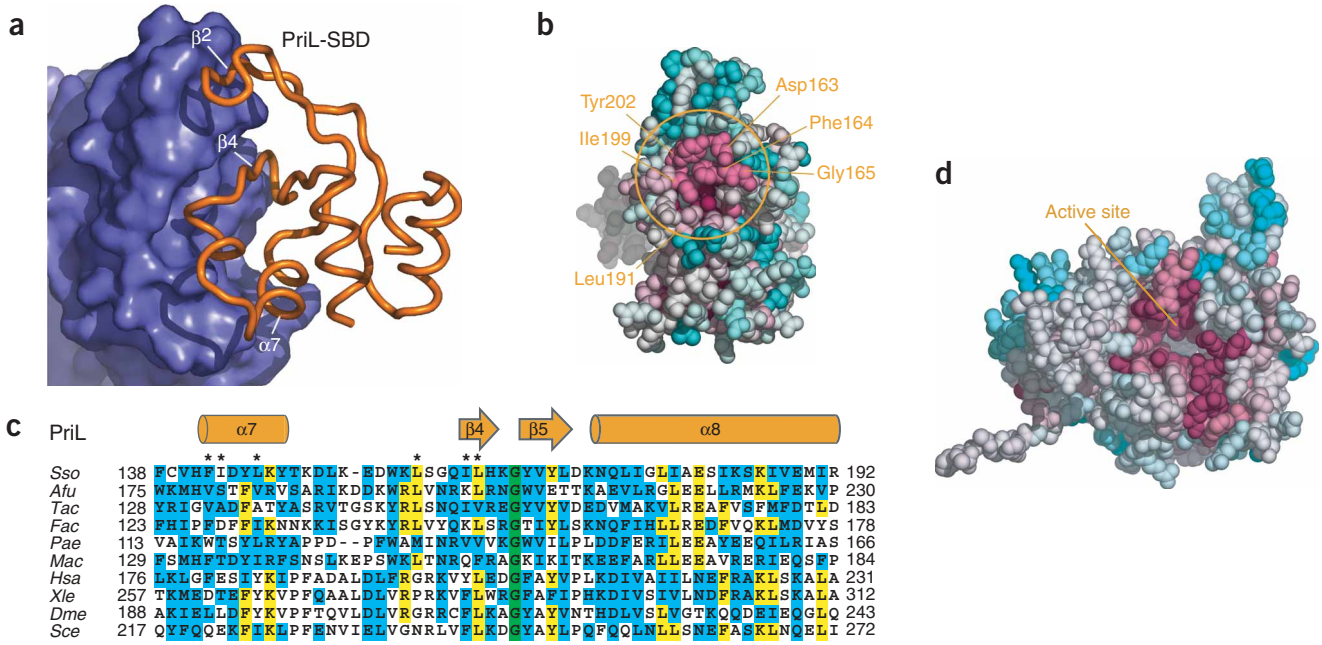

Figure 3 The PriS-PriL interface of the *Sso* core primase. (**a**) Overview of the interface. Blue molecular surface, PriS; orange tube, PriL-SBD. Secondary structure elements of PriL-SBD that are in contact with PriS are labeled. (**b**) Evolutionary conservation of the PriL-binding surface in PriS, indicated by transition in color hues: magenta, most conserved residues; cyan, least conserved. The coloring derives from a structure-based sequence alignment of 24 archaeal and 6 eukaryotic PriS sequences carried out by web server ConSurf (http://www.consurf.tau.ac.il). (**c**) Multiple sequence alignment of the conserved portion of PriL-SBD in archaeal and eukaryotic PriLs, colored according to sequence conservation: green, completely conserved residues; yellow, identical in over half of the sequences; cyan, similar. *Sso*, *S. solfataricus*; *Afu*, *Archeoglobus fulgidus*; *Tac*, *Thermoplasma acidophilum*; *Fac*, *Ferroplasma acidarmanus*; *Pae*, *Pyrobaculum aerophilum*; *Mac*, *Methanosarcina acetivorans*; *Hsa*, *Homo sapiens*; *Xle*, *Xenopus laevis*; *Dme*, *Drosophila melanogaster*; *Sce*, *Saccharomyces cerevisiae*. Asterisks mark conserved hydrophobic PriL residues that interact with PriS. Secondary structure elements in *Sso* PriL are shown above the alignment. (**d**) Sequence conservation at the PriS active site, colored as in **b**.

X-ray analysis. The crystal structure of the *Sso* core primase was determined to a resolution of 3.3 Å by exploiting the anomalous signal of the zinc atom present in the small subunit (**Supplementary Fig. 1** online).

The large subunit of the core primase

Our crystallographic analysis of the *Sso* core primase shows that the large subunit adopts a multi-domain architecture (**Fig. 2a**). An entirely α-helical domain (PriL-α) spanning residues 1–94 and 183–209 constitutes the bulk of the large subunit. Hydrophobic helix α4 is buried within a seven-membered helical bundle, for which fold recognition analysis in Dali[26] (http://www.ebi.ac.uk/dali/) did not show convincing similarity to any fold already present in the Protein Data Bank. The smaller domain responsible for interaction with PriS protrudes from PriL-α helices α5 and α8 (PriL-SBD; residues 95–182), and comprises three helices supporting a four-membered, highly twisted antiparallel β-sheet. A disulfide bridge crosslinks Cys120 and Cys139 near the base of the long hairpin formed by strands β1, β2 and β3, which becomes part of the subunit interface in the complex. PriL-α and PriL-SBD share the long helix α8, but seem otherwise unrestrained relative to each other. The mobility of PriL-α within the core primase might have functional implications for the mechanism of DNA-dependent RNA synthesis. The C-terminal segment of PriL (PriL-CTD; residues 212–307) is proteolytically sensitive and is absent in our current structure. Experimental evidence shows that PriL-CTD has a high affinity for single-stranded DNA, suggesting that it might have a role in template DNA binding[27]. On the basis of sequence conservation, the multi-domain architecture described here for the large subunit of the *Sso* core primase is likely to be a shared feature of archaeal and eukaryotic core primases.

The primase fold and active site of the small subunit

In the *Sso* core primase, PriS adopts the multi-layer, slab-like conformation observed earlier for the isolated subunit[20], with two nearly orthogonal, inner-tier β-sheets hosting the active side and two α-helical outer tiers (**Fig. 2**). A structural comparison between the catalytic subunits of the euryarchaeote *P. furiosus*[20] and the crenarchaeote *S. solfataricus* provides an improved definition of the evolutionarily conserved elements constituting the archaeal primase fold (PriS-prim) (**Fig. 2b**; see **Supplementary Fig. 2** online). The *Sso* PriS-prim fold spans ~220 residues (r.m.s. deviation of 1.6 Å over 181 C_α atoms) and includes the two inner-tier β-sheets and surrounding α-helices.

Such comparison further reveals large differences in the PriS regions lying outside the conserved PriS-prim fold. The α-helical domain observed in the *Pfu* structure has shrunk in the *Sso* structure to a shorter sequence of irregular conformation (PriS-coil) folded around Trp221. A mixed α/β domain is also uniquely present at the C terminus of the *Sso* PriS (PriS-CTD). The PriS-CTD seems to support and position the elongated hairpin structure leading into the zinc-binding motif (PriS-Zn).

As expected, the architecture of the active site of the *Sso* core primase is very similar to that observed in the small subunit of the *Pfu* and *Pho* primases (**Fig. 2c**). In the *Sso* core primase structure, two sulfate ions reside in positions that are equivalent to those occupied by the β- and γ-phosphate groups of the nucleotide in the structure of the

Table 1 Yeast two-hybrid analysis of the PriS-PriL interface

	pGBKT7	pGADT7
1	Empty	Empty
2	PriL	Empty
3	Empty	PriS
4	PriL	PriS
5	Empty	PriS F164E
6	PriL	PriS F164E
7	Empty	PriS F164E I199K
8	PriL	PriS F154E I199K
9	PriL F142E L163E	Empty
10	PriL F142E L163E	PriS
11	Empty	PriS G165I
12	PriL	PriS G165I
13	PriL G166I	Empty
14	PriL G166I	PriS
15	PriL G166I	PriS G165I

Sample numbers should be used to interpret the yeast two-hybrid experiment (**Fig. 4b**).

Pho PriS–UTP complex[21]. However, the position of *Sso* PriS Asp103, a member of the catalytic triad that also includes Asp101 and Asp235, is clearly different from that of the equivalent Asp97 in *Pfu* and *Pho* PriS. Rather, *Sso* Asp103 occupies a position closer to that of Glu113 in the bifunctional *Sis* primase-polymerase. The shift in position might be correlated with the different size and orientation adopted by the adjacent, highly conserved loop linking β-strands β7 and β8. The loop has the consensus sequence SGXRG (where X is any amino acid) and is involved in nucleotide binding[19,21]. In *Sso* and *Sis* primases, X is an asparagine, but it can also be a glycine or a serine, whereas in *Pfu* and *Pho* primases X is absent; in eukaryotic primases X is an arginine. Whether the different position of Asp103 implies a difference in the mechanism of nucleotide processing is at present unknown.

Architecture and subunit interface of the core primase

The *Sso* PriS–PriL complex has a curved, elongated shape reminiscent of an arch or a cashew nut, with overall dimensions of ~100 × 50 × 35 Å^3. This shape is acquired by buttressing the PriL-SBD of PriL against the narrow face of the slab-like PriS structure (**Figs. 2** and **3**). Helix α7, strands β2 and β3 and strands β4 and β5 of PriL-SBD form a three-tier stack of secondary structure elements that fits in a shallow concavity on the PriS surface created by the juxtaposed PriS-prim and PriS-coil domains. The PriS elements interacting with PriL comprise the carboxyl cap of helix α4, the linker between strand β8 and helix α5, helix α5 in PriS-prim and residues within the 207–222 sequence of the conformationally irregular PriS-coil. Despite its evolutionary divergence, the part of PriS-coil contributing to the interaction surface with PriL in the *Sso* core primase has retained a similar three-dimensional conformation in *Pfu* PriS. At ~1,800 Å^2, the total surface area buried at the PriS-PriL interface is within the expected range for protein-protein interactions.

The protein-protein interface in the heterodimeric primase is mainly, but not exclusively, hydrophobic (**Fig. 3b**,**c**). Hydrophilic interactions between PriS and PriL are limited to one important contact between the carboxyl cap of PriS helix α4 and the PriL residues connecting strands β4 and β5 (**Fig. 4a**). The main chain carbonyl moieties of Asp162 and Asp163 in PriS hydrogen bond to the backbone amides of Lys165 and Gly166, respectively. This point of close contact between subunits is made possible by the presence of two conserved glycines, Gly165 in PriS and Gly166 in PriL. Although not directly involved in PriS-PriL recognition, other polar interactions of PriS residues important for subunit association are reported below.

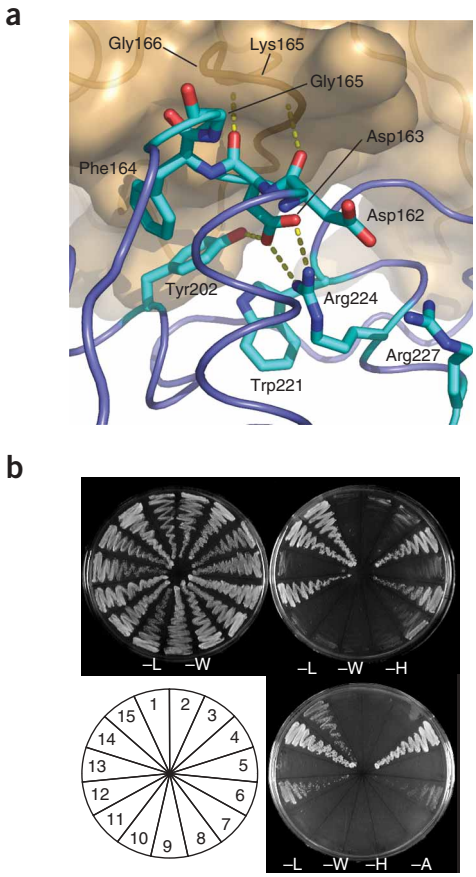

Figure 4 Analysis of the PriS-PriL interface. (**a**) Hydrophilic interactions at the PriS-PriL interface. Hydrogen bonds are yellow dashed lines. The carbonyl groups of Asp162 and Asp163 at the C terminus of PriS helix α4 are hydrogen-bonded to the main chain amides of PriL residues Lys165 and Gly166. The side chains of Asp162 and Asp163 further interact electrostatically with Arg224 and Arg227. (**b**) Structure-based yeast two-hybrid analysis of the *Sso* PriS-PriL interface. Single and double mutations in PriS and PriL disrupt the interaction between the core primase subunits. Control plate: –Leu, –Trp; selective plates: –Leu, –Trp, –His and the more stringent –Leu, –Trp, –His, –Adenine. Numbered samples are identified in **Table 1**.

We probed the stability of the *Sso* PriS-PriL association in the yeast two-hybrid assay, using site-directed mutagenesis that reversed the nature of hydrophobic residues at the subunit interface (**Fig. 4b** and **Table 1**). The single most effective mutation in disrupting the PriS-PriL interaction was PriS F164E, which considerably weakened the subunit association. In the complex, Phe164 becomes buried at the subunit interface and is packed against the side chains of PriL residues Phe142, Leu158 and Leu163. The single mutation PriS G165I, targeting a conserved glycine at a point of close contact between subunits, also achieved a partial destabilization of the complex. Double mutations of F164G I199K in PriS and F142E L163E in PriL abolished the PriS-PriL interaction.

PriS-PriL interface is conserved in eukaryotic primases

To confirm the general relevance to archaeal and eukaryotic primases of the observed mode of subunit association in the *Sso* core primase,

we analyzed functional sites on the surface of the PriS structure with the ConSurf server[28] (http://www.consurf.tau.ac.il) (**Fig. 3b,d**). The analysis correctly recognized the active site of PriS as the most conserved region on the protein surface. It also identified a second conserved region, comprising the sequence Asp163-Phe-Gly165 and residues Leu191, Ile199 and Tyr202, which coincides with the PriS-PriL interface of our crystallographic model. Thus, phylogenetic mapping of the PriS surface strongly supports the conclusion that the interaction and relative orientation of the subunits in the *Sso* core primase are shared by archaeal and eukaryotic core primases.

In the mouse primase, mutation of conserved acidic PriS residues Glu148 and Asp149 to alanine reduced its affinity to PriL[25]. The *Sso* core primase structure reveals that the equivalent *Sso* PriS residues, Asp162 and Asp163 on helix α4, are located at or near to the subunit interface and, together with Glu158, form an electrostatic link between PriS-prim helix α4 and Arg224 and Arg227 of the PriS-coil domain (**Fig. 4a**). Highly conserved Asp163 has a particularly important role, as it forms a salt link with Arg224 of PriS-coil and is hydrogen bonded to conserved Tyr202 on helix α5. Thus, Asp162 and Asp163 contribute to the PriS-PriL interaction by providing a necessary structural support for the architecture of the PriS PriL-binding site.

The structure of the *Sso* core primase offers a rationale for the enhanced sensitivity to genotoxic agents and Rad53-dependent S-phase checkpoint failure of a yeast primase mutant strain[24]. The phenotype was caused by the insertion of a Gly-Ser dipeptide between Asn186 and Val187 of the catalytic subunit of yeast primase. Mapping the insertion on the structure of the *Sso* core primase places the mutation on PriS helix α5, at a point of close contact between subunits in the PriS-PriL interface. Thus, our structural data suggest that the observed phenotype is attributable to destabilization of the PriS-PriL interaction, which is caused by a disruption of the subunit interface. Indeed, yeast two-hybrid analysis of mutant *Sso* PriS bearing a Gly-Ser insertion between the equivalent residues Glu198 and Ile199 shows that the mutation appreciably weakens, but does not abolish, the interaction with PriL (**Fig. 5**). Taken together, these observations argue for a role of the large subunit in the regulation of primase activity after DNA damage. Furthermore, the observed sensitivity of mutant yeast cells to genotoxic agents suggests that small molecules designed to disrupt the PriS-PriL interaction could be of therapeutic value in enhancing the sensitivity of diseased tissue to chemotherapy.

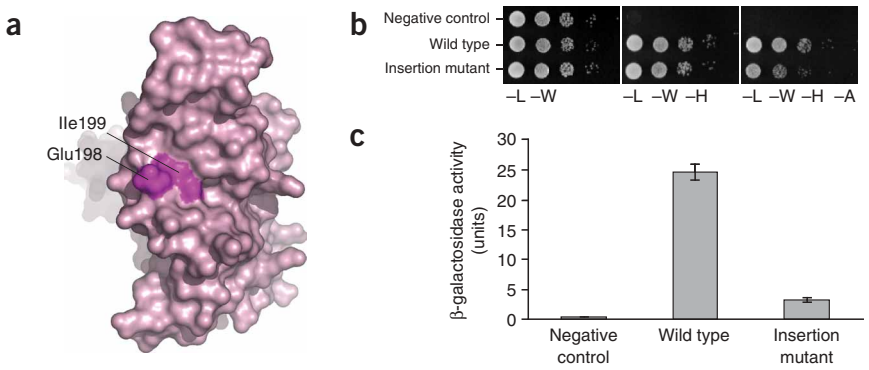

Figure 5 The Gly-Ser dipeptide insertion between PriS residues Glu198 and Ile199 weakens the interaction of the primase subunits. (**a**) *Sso* PriS, with Glu198 and Ile199 indicated in magenta. (**b**) Yeast two-hybrid analysis of PriS insertion mutant (PriS-GS). Double transformants were diluted in a 1 in 10 serial dilution. Negative control: pGBKT7 and pGADT7-PriS-GS. Wild type: pGBKT7-PriL and pGADT7-PriS. Insertion mutant: pGBKT7-PriL and pGADT7-PriS-GS. Control plate: –Leu, –Trp. Selective plates: –Leu, –Trp, –His and the more stringent –Leu, –Trp, –His, –Adenine. (**c**) β-galactosidase activity of the PriS-GS insertion mutant compared with the wild-type primase. β-galactosidase activity was assessed using chlorophenol red β-D-galactopyranoside (CRPG) as a substrate. One unit of β-galactosidase is defined as the amount that hydrolyzes 1 μmol of CRPG per minute per cell.

Role of the zinc-binding motif in primase function

Archaeal and eukaryotic primases contain one zinc ion in their small subunits. Their zinc-binding motifs (PriS-Zn) are not similar in sequence or structure, and even the point of insertion in the polypeptide chain of PriS-prim varies (**Fig. 6a**). In the *Sso* and *Pfu* primases, the PriS-Zn is inserted at the same point of the PriS-prim polypeptide chain. However, whereas the *Pfu* PriS-Zn is directly packed against the bulk of the PriS-prim, the *Sso* PriS-Zn is connected to the PriS-prim via a long stem formed by strands β5 and β6. The stem is crosslinked at its base by a disulfide bridge linking Cys108 and Cys145. Residues Cys116, Cys119, Cys128 and Asp131 coordinate the zinc atom.

The functional implications of the different positions adopted by the PriS-Zn in the *Sso* and *Pfu* primases are at present unclear. We took advantage of the markedly solvent-exposed nature of the *Sso* PriS-Zn to replace residues 116–134, spanning the entire PriS-Zn, with the tripeptide Ser-Asn-Gly, designed to induce the formation of a tight turn between strands β5 and β6 of the stem. To our surprise, the mutant primase lacking PriS-Zn (Pri-SNG) had higher activity compared with the wild-type enzyme, as well as a bias toward shorter products of 30–100 nucleotides (**Fig. 6b,c**). The altered enzymatic activity of Pri-SNG supports a role for the zinc-binding motif in the interaction of the primase with the DNA template. Thus, the diminished size of the Pri-SNG products suggests that removal of the PriS-Zn might weaken the grip of the enzyme on the template strand, decreasing its processivity. The conserved zinc-binding domain of prokaryotic and bacteriophage primases has also been implicated in single-stranded DNA binding[29,30].

Implications for template DNA binding and RNA synthesis

The structure of the *Sso* core primase shows that the large subunit is removed from the active site, which remains fully accessible. A direct involvement of the large subunit in the catalytic steps necessary for RNA synthesis (nucleotide binding, dinucleotide formation and primer extension) is therefore unlikely. To gain insight into core primase function, we analyzed the possible mode of interaction of the core primase with a DNA template and an RNA primer. An extended DNA-RNA helix was created in X3DNA[31], based on the helical parameters of a decameric hybrid DNA-RNA helix[32], and manually positioned on the primase (**Fig. 6d**). The docking process was aided by knowledge of the active site location as deduced in the structural analysis of the *Pho* PriS–UTP complex[21]. Docking was further helped by knowledge of the structure of the DNA polymerase β-DNA–ddCTP complex[33].

In the resulting model, the DNA template (lagging strand) traverses the length of the core primase, so that during polymerization PriS-Zn would be near the 5′ end of the template strand, whereas PriL-CTD would be near the 3′ end of the template strand, in the direction of the

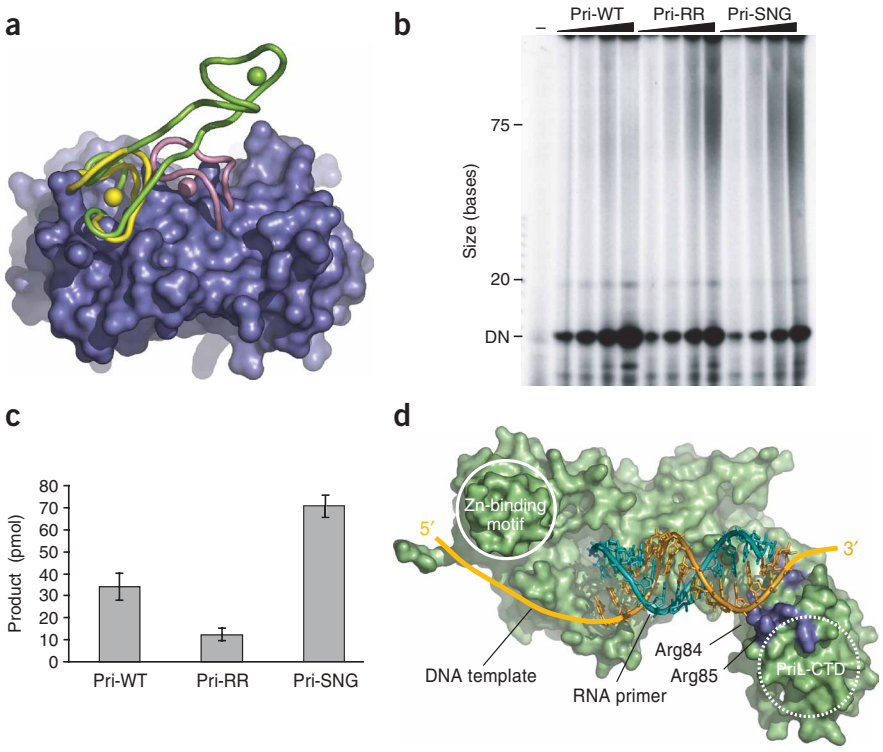

Figure 6 Interaction of the *Sso* core primase with DNA template and RNA primer. (**a**) The zinc-binding motif of archaeal primases. The small subunits of *Sso*, *Pho* (PDB entry 1V33) and *Sis* (PDB entry 1RNI) primases were superimposed. The *Sso* PriS-prim is shown as a molecular surface. The zinc-binding motifs of *Sso* (green), *Pho* (yellow) and *Sis* (pink) are narrow tubes and their zinc atoms are spheres. (**b**) Comparison of the enzymatic activity of wild-type *Sso* core primase (Pri-WT) with that of the RR and SNG mutants (Pri-RR and Pri-SNG, respectively). Experiments were performed as for **Figure 1c**. DN, dinucleotide product. (**c**) Comparison of the quantities of product synthesized by Pri-WT and the RR and SNG mutants. Experiments were performed as for **Figure 1c**, with 1.2 μM primase concentration. The radiolabeled products were quantified by filter binding and liquid scintillation counting. Each bar represents the average of five independent values with s.e.m. indicated. (**d**) Model of the *Sso* core primase–DNA template–RNA primer complex. The protein component of the complex is depicted as a molecular surface. The phosphate backbones of DNA and RNA are orange and cyan tubes, respectively. The proposed trajectory of the template DNA across the surface of the core primase is drawn. The position of PriS-Zn and the putative position of PriL-CTD are indicated by solid and dashed circles, respectively. The side chains of basic residues on and near PriL α5 are blue. The positions of PriL residues Arg84 and Arg85 are indicated.

replication fork. The proposed trajectory of the template DNA strand is in agreement with our biochemical data for the Pri-SNG mutant, which support a role for the PriS-Zn in the primase interaction with the DNA template. It is further supported by the experimental observation that the PriL-CTD has a high affinity for single-stranded DNA[27].

In the model, the DNA-RNA helix extends toward the large subunit. The putative area of contact contains a cluster of basic residues centered on PriL helix α5. Inspection of the putative DNA-RNA helix–PriL interface shows that two solvent-exposed arginines, Arg84 and Arg85, are well-positioned to interact with the phosphate backbone of the DNA-RNA duplex. Indeed, a double PriL R84A R85A primase mutant (Pri-RR) showed a marked reduction in the size and amount of RNA product synthesized (**Fig. 6b,c**).

The model further suggests that interactions between the large subunit and the RNA primer become possible once the product reaches a size within the observed physiological range (7–14 bases). Additional contacts might cause the primase–DNA template–RNA primer complex to pause and allow hand-off of the primer to a DNA polymerase, as proposed earlier[34]. Thus, the model is consistent with the experimental evidence that the large subunit is implicated in the negative regulation of primer synthesis[10]. The paradoxical ability of archaeal primases to synthesize a wide size range of DNA and RNA products *in vitro*[17,19] might reflect an absence of archaeal orthologs of DNA polymerase α and may suggest a dual role in RNA primer synthesis and elongation. Alternatively, this apparent discrepancy might be explained by the absence *in vitro* of additional replication factors contributing to the regulatory mechanism limiting the size of the growing RNA chain.

DISCUSSION

Here we have presented a structural and functional investigation of a proteolytically resistant version of the core primase from the archaeal organism *S. solfataricus*. The relevance to archaeal and eukaryotic primases of the observed mode of subunit association is supported by the evolutionary conservation of interface residues. In addition, the structure provides a rationale for mutations that cause a DNA damage phenotype in yeast and a decrease in the affinity between the subunits of the mouse primase.

Furthermore, our findings provide insight into the role of the large subunit in core primase function. The orientation of the large subunit relative to the active site in the small subunit argues against a direct contribution of the large subunit to the multi-step reaction of catalysis. Through a combination of DNA-RNA modeling and structure-based mutagenesis, we provide evidence in support of a two-fold role of the large subunit. By providing additional points of contact with the DNA template, the large subunit would increase the stability of the primase–DNA complex and help keep the primase in the correct orientation for synthesis initiation. Owing to its position within the heterodimeric primase relative to the nascent RNA primer, the large subunit would also be poised to take part in the mechanism of RNA counting and further primer processing (that is, transfer to DNA polymerase α for elongation).

Our findings represent an advance toward a complete description at the atomic level of the protein-protein interactions that govern the DNA replication apparatus and provide a framework for further study of the core primase function in DNA replication. Future research will address the mechanism of RNA synthesis and the coordination of core primase function with the other components of the DNA replication apparatus.

METHODS

Purification and crystallization of the *Sso* core primase. The intact and proteolytically resistant versions of the *Sso* core primase were bacterially expressed and purified as described in ref. 19. Crystallization was induced

Table 2 X-ray diffraction data and refinement statistics

	Native 1			Native 2
Data collection				
Space group	$P4_22_12$			
Cell dimensions				
a, *b*, *c* (Å)	193.3, 193.3, 213.0			
	Peak	*Inflection*	*Remote*	*Peak*
Wavelength (Å)	1.2824	1.2829	1.2769	1.2824
Resolution (Å)	3.56	3.56	3.57	3.33
R_{sym} (%)	14.1 (91.0)	14.5 (99.9)	14.9 (94.7)	13.3[a]
Redundancy	6.8 (6.3)	5.3 (5.0)	7.0 (6.6)	12.0 (12.2)
$I / \sigma I$	12.6 (1.9)	11.3 (1.5)	13.1 (2.0)	17.7 (2.1)
χ^2	0.97 (0.91)	1.21 (1.08)	1.06 (1.01)	1.08 (1.04)
Completeness (%)	99.1 (91.0)	98.6 (85.5)	98.7 (82.0)	99.6 (93.1)
Refinement				
Resolution (Å)				40–3.33
No. reflections				59,248
R_{work} / R_{free}				23.6 / 24.7
No. atoms				
Protein				8,799
Zn				2
Sulfate				39
Water				12
B-factor (Å^2)				89.9
R.m.s. deviations				
Bond lengths (Å)				0.005
Bond angles (°)				0.845

Values in parentheses are for the highest-resolution shell. χ^2 represents the goodness of fit between the data and the error model (as defined in SCALEPACK[36]).
[a]The R_{sym} value for the highest-resolution shell exceeded 100%.

by hanging drop vapor diffusion at 18 °C, by mixing equal volumes of concentrated (170 μM) protein solution and 2.0 M $(NH_4)_2SO_4$, 0.1 M Tris-HCl (pH 8.5) buffer.

Phasing and refinement. The *Sso* core primase crystallized with two heterodimers in the asymmetric unit (**Table 2**). Crystals diffracted to ~3.0 Å, with considerable fall-off beyond 3.5 Å. The structure was determined using the anomalous signal of the zinc atom present in the small subunit. MAD data were collected at beamline ID29 of the European Synchrotron Research Facility (Grenoble, France). The positions of the two zinc atoms in the asymmetric unit were determined with direct methods as implemented in Shake-and-Bake[35]. An initial set of experimental phases calculated in SHARP[36] produced an electron density map with detectable protein-solvent boundaries but no interpretable secondary structure. The solvent modification scripts available within the SHARP[36] user interface markedly improved the electron density, allowing the unambiguous tracing of the small and large polypeptide chains in COOT[37]. The crystallographic model was initially refined using REFMAC[38] and considerably improved by BUSTER-TNT[39]. The refined model comprises 1,062 amino acids, 8,799 nonhydrogen atoms, 2 zinc ions, 39 sulfate ions and 12 water molecules. Small-subunit residues 1, 2 and 330 of chain A and 1–10 of chain C, and large-subunit residues 1–2 and 210–212 of chain B and 1, 127–132 and 210–212 of chain D are disordered or poorly ordered in the density map and are not included in the model. Of the protein residues, 98.4% were in the favored or allowed regions of the Ramachandran plot and 1.6% in the outlier regions. Figures were prepared with PyMOL (http://pymol.sourceforge.net).

Primase assay and yeast two-hybrid analysis. The enzymatic activities of the various *Sso* core primase constructs (Pri-WT, Pri-trunc, Pri-RR and Pri-SNG) were assayed as described in ref. 19. The yeast two-hybrid analysis of the PriS-PriL interaction was done as described in ref. 19. Expression levels of the PriS and PriL fusion proteins in yeast were verified by western blotting with anti-cMyc and anti-HA antisera (data not shown) according to the manufacturer's instructions (Clontech).

Accession codes. Protein Data Bank: Coordinates have been deposited with accession code 1ZT2.

Note: Supplementary information is available on the Nature Structural & Molecular Biology website.

ACKNOWLEDGMENTS
This research was supported by a Wellcome Trust senior research fellowship award to L.P. and by the Medical Research Council in the laboratory of S.D.B. We thank X.-J. Lu for help with X3DNA.

COMPETING INTERESTS STATEMENT
The authors declare that they have no competing financial interests.

Published online at http://www.nature.com/nsmb/
Reprints and permissions information is available online at http://npg.nature.com/reprintsandpermissions/

1. Kornberg, A. & Baker, T.A. *DNA Replication* 2nd edn (W.H. Freeman and Company, New York, 1992).
2. Frick, D.N. & Richardson, C.C. DNA primases. *Annu. Rev. Biochem.* **70**, 39–80 (2001).
3. Iyer, L.M., Koonin, E.V., Leipe, D.D. & Aravind, L. Origin and evolution of the archaeo-eukaryotic primase superfamily and related palm-domain proteins: structural insights and new members. *Nucleic Acids Res.* **33**, 3875–3896 (2005).
4. Foiani, M., Lucchini, G. & Plevani, P. The DNA polymerase α-primase complex couples DNA replication, cell-cycle progression and DNA-damage response. *Trends Biochem. Sci.* **22**, 424–427 (1997).
5. Arezi, B. & Kuchta, R.D. Eukaryotic DNA primase. *Trends Biochem. Sci.* **25**, 572–576 (2000).
6. Santocanale, C., Foiani, M., Lucchini, G. & Plevani, P. The isolated 48,000-dalton subunit of yeast DNA primase is sufficient for RNA primer synthesis. *J. Biol. Chem.* **268**, 1343–1348 (1993).
7. Copeland, W.C. & Wang, T.S. Enzymatic characterization of the individual mammalian primase subunits reveals a biphasic mechanism for initiation of DNA replication. *J. Biol. Chem.* **268**, 26179–26189 (1993).
8. Schneider, A. *et al.* Primase activity of human DNA polymerase α-primase. Divalent cations stabilize the enzyme activity of the p48 subunit. *J. Biol. Chem.* **273**, 21608–21615 (1998).
9. Foiani, M., Santocanale, C., Plevani, P. & Lucchini, G. A single essential gene, PRI2, encodes the large subunit of DNA primase in *Saccharomyces cerevisiae*. *Mol. Cell. Biol.* **9**, 3081–3087 (1989).
10. Arezi, B., Kirk, B.W., Copeland, W.C. & Kuchta, R.D. Interactions of DNA with human DNA primase monitored with photoactivatable crosslinking agents: implications for the role of the p58 subunit. *Biochemistry* **38**, 12899–12907 (1999).
11. Zerbe, L.K. & Kuchta, R.D. The p58 subunit of human DNA primase is important for primer initiation, elongation, and counting. *Biochemistry* **41**, 4891–4900 (2002).
12. Edgell, D.R. & Doolittle, W.F. Archaea and the origin(s) of DNA replication proteins. *Cell* **89**, 995–998 (1997).
13. Dionne, I. *et al.* DNA replication in the hyperthermophilic archaeon *Sulfolobus solfataricus*. *Biochem. Soc. Trans.* **31**, 674–676 (2003).
14. Grabowski, B. & Kelman, Z. Archeal DNA replication: eukaryal proteins in a bacterial context. *Annu. Rev. Microbiol.* **57**, 487–516 (2003).
15. Desogus, G., Onesti, S., Brick, P., Rossi, M. & Pisani, F.M. Identification and characterization of a DNA primase from the hyperthermophilic archaeon *Methanococcus jannaschii*. *Nucleic Acids Res.* **27**, 4444–4450 (1999).
16. Makarova, K.S. *et al.* Comparative genomics of the archaea (euryarchaeota): evolution of conserved protein families, the stable core, and the variable shell. *Genome Res.* **9**, 608–628 (1999).
17. Bocquier, A.A. *et al.* Archaeal primase: bridging the gap between RNA and DNA polymerases. *Curr. Biol.* **11**, 452–456 (2001).
18. Liu, L. *et al.* The archaeal DNA primase: biochemical characterization of the p41-p46 complex from *Pyrococcus furiosus*. *J. Biol. Chem.* **276**, 45484–45490 (2001).
19. Lao-Sirieix, S.H. & Bell, S.D. The heterodimeric primase of the hyperthermophilic archaeon *Sulfolobus solfataricus* possesses DNA and RNA primase, polymerase and 3′-terminal nucleotidyl transferase activities. *J. Mol. Biol.* **344**, 1251–1263 (2004).
20. Augustin, M.A., Huber, R. & Kaiser, J.T. Crystal structure of a DNA-dependent RNA polymerase (DNA primase). *Nat. Struct. Biol.* **8**, 57–61 (2001).
21. Ito, N., Nureki, O., Shirouzu, M., Yokoyama, S. & Hanaoka, F. Crystal structure of the *Pyrococcus horikoshii* DNA primase-UTP complex: implications for the mechanism of primer synthesis. *Genes Cells* **8**, 913–923 (2003).
22. Steitz, T.A., Smerdon, S.J., Jager, J. & Joyce, C.M. A unified polymerase mechanism for nonhomologous DNA and RNA polymerases. *Science* **266**, 2022–2025 (1994).
23. Lipps, G., Weinzierl, A.O., von Scheven, G., Buchen, C. & Cramer, P. Structure of a bifunctional DNA primase-polymerase. *Nat. Struct. Mol. Biol.* **11**, 157–162 (2004).

24. Marini, F. *et al.* A role for DNA primase in coupling DNA replication to DNA damage response. *EMBO J.* **16**, 639–650 (1997).
25. Copeland, W.C. & Tan, X. Active site mapping of the catalytic mouse primase subunit by alanine scanning mutagenesis. *J. Biol. Chem.* **270**, 3905–3913 (1995).
26. Holm, L. & Sander, C. Dali: a network tool for protein structure comparison. *Trends Biochem. Sci.* **20**, 478–480 (1995).
27. Matsui, E. *et al.* Distinct domain functions regulating *de novo* DNA synthesis of thermostable DNA primase from hyperthermophile *Pyrococcus horikoshii*. *Biochemistry* **42**, 14968–14976 (2003).
28. Glaser, F. *et al.* ConSurf: identification of functional regions in proteins by surface-mapping of phylogenetic information. *Bioinformatics* **19**, 163–164 (2003).
29. Pan, H. & Wigley, D.B. Structure of the zinc-binding domain of *Bacillus stearothermophilus* DNA primase. *Structure Fold. Des.* **8**, 231–239 (2000).
30. Kato, M., Ito, T., Wagner, G. & Ellenberger, T. A molecular handoff between bacteriophage T7 DNA primase and T7 DNA polymerase initiates DNA synthesis. *J. Biol. Chem.* **279**, 30554–30562 (2004).
31. Lu, X.J. & Olson, W.K. 3DNA: a software package for the analysis, rebuilding and visualization of three-dimensional nucleic acid structures. *Nucleic Acids Res.* **31**, 5108–5121 (2003).
32. Horton, N.C. & Finzel, B.C. The structure of an RNA/DNA hybrid: a substrate of the ribonuclease activity of HIV-1 reverse transcriptase. *J. Mol. Biol.* **264**, 521–533 (1996).
33. Pelletier, H., Sawaya, M.R., Kumar, A., Wilson, S.H. & Kraut, J. Structures of ternary complexes of rat DNA polymerase β, a DNA template-primer, and ddCTP. *Science* **264**, 1891–1903 (1994).
34. Sheaff, R.J. & Kuchta, R.D. Mechanism of calf thymus DNA primase: slow initiation, rapid polymerization, and intelligent termination. *Biochemistry* **32**, 3027–3037 (1993).
35. Weeks, C.M. & Miller, R. The design and implementation of SnB version 2.0. *J. Appl. Crystallogr.* **32**, 120–124 (1999).
36. La Fortelle, E.D. & Bricogne, G. Maximum-likelihood heavy-atom parameter refinement for multiple isomorphous replacement and multiwavelength anomalous diffraction methods. *Methods Enzymol.* **276**, 472–494 (1997).
37. Emsley, P. & Cowtan, K. Coot: Model-building tools for molecular graphics. *Acta Crystallogr. D* **60**, 2126–2132 (2004).
38. Collaborative Computational Project, Number 4. The CCP4 suite: programs for protein crystallography. *Acta Crystallogr. D* **50**, 760–763 (1994).
39. Blanc, E., Roversi, P., Vonrhein, C., Flensburg, S.M. & Bricogne, G. Refinement of severely incomplete structures with maximum likelihood in BUSTER-TNT. *Acta Crystallogr. D* **60**, 2210–2221 (2004).