

The EM structure of human DNA polymerase γ reveals a localized contact between the catalytic and accessory subunits

Elena Yakubovskaya¹, Mark Lukin¹, Zhixin Chen¹, John Berriman^{2,6}, Joseph S Wall³, Ryuji Kobayashi⁴, Caroline Kisker⁵ and Daniel F Bogenhagen^{1,*}

¹Department of Pharmacological Sciences, State University of New York at Stony Brook, Stony Brook, NY, USA, ²New York Structural Biology Center, New York, NY, USA, ³Department of Biology, Brookhaven National Laboratory, Upton, NY, USA, ⁴Department Molecular Pathology, University of Texas MD Anderson Cancer Center, Houston, TX, USA and ⁵Rudolf Virchow Center for Experimental Biomedicine, Institute for Structural Biology, University of Würzburg, Würzburg, Germany

We used electron microscopy to examine the structure of human DNA pol γ , the heterotrimeric mtDNA replicase implicated in certain mitochondrial diseases and aging models. Separate analysis of negatively stained preparations of the catalytic subunit, pol γ A, and of the holoenzyme including a dimeric accessory factor, pol γ B₂, permitted unambiguous identification of the position of the accessory factor within the holoenzyme. The model explains protection of a partial chymotryptic cleavage site after residue L₅₄₉ of pol γ A upon binding of the accessory subunit. This interaction region is near residue 467 of pol γ A, where a disease-related mutation has been reported to impair binding of the B subunit. One pol γ B subunit dominates contacts with the catalytic subunit, while the second B subunit is largely exposed to solvent. A model for pol γ is discussed that considers the effects of known mutations in the accessory subunit and the interaction of the enzyme with DNA.

The EMBO Journal (2007) 26, 4283–4291. doi:10.1038/sj.emboj.7601843; Published online 30 August 2007

Subject Categories: structural biology

Keywords: DNA polymerase γ ; electron microscopy; mitochondrial DNA; processivity; replication

Introduction

DNA polymerase γ (pol γ) serves as both the replicative and repair polymerase in mitochondria (Kaguni, 2004; Graziewicz *et al*, 2006). This enzyme has been implicated in generation of mutations in mtDNA, which show a signature of transition and transversion mutations similar to those produced by pol

γ *in vitro* (Zheng *et al*, 2006). mtDNA mutations cause a number of metabolic diseases and contribute to the development of degenerative disorders as well as to aging. In recent years, evidence has accumulated that the gene encoding the catalytic subunit of pol γ is an important disease locus, as individuals bearing mutant alleles of the pol γ A gene accumulate mtDNA mutations at an accelerated rate, giving rise to disorders such as progressive external ophthalmoplegia and Alper's syndrome (Longley *et al*, 2005). An extreme example of the effect of defective pol γ A is provided by recent work showing that mice bearing a pol γ lacking the editing 3'–5' exonuclease have a shortened lifespan and many hallmarks of premature aging (Trifunovic *et al*, 2004; Kujoth *et al*, 2005).

Pol γ in vertebrates is a heterotrimer with one copy of the larger catalytic subunit, pol γ A, and a dimer of an accessory subunit, pol γ B (Yakubovskaya *et al*, 2006). Vertebrate pol γ enzymes are so highly conserved in size and sequence that both the mouse and frog pol γ B proteins have been shown to stimulate replication by the 139-kDa human catalytic subunit (Carrodeguas *et al*, 1999). The X-ray crystal structure of the murine accessory subunit revealed that it is a dimer of 54-kDa subunits that adopts a structure closely resembling those of bacterial tRNA synthetases (Carrodeguas *et al*, 2001). As expected from the high degree of sequence conservation, the human pol γ B has also been shown to resemble tRNA synthetases (Fan *et al*, 2006).

The pol γ A catalytic subunit shares homology with the other members of the family A of DNA polymerases, but is considerably larger than the related T7 DNA pol or the Klenow fragment of *Escherichia coli* DNA pol I (Ito and Braithwaite, 1990; Ye *et al*, 1996; Graziewicz *et al*, 2004). Much of this size difference is accounted for by an expanded spacer region separating the N-terminal 3'–5' exonuclease from the C-terminal polymerase domain. To obtain structural information on the catalytic subunit and its interaction with the accessory subunit, we studied the structure of pol γ using EM image reconstruction.

Under ideal conditions, EM is capable of providing very detailed information on protein structure, with a resolution approaching 12 Å (for stained specimens) and 4 Å (for cryo-preserved specimens), especially for complexes showing a high degree of symmetry (Henderson, 2004). The higher-resolution cryo images are unstained and of much lower contrast, requiring analysis of many thousands of images. As a small asymmetrical object, pol γ does not provide an optimal candidate for EM analysis. Nevertheless, we have been able to obtain structures of approximately 17-Å resolution for both the catalytic subunit and the holoenzyme using negatively stained specimens. These structures have permitted us to identify the precise location of the pol γ B dimer within the holoenzyme, and have provided the first detailed understanding of the structural relationship between the pol γ subunits.

*Corresponding author. Department of Pharmacological Sciences, State University of New York at Stony Brook, bst8-140, Stony Brook, NY 11794-8651, USA. Tel.: +1 631 444 3068; Fax: +1 631 444 3218; E-mail: dan@pharm.sunysb.edu

⁶Present address: Division of Physical Biochemistry, The National Institute for Medical Research, The Ridgeway, Mill Hill, London NW7, UK

Received: 25 April 2007; accepted: 8 August 2007; published online: 30 August 2007

Results

Structure of the pol γ A catalytic subunit and its position in the pol γ holoenzyme

We decided to study the EM structures of the catalytic subunit and the holoenzyme in parallel to permit confident assignment of the position of the known pol γ B structure within the holoenzyme. We analyzed 8764 images of pol γ A preserved in uranyl acetate stain and reconstructed its 3D structure using reference-free alignment and common-line Euler angle assignment (Supplementary Figures S1 and S4), followed by iterative projection matching refinement of the initial model as implemented in the EMAN software package. Projections of the final structure match well with corresponding raw images (Figure 1B) and the distribution of raw particles between classes demonstrates that most particle orientations are well represented in the raw image set (Supplementary Figure S1C). The structure obtained following eight cycles of refinement (Figure 1C) has a globular C-shape that differs significantly from other family A polymerases such as T7 DNA polymerase and the Klenow fragment of pol I.

Structure of the pol γ AB₂ holoenzyme

The parallel analysis of negatively stained pol γ holoenzyme provided a clear contrast to the structure of the A subunit alone. In this case, analysis of raw images showed that the holoenzyme adopted a two-lobed shape that is consistent with a stoichiometry of one 108-kDa B₂ dimer and one 139-kDa catalytic A subunit (Figure 2).

The pol γ holoenzyme was imaged in the presence of a 24/46-mer DNA primer-template, because grids prepared in the absence of DNA showed a substantial amount of disordered pol γ . To assure that the pol γ -DNA complex remains intact during preparation of EM specimens, and to estimate the fraction of pol γ bound to DNA, separate grids were prepared from a sample containing pol γ , the template-primer

labeled with a 5'-biotin residue on the primer strand and avidin (4:4:1 molar ratio, to take advantage of the four-fold symmetry of the avidin-biotin interaction). Imaging of these uranyl acetate-stained grids demonstrated that >90% of pol γ molecules are contained in compact star-like structures composed of two to four pol γ molecules (data not shown). This confirmed that the enzyme remained bound to DNA during specimen preparation.

Iterative classification and alignment of the untilted image set initially produced 39 classes, 31 of which correspond to molecules adsorbed on the carbon film with their major axis oriented more or less parallel to the grid plane (Supplementary Figure S2a). These classes represent pol γ molecules differently rotated around the major axis and form a continuous set of orientations. To obtain additional views of the protein, a set of raw image pairs was collected at 0° and 50° tilt angles (see Supplementary Figure S3). After multi-reference alignment of untilted images, we selected 3195 dumb-bell-shaped molecules for analysis with their corresponding tilted counterparts. This combined set of 6390 untilted/tilted images was used for subsequent model building and 3D reconstruction, along with 2262 untilted images from our initial image set during the final iteration steps (a total of 8652 images). Iterative projection matching refinement against this set produced a bilobed structure at ~17-Å resolution (Figure 2). Throughout the reconstruction procedure, we were concerned with the possibility that the pseudosymmetry inherent in the two-lobed dumb-bell shape of the holoenzyme might lead to the final structure having artificially exaggerated symmetry due to ambivalent particle alignment. However, monitoring the evolution of particle orientation during refinement (see Materials and methods) demonstrated that the contribution of ambivalent alignment decreased sharply during the first two to three iterations and became negligible during further iterations. Additional characterization of the reconstruction is shown in Supplementary Figure S4.

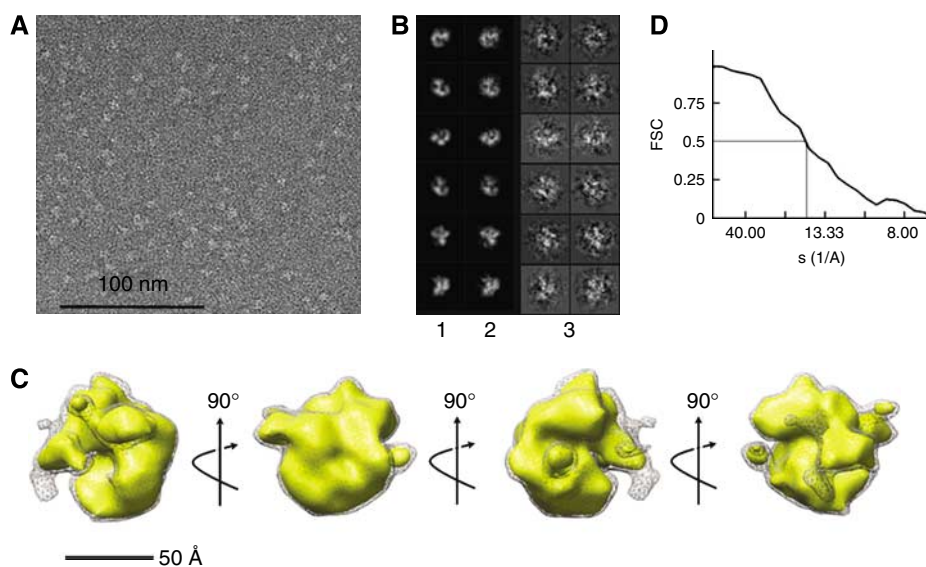


Figure 1 Structure of the pol γ catalytic subunit. **(A)** A field of an EM image of pol γ A stained with uranyl acetate is shown. **(B)** Representative projections (1), corresponding class averages (2) and CTF-corrected raw images (3) used for reconstruction. The box size is 167 Å. **(C)** The final structure of the pol γ catalytic subunit filtered to 15 Å is shown, with each image rotated by 90° with respect to its neighbors, with the axis and direction of rotation shown by arrows. The gray mesh encloses a volume corresponding to the entire 139 kDa, the gold shape corresponds to ~75% of the full mass (105 kDa). **(D)** The Fourier shell correlation curve indicates a resolution of ~17 Å.

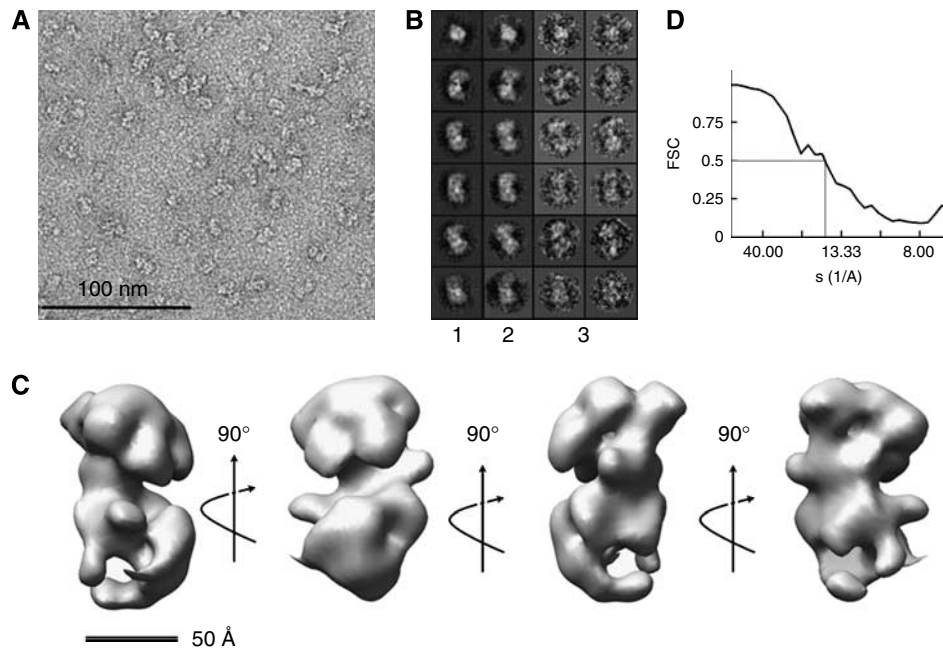


Figure 2 Structure of the pol γ holoenzyme complex. **(A)** An area of an image of template–primer-bound pol γ stained with uranyl acetate is shown using a 3- μm defocus raw micrograph to provide improved clarity. **(B)** Representative projections (1), corresponding class averages (2) and CTF-corrected raw images (3) used for reconstruction. The box size is 223 Å. **(C)** The final structure of the pol γ holoenzyme is illustrated by a series of images, each rotated by 90° with respect to its neighbors, with the axis and direction of rotation shown by arrows. **(D)** The Fourier shell correlation curve indicates a resolution of ~ 17 Å.

Assignment of subunit positions in pol γ models

Simple visual inspection of the dumb-bell-shaped holoenzyme structure suggested that the upper lobe appears to permit a good fit to the crystal structure of the human or mouse accessory subunit dimer (Carrodeguas *et al*, 2001; Fan *et al*, 2006). The lower part adopts a shape resembling the catalytic subunit shown in Figure 1. These subunit structures are shown docked into the holoenzyme in Figure 3. Both manual and computer-aided (SITUS software; <http://situs.biomachina.org/ftut.html>) docking of the pol γB_2 resulted in the same unique orientation. The pol γA EM structure is rather globular (Figure 1), but its size and surface features permitted this structure to be docked in the lower lobe of the holoenzyme. Remarkably, the orientation of pol γB_2 within the holoenzyme indicates that its predominant contacts with the catalytic subunit appear to be largely restricted to one of two B subunits (Figure 3). This is consistent with a number of previous observations, as discussed below.

Probing the contacts between pol γ subunits by limited proteolysis

As an independent method to explore the binding relationship between the pol γA and B_2 subunits, we used limited proteolysis to identify protease cleavage sites that might be protected in the holoenzyme structure. We found that chymotrypsin was able to cleave pol γA at three highly accessible sites (Figure 4), which might be considered to divide the enzyme into three domains. The fact that only the A subunit contained a His₆ tag, in this case at the N-terminus, enabled us to identify the three major N-terminal fragments produced by chymotrypsin, polypeptides, labeled b, d and f in Figure 4B. Three major fragments lacking a His₆ tag were found to react with an antibody directed against a conserved sequence within the pol γ polymerase domain,

KVFNYGRIYGAGC, representing residues 947–959 in human pol γ (data not shown). These C-terminal fragments are labeled a, c and e in the Coomassie blue-stained image in Figure 4A. Edman degradation was used to obtain N-terminal sequences of these C-terminal fragments. Fragment a was found to have the sequence IAAKQGKHKQPPT; fragments c and e started with QKLKGTTELL and TARGGPX(D)TQ, respectively. Therefore, this analysis revealed limited chymotrypsin cleavage of pol γA after residues W₃₁₂, L₅₄₉ and L₇₁₉. It is apparent in Figure 4 that the L₅₄₉ cleavage site is selectively protected from digestion when the A subunit is bound to the B subunit, since fragments c (lane A3) and d (lane B3) are greatly reduced in intensity when the holoenzyme is treated with chymotrypsin (compare with lanes A6 and B6). Using a program that calculates the likelihood of local disorder in protein structure, VSL2 (Obradovic *et al*, 2005), we found three intrinsically disordered internal regions in the pol γ catalytic subunit, each of which coincided with the positions of chymotrypsin cleavage (Figure 4C; Supplementary Figure S5). The chymotrypsin cleavage site protected by binding of pol γB_2 resides within the largest disordered region (residues 478–581). Our protease protection results are consistent with the observation that pol γA bearing a mutation at residue 467 is deficient in its ability to bind pol γB_2 (Chan *et al*, 2005). Taken together, these results imply that pol γB_2 interacts most intimately with the spacer region of pol γA separating the exo and pol domains, possibly spanning residues that extend from residue 467 to L₅₄₉.

The position of the pol γA N-terminus in relation to the DNA template

To obtain further insight into the structure of the catalytic subunit, we took advantage of the ability to label the N-terminal His₆ tag of pol γA with a 1.4 nm gold cluster

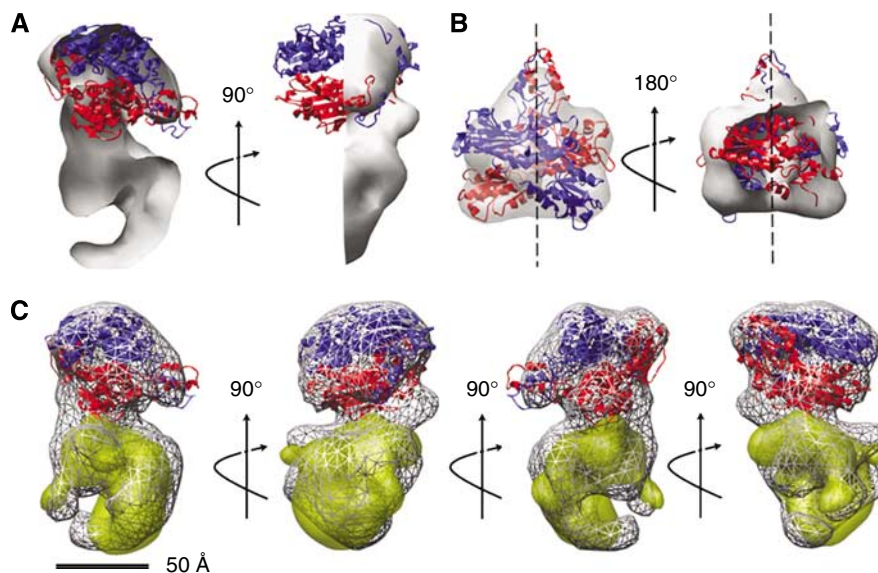


Figure 3 Subunit localization in the 3D model of the pol γ holoenzyme. Side (A) and top (B) views of the holoenzyme model, with the crystal structure of human DNA pol γB_2 docked into the upper lobe. The EM model is clipped to show the symmetrical fit of pol γB_2 (red and blue ribbons) more clearly. The symmetry axis of pol γB_2 is shown by a dashed line. The illustration on the right of panel B shows the upper, pol γB_2 , lobe from below with the lower lobe removed. The contact area between the B dimer (top) and the pol γ catalytic subunit (bottom) is formed mostly by one (red) subunit of the dimer. (C) Crystal structure of pol γB_2 (red and blue ribbons) and EM structure of pol γA (gold) docked simultaneously into the pol γ holoenzyme EM structure. The orientation of pol γA filtered to 20 Å and 105 kDa is as shown in Figure 1.

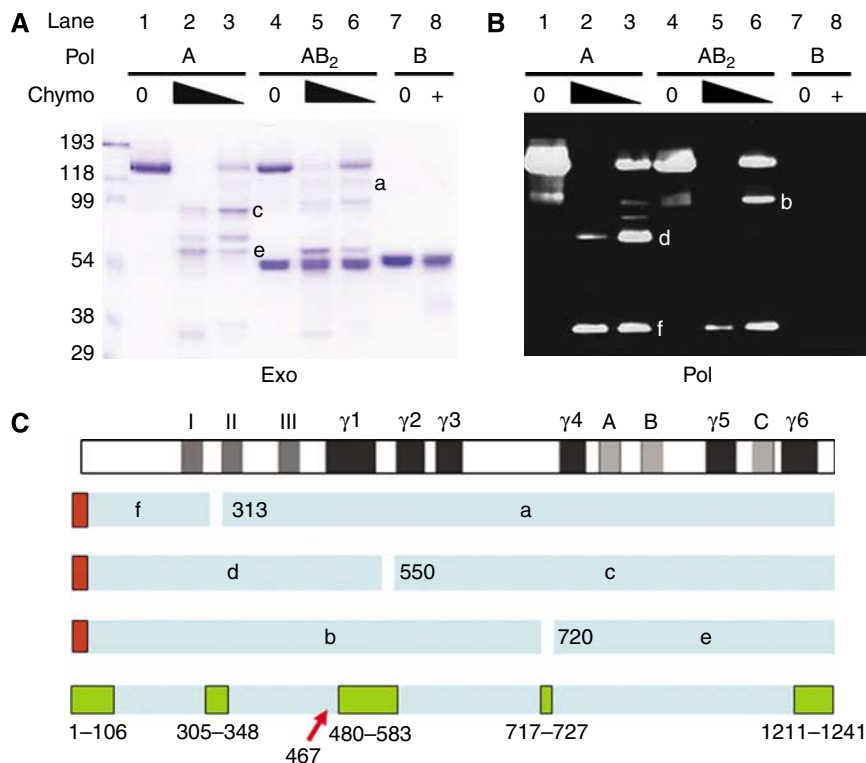


Figure 4 The EM model of pol γ explains the protection by pol γB_2 of a chymotrypsin cleavage site in pol γA . Partial proteolysis of pol γ components was used to identify protease-resistant domains. Chymotrypsin was used to treat either N-terminally His-tagged pol γA alone (lanes 1–3), or the same tagged pol γA complexed with untagged pol γB (lanes 4–6), or pol γB alone (lanes 7, 8). Reactions were stopped by the addition of PMSF and polypeptides were subjected to SDS-PAGE. (A) Coomassie-stained gel analysis. (B) Immunoblot to detect the N-terminal proteolytic fragments using an anti-His tag antibody. Six major fragments of pol γA were identified as products of three proteolytic cleavages. Fragments b, d and f identified in panel B retained the N-terminal His-tag (red boxes in panel C), whereas fragments a, c and e identified in panel A did not. No discrete degradation products of pol γB were observed under these conditions. (C) An interpretation of the pol γA fragments with respect to major sequence features. Subsequent N-terminal sequence analysis of fragments a, c and e identified cleavage sites as shown in the figure and discussed in the text. Disordered regions predicted by VSL2b are shown as green boxes at the bottom for comparison. The position of the pol γA mutation at residue 467 affecting binding to the accessory factor is indicated by a red arrow.

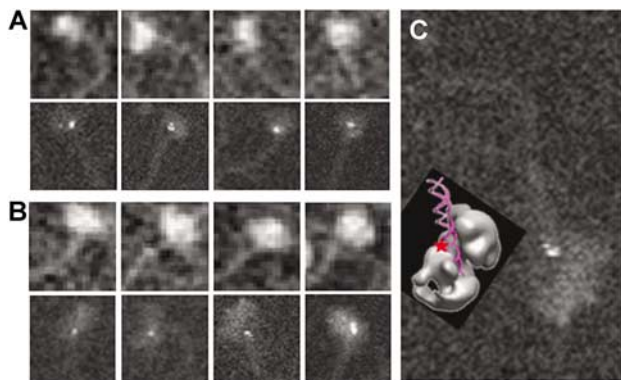


Figure 5 DNA binding by pol γ visualized using STEM. Images of pol γ A (A) or pol γ holoenzyme (B) bound to 350-bp-long double-stranded DNA with a 4-nt-long 5'-overhang obtained using STEM (only a part of the DNA is included in the box). The top row of each panel shows unlabeled protein. The bottom row represents protein bearing a 1.4 nm gold cluster at the N-terminus of the pol γ A polypeptide. The box size is 250 Å. The contrast of images of gold-labeled protein was scaled to make the gold label more visible. (C) Enlarged STEM image of DNA-bound pol γ . The EM model is shown in the inset for comparison with the putative position of the gold label marked by the red asterisk.

that can be visualized using scanning transmission electron microscopy (STEM). In contrast to brightfield TEM, this dark-field STEM technique can simultaneously visualize the projected mass of protein, DNA and a nanogold cluster. Although the resolution of STEM images is lower than that of TEM, we found that STEM was able to complement our other EM analyses. STEM images of pol γ A or holoenzyme labeled with Ni^{2+} -NTA-gold attached to the His₆-tag at the N-terminus of γ A are shown in Figure 5A and B, respectively. The polymerase was observed to be bound to a 300-bp-long double-stranded DNA restriction fragment serving as primer-template. The N-terminal gold label is positioned near the center of the pol γ molecule very close to the point where the double-stranded template issues from pol γ . The central orientation of the N-terminus is consistent with the model shown in Figure 1, if we consider that the extreme N-terminus may be folded into the center of the exo domain. It is possible that the DNA is engaged within the exo site of the enzyme in these STEM images.

Discussion

We present the first experimentally based model of human pol γ determined by EM methods. The shape of the enzyme and the relative positions of its catalytic and accessory subunits were resolved at 17-Å resolution. This model, in combination with other structural and functional data, allows us to draw significant conclusions regarding the structural organization of pol γ , its interaction with DNA and its enzymatic activity.

EM images show that the pol γ AB₂ heterotrimer resembles a distorted dumb-bell composed of two nearly equal parts. The catalytic and dimeric accessory subunits are connected through a relatively confined binding interface detectable as a bridge between the two subunits. The first determination of the primary structure of a higher-eukaryotic pol γ A (Ye *et al*, 1996) identified three consecutive domains within the pol γ A sequence: the N-terminal exo domain, a relatively large

spacer domain (in comparison with other family A polymerases) and the pol domain. The EM structure of pol γ (Figure 2) and its catalytic subunit alone (Figure 1) are entirely consistent with a model in which the exonuclease and polymerase domains are separated by the long spacer that contacts the accessory subunit dimer. Combined EM and limited proteolysis data (Figure 4), as well as previous mutational data (Chan *et al*, 2005), support the suggestion that the accessory subunit of pol γ interacts significantly with the spacer region of pol γ A. Higher-resolution structural information for pol γ A is required in order to provide a more detailed view of the portions of the catalytic subunit that contact pol γ B₂. The binding of pol γ B₂ to pol γ A alters many of the properties of the polymerase, including its affinity for both primer-template and nucleotides, as well as its processivity (reviewed in Kaguni, 2004; Graziewicz *et al*, 2006). It is possible that some portion of pol γ A undergoes a conformational change upon binding pol γ B₂. For example, a flexible loop in pol γ A that is not well ordered in the EM structure in Figure 1 may extend outward from the globular core of the catalytic subunit to contact pol γ B₂, as has been observed for the contact between the polymerase and thioredoxin subunits of T7 DNA pol (Doublet *et al*, 1998).

Unambiguous docking of the pol γ B₂ X-ray structure into the EM structure of pol γ suggests that no large-scale conformational transitions take place within the B₂ subunit upon binding to the catalytic subunit. One accessory subunit forms extensive contacts with the catalytic subunit while the second B subunit is mostly exposed to solvent (Figure 3). The insight provided by this model explains, for the first time, the properties of a well-studied deletion derivative of pol γ B that binds pol γ A as a monomer. This mutant, pol γ B Δ I4, is not readily able to dimerize, since it lacks the two-helix bundle that makes a major contribution to the dimerization interface in pol γ B₂ (Carrodeguas *et al*, 2001). We have recently shown that the kinetics and thermodynamics of the association of pol γ A with the dimeric (B₂) and monomeric (B Δ I4) forms of the accessory subunit are nearly equivalent (Yakubovskaya *et al*, 2006). Combining these earlier results with the EM structure suggests that the second subunit of the accessory dimer makes only a minor contribution to the primary interface with pol γ A. This model also provides insight into the peculiar situation of *Drosophila* pol γ , which is known to be a heterodimer containing only one copy of the accessory subunit (Olson *et al*, 1995). As noted previously (Yakubovskaya *et al*, 2006), *Drosophila* pol γ B differs from the human pol γ B in several respects, including a lack of amino acids corresponding to those deleted in the human pol γ B Δ I4 mutant. Based on this model, it is easy to appreciate how the *Drosophila* pol γ could have lost the second copy of the accessory subunit during evolution to become a heterodimer.

In our initial determination of the crystal structure of mouse pol γ B₂, we identified three domains in pol γ B (Carrodeguas *et al*, 2001). Docking the B₂ structure within the EM model demonstrates that the N-terminal domain 1 of pol γ B mediates most contacts with pol γ A. As shown in Figure 6, the contact area is formed by three α -helices (A, B and G) and three β -strands (7, 8 and 9) belonging to domain 1 (nomenclature as in Carrodeguas *et al*, 2001; see Supplementary Figures S6). Through efforts in our laboratory

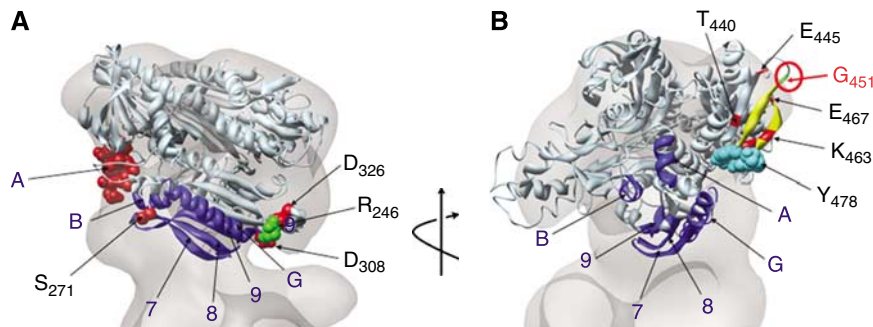


Figure 6 Positions of pol γ B mutations affecting its ability to enhance pol γ A catalytic activity. **(A)** Helices B, G and β -strands 7, 8, 9 that form the pol γ A to pol γ B₂ interface are shown in blue. Helix A also contributes to the binding interface, but is shown in red since its deletion in mutant dN81 greatly impairs pol γ activity. Other deleterious mutations adjacent to this region are also shown in red: S₂₇₁A affects the N-terminus of strand 7, D₃₀₈A/E₃₁₀A affect the C-terminus of strand 9, D₃₂₆ is involved in a charge interaction with R₂₄₆ (green) at the N-terminus of helix G (Fan *et al*, 2006). **(B)** The model in panel A is rotated to show additional pol γ B mutations that seriously affect pol γ activity. Helices and β -strands forming the interface, including helix A, are shown in blue. The β -strand 18–19 hairpin is shown in yellow. The deletion mutation dC30, which fails to stimulate pol γ , completely removes this hairpin; other deleterious mutations affecting amino acids situated either in this hairpin (K₄₆₃, H₄₆₇; Fan *et al*, 2006), or in close proximity to its base (T₄₄₀, E₄₄₅, T₄₄₇; Fan *et al*, 2006), D₄₇₃, F₄₇₄, K₄₇₇, Y₄₇₈; Pinz and Bogenhagen, 2006) are shown in red. The position of G₄₅₁ is encircled and shown by a red arrow. Residues 458–460 in the loop between β -strands 18 and 19 that are adjacent to the pol γ A to pol γ B interface are shown as cyan spheres.

and others, a series of pol γ B mutations is available that allows us to interpret aspects of the proposed binding model.

Pol γ B mutations with greatly decreased ability to stimulate pol γ A can be subdivided into two classes. Mutations of the first class are positioned in close proximity to the interface between the pol γ A and pol γ B subunits (Figure 6). Deletion of 15 amino-acid residues between positions 66–81 from the N-terminus (dN81), thus removing helix A, severely affects pol γ B activity. Point mutations within the pol γ A-binding interface also have substantial effects on activity, as illustrated by the S₂₇₁A, D₃₀₈A/E₃₁₀A and D₃₂₆A/R₃₂₈A mutations (Figure 6A) described by Fan *et al* (2006).

A second class of pol γ B mutations affects the hairpin formed by β -strands 18 and 19. This hairpin is absolutely required for γ B₂ to stimulate γ A activity (Carrodegua and Bogenhagen, 2000; Pinz and Bogenhagen, 2006). Point mutations affecting this hairpin include E₄₄₅A/T₄₄₇A and K₄₆₃A/H₄₆₇A mutations described by Fan *et al* (2006), as well as our mutations D₄₇₃A/K₄₇₇A and F₄₇₄A/Y₄₇₈A (Pinz and Bogenhagen, 2006) (Figure 6B). The most striking example of this class is the naturally occurring G₄₅₁E substitution that almost completely abolishes the ability of pol γ B to bind pol γ A and to stimulate polymerase activity. In human and murine pol γ B, Gly₄₅₁ is located in domain 3 at the N-terminus of β -strand 18 and adopts a conformation falling into the top right quadrant of the Ramachandran plot ($\phi = 150^\circ$, $\psi = 52^\circ$) that is restricted for all amino-acid residues except glycine. Consequently, introduction of the glutamate residue at this position could considerably distort the β -strand 18 and create a steric clash with residues at the end of β -strand 19. The dramatic effect of the Gly₄₅₁ substitution with Glu illustrates the requirement for proper β -strand 18–19 hairpin orientation for efficient pol γ B binding to pol γ A. Although the β -strand 18–19 hairpin is located at a distance from the A to B subunit interface, its loop (cyan spheres in Figure 6B) contacts the N-terminus of strand β 7 and, therefore, can interfere with the interaction with the pol γ A subunit.

In contrast to the close fit of the accessory subunit within the holoenzyme, docking the EM model of the catalytic

subunit is less definitive, since this has a rather globular shape. Although some similarity is observed between the bottom part of the pol γ EM model and the catalytic subunit, the pol γ A EM model scaled to 139 kDa is slightly greater than the volume of the larger lobe of the holoenzyme. In addition, as shown in Figure 3, some rather small segments of the EM structure of the holoenzyme are not precisely filled by simultaneously docking the A and B₂ structures. We cannot entirely rule out the possibility that the small DNA primer-template oligonucleotide contributes to this small discrepancy in volumes, since this was included only for preparations of the holoenzyme, not for the free catalytic subunit, although the literature suggests that its contribution should be very small (see, for example, Boskovic *et al*, 2003). We propose the alternative possibility that this discrepancy can be reconciled taking into account our proteolysis data and VSL2 predictions suggesting that large intrinsically disordered internal regions are present in the pol γ A subunit (Figure 4C; Supplementary Figure S5). Disordered regions are known to have a number of functions in proteins, including molecular recognition and binding. These regions undergo disorder-to-order transitions upon binding to their partners (Sickmeier *et al*, 2007), as has been observed for the contact between the polymerase and thioredoxin subunits of T7 DNA pol (Supplementary Figure S5; Doublet *et al*, 1998). Such a disordered loop might be poorly visible in the uranyl acetate-stained EM images of the isolated pol γ A subunit, due to variation in its position and penetration of the stain into the interior of the disordered regions. However, if this segment becomes ordered upon binding to pol γ B₂, this conformational stabilization would make it clearly visible in the holoenzyme structure. Thus, this effect may explain the observed volume discrepancies as well as the origin of the dense bridge connecting the two lobes that is clearly seen on some individual pol γ images and on most class averages (Supplementary Figure S2), and that cannot be easily explained by either pol γ B₂ or pol γ A alone. A conformational transition in pol γ A might help explain the pervasive changes in the properties of the enzyme when bound to its accessory

subunit. Binding pol γB_2 has been shown to increase affinity for primer–template and nucleotides, as well as to decrease fidelity of the catalytic subunit (Johnson *et al*, 2000; Longley *et al*, 2001), probably due to the relative displacement of polymerase and exonuclease domains.

Our model of the interaction between pol γB_2 and pol γA does not immediately explain the fact that in both this work and our earlier hydrodynamic studies (Yakubovskaya *et al*, 2006) we see no evidence for binding a second copy of pol γA to the pol γB protomer that is relatively exposed to solvent. One possible explanation for this phenomenon is that there may be a subtle conformational change in the pol γB dimer upon binding to pol γA that prevents interaction with a second copy of pol γA . This appears to be unlikely, since the pol γB_2 crystal structure fits into the EM structure of the holoenzyme without invoking conformational change. However, due to the low resolution of the EM structure, a subtle conformational change in pol γB_2 cannot be ruled out. A second possibility to explain the inability of a second molecule of pol γA to bind to pol γB_2 is that the nature of the surface of pol γA that interacts with pol γB_2 is poorly understood and may, as discussed above, involve a conformational change. The binding of a second copy of pol γA could be hindered by a steric clash.

Our EM studies to date do not provide definitive information on the path of DNA within pol γ . The small DNA oligonucleotide used as primer–template in EM analysis of the holoenzyme may be quite difficult to visualize under our experimental conditions (Griffith *et al*, 1999). Although observation of a DNA core was reported in the stained EM structure of a DNA-bound clamp–clamp loader complex obtained at 12-Å resolution (Miyata *et al*, 2005), the DNA was not resolved in a similar structure determined at lower resolution (Miyata *et al*, 2004). In our pol γ structure, with a resolution of about 17 Å, we could not directly observe the short primer–template oligonucleotide used in our studies. Therefore, we employed a second EM technique, STEM,

capable of visualizing both DNA and protein. STEM analysis revealed that the double-stranded DNA emerged from the center of pol γ , most commonly at an angle of ~ 40 – 60° relative to the long axis of pol γ .

We inspected our pol γ model to identify surface features that might participate in binding DNA primer–templates. Two prominent ~ 20 -Å-wide clefts at the interface between pol γA and pol γB_2 appear to be capable of accommodating a double-stranded DNA to generate two alternative models shown as rotated pairs in panels A and B of Figure 7. In order to consider whether these grooves could participate in guiding DNA into the active site of the polymerase, we asked how our EM data might accommodate the hypothetical (*in silico*) model of the pol γ active site, which presumes that amino-acid residues 871–1178 of pol γ will resemble the active site of the distantly-related T7 DNA polymerase (Graziewicz *et al*, 2004). A large fraction of the amino-acid residues conserved between pol γA and T7 DNA pol is located at the polymerase active site. Despite the relatively small size of this active site model, it can be docked into our EM structure in only a few orientations, if we assume that the DNA extends outward from the catalytic center in the manner suggested by Graziewicz *et al* (2004). The model of the polymerase center fits most easily in the position shown in the structural models in Figure 7A' and B'. Comparison of alternative models for docking of the polymerase active site suggests that the DNA emerging from the polymerase center may be accommodated more easily in one cleft, as shown in Figure 7A and A', than the other. It is interesting to speculate that the binding of DNA in the basic groove formed by the addition of pol γB_2 to pol γA may impede the ability of the enzyme to engage its exo function to edit polymerase errors. This may help explain the observation that the pol γ holoenzyme has lower replication fidelity than the pol γA catalytic subunit alone (Longley *et al*, 2001).

Recently Fan *et al* (2006) presented a hypothetical model for the structure of the human pol γ –DNA complex. This model

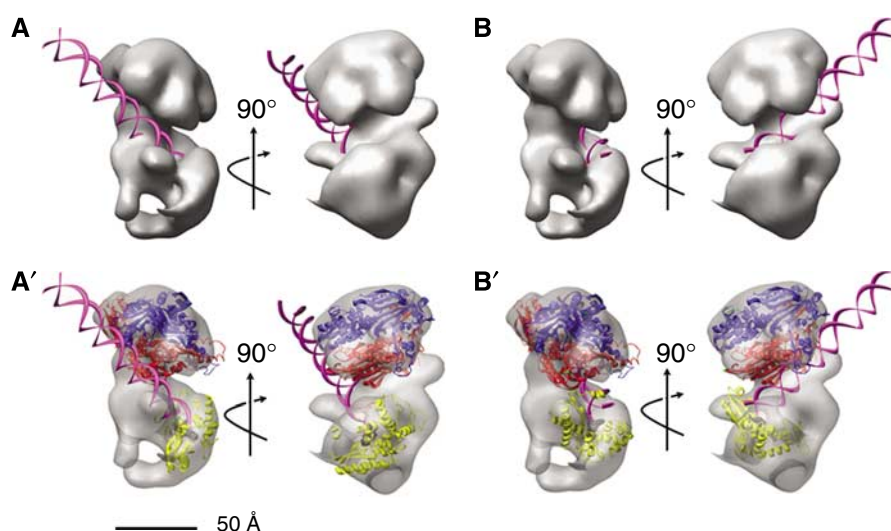


Figure 7 Hypothetical models of primer–template DNA binding to pol γ . (A, B) Models of DNA docked into clefts A and B of the EM structure of pol γ . The DNA helix is shown in magenta. Two rotated views of each model are shown. (A', B') Additional views of the models in panels A and B, with surfaces made transparent to show the following: the accessory subunit dimer (red and blue ribbons), the DNA (magenta ribbon), the predicted structure of the active center (residues 871–1178, yellow ribbon; Graziewicz *et al*, 2004) and the dNTP bound to the active site (black spheres). The direction of DNA emanating from the active site is determined by the active center configuration, but no assumptions are made concerning the possible distortion of DNA also shown.

assumes that the 1239-residue human pol γ A resembles the 704-residue T7 DNA polymerase, but is limited by the fact that these proteins show only 8.5% sequence identity. While the homology between pol γ A and T7 DNA polymerase is sufficient to permit modeling in the polymerase active site, as noted above, the primary sequence conservation in other regions is lower than 6%. Despite this limitation, Fan *et al* (2006) suggested that human pol γ A and B₂ subunits share an extensive binding interface that permits the holoenzyme to substantially encircle the nascent DNA. However, our present structural data, as well as previous studies (Yakubovskaya *et al*, 2006) demonstrating that one B subunit contributes most of the interaction energy with pol γ A, fail to support the clamp-type pol γ -DNA complex formation proposed by Fan *et al* (2006).

In summary, the EM model for the pol γ holoenzyme presented here provides considerable insight into the structure and function of the polymerase, explaining features of this enzyme that distinguish it from other family A DNA polymerases. In contrast to the Klenow fragment of *E. coli* DNA polymerase I, human pol γ contains a large spacer region separating its exo and pol domains. We show, for the first time, that this spacer mediates a focal contact with the dimeric accessory subunit. The facile placement of the accessory factor within the EM structure of the holoenzyme provides a model that explains several experimental observations: First, the model shows that pol γ A forms tight contacts principally with one pol γ B protomer, and permits analysis of the structural motifs of pol γ B that mediate this interaction. Second, the model is consistent with the altered properties of numerous pol γ mutants. Finally, the model permits us to predict how known DNA-binding domains of pol γ B may help guide the path of DNA along the holoenzyme in a way that enhances the processivity of the polymerase, but may impair its fidelity.

Materials and methods

Specimen preparation and EM imaging

Proteins were purified as described (Yakubovskaya *et al*, 2006). Pol γ holoenzyme and pol γ A (initial concentration 1 mg/ml) were diluted to 25 μ g/ml with binding buffer 20 mM Hepes, pH 8.0, 150 mM KCl, 1 mM EDTA, 5 mM DTT buffer. For the DNA:poly complex, enzyme was mixed with 26:45-mer oligonucleotide (Qiagen) in a 1:2 molar ratio and diluted to 25 μ g/ml protein with binding buffer. Immediately after dilution, 3- μ l aliquots of protein were applied onto 300-mesh copper grids (Ted Pella) coated with carbon film that were glow-discharged for 10 s. After 1 min the protein solution was blotted with filter paper and stained with 2% (w/v) uranyl acetate solution.

Samples were imaged under low-dose ($10\text{--}15\text{ e}^-/\text{\AA}^2$) conditions using an FEI Tecnai TF20 microscope operating at 200 kV. Micrographs with actual magnification 76 000 were taken at 0° or 50° angles, with underfocus values between 0.85 and 2.7 μ m, using a 4k \times 4k CCD camera (TVIPS F415) with an actual pixel size of 1.744 \AA /pixel. Subsequent two-fold pixel averaging produced a final sampling of 3.488 \AA /pixel.

Image processing and 3D reconstruction

The EMAN 1.8 software package was used for particle selection, classification and 3D reconstruction (Ludtke *et al*, 1999). Particles from untilted micrographs were selected semi-automatically using the BOXER program and then manually examined to remove poor quality images, typically about 5%. The box size was 48 \times 48 for pol γ A and 64 \times 64 for the holoenzyme. Tilted pairs were selected using xmipp_{mark} (Scheres *et al*, 2007; Xmipp program package). After selection, all images were high pass filtered at $1/150\text{ \AA}^{-1}$. Determination of defocus and CTF correction was performed using the CTFIT program

(Ludtke *et al*, 1999) as described in <http://blake.bcm.tmc.edu/eman/eman1/>. Raw image sets were subjected to a reference-free alignment and correspondence analysis using the *refine2d* python macros (EMAN 1.8 program package). Hierarchical ascendant classification was performed using Spider (Frank *et al*, 1996). Classes obtained using the *refine2d.py* were selected manually for initial model generation using the Fourier common-line procedure. Subsequent iterative refinement of starting models against the data set was performed using the REFINER routine as described (Brink *et al*, 2004). The *classiter* parameter was set to 8 to compensate for possible model bias and decreased gradually to a value of 3 in the last three iterations. Since pol γ has a two-fold pseudosymmetry, the orientation of each particle was monitored during the sixth–eighth iteration, and particles changing their orientation by more than 45° were excluded from subsequent iteration steps (820 particles were discarded). Stabilization of a Fourier Shell Coefficient (FSC) between successive iteration steps was used as a convergence criterion, which usually required 9–12 iterations. To estimate the resolution of the final structure, two structures were calculated on the basis of even or odd halves of the final class set using the EOTEST routine. According to the 0.5 FSC criterion, the resolution of the final structures was $\sim 17\text{ \AA}$, however, since the FSC tends to overestimate the resolution for structures of small molecules ($< 100\text{ \AA}$) (Yang *et al*, 2003), all final 3D models were low pass filtered to 20 \AA . The UCSF CHIMERA program (Pettersen *et al*, 2004) was used to create 3D graphical representations. The volumes for surface rendering were calculated using a known standard protein density of 0.81 Da/ \AA^2 and molecular mass of 246 kDa (for pol γ) or 105 kDa (for the pol γ A subunit). The last value was calculated on the basis of the known molecular weight of the pol γ A subunit (139 kDa), taking only the ordered part of the molecule into account. Size and position of disordered regions were calculated using the VSL2b predictor (Obradovic *et al*, 2005), using an output window size of 41 and a standard 0.5 criterion. The crystal structure of the accessory subunit dimer of human pol γ (PDB accession number 2G4C) was positioned interactively as a rigid body using CHIMERA or automatically using a SITUS *colores* command. The models of the catalytic subunit and the holoenzyme have been deposited in the EBI database (<http://www.ebi.ac.uk/msd/index.html>) with accession numbers EMD-1409 and EMD-1410, respectively.

Scanning transmission electron microscopy

A 100- μ l volume of 30 nmol Ni-NTA-Nanogold (Nanoprobes) was mixed with 100 μ l of 1 mg/ml pol γ A or pol γ holoenzyme with the His-tag at the N-terminus of the catalytic subunit, and incubated at 4°C for 3 h. Samples were diluted to 1 ml with 20 mM Hepes, pH 8.0, 300 mM KCl, 20 mM NiCl₂, 1 mM EDTA, 5 mM DTT buffer and applied to a 100 000 *M_r* Vivaspın ultrafiltration device to remove excess gold. After concentrating the sample to 200 μ l, an additional 1 ml of the buffer was added for each of three repetitions. After the final concentration to 200 μ l was achieved, the UV-Vis spectrum was collected and the number of gold clusters bound per base was calculating as described (Hainfeld, 1987). For STEM, samples from the column were fixed by adding 0.15% glutaraldehyde on ice for 30 min before grid making. After washings and application of an internal standard, tobacco mosaic virus, the sample was applied directly to thin carbon films on holey grids and allowed to stand for 1 min. The samples were wicked and washed with 20 mM ammonium acetate several times (washing also stops the fixation and removes free glutaraldehyde), frozen rapidly in liquid nitrogen slush, and freeze-dried overnight. STEM visualization was performed as described (Wall *et al*, 1998). Digital dark-field images of these freeze-dried samples were recorded with 512 \times 512 pixels at raster steps from 2.0 to 0.25 nm per pixel.

Limited proteolysis of human pol γ A and the holoenzyme

Holoenzyme or pol γ A (1 mg/ml) was incubated with equal volumes of serially diluted chymotrypsin, 10 or 3.3 μ g/ml, at 22°C for 5 min. The reaction was stopped by adding 2 μ l of 20 mM PMSF to 20- μ l reaction mixtures. Boiled samples were loaded onto 10% SDS-PAGE, subjected to electrophoresis and the gel was stained with Coomassie blue. In some cases, following SDS-PAGE, proteolysis products were also transferred onto polyvinylidene difluoride (PVDF) membranes. N-terminal fragments were detected using a 1:5000 dilution of mouse monoclonal His-probe antibody conjugated to horseradish peroxidase (Santa Cruz), and chemiluminescent imaging. Alternatively, fragments containing the pol domain were detected using a 1:5000 dilution of rabbit polyclonal

antibody directed against a conserved sequence in the Pol γ active site, KVFNYGRIYGAGC (Ye *et al*, 1996; residues 947–959 in human). Bound antibody was detected with goat anti-rabbit IgG secondary antibody and visualized using BCIP/NBT phosphatase substrate (K & P). A parallel blot of a partial chymotryptic digest of pol γ A was stained with Coomassie blue and the C-terminal proteolytic fragments were sequenced from the N-terminus by Edman degradation using an automated protein sequencer (Procise 494, Applied Biosystems, CA).

Supplementary data

Supplementary data are available at *The EMBO Journal* Online (<http://www.embojournal.org>).

References

- Boskovic J, Rivera-Calzada A, Maman JD, Chacon P, Willison KR, Pearl LH, Llorca O (2003) Visualization of DNA-induced conformational changes in the DNA repair kinase DNA-PKcs. *EMBO J* **22**: 5875–5882
- Brink J, Ludtke SJ, Kong Y, Wakil SJ, Ma J, Chiu W (2004) Experimental verification of conformational variation of human fatty acid synthase as predicted by normal mode analysis. *Structure* **12**: 185–191
- Carrodegua JA, Bogenhagen DF (2000) Protein sequences conserved in prokaryotic aminoacyl-tRNA synthetases are important for the activity of the processivity factor of human mitochondrial DNA polymerase. *Nucleic Acids Res* **28**: 1237–1244
- Carrodegua JA, Kobayashi R, Lim SE, Copeland WC, Bogenhagen DF (1999) The accessory subunit of *X. laevis* mitochondrial DNA polymerase γ increases processivity of the catalytic subunit of human DNA polymerase γ and is related to class II amino acyl tRNA synthetases. *Mol Cell Biol* **19**: 4039–4046
- Carrodegua JA, Theis K, Bogenhagen DF, Kisker C (2001) Crystal structure and deletion analysis show that the accessory subunit of mammalian DNA polymerase γ , pol γ B, functions as a homodimer. *Mol Cell* **7**: 43–54
- Chan SSL, Longley MJ, Copeland WC (2005) The common A467T mutation in the human mitochondrial DNA polymerase (POLG) compromises catalytic efficiency and interaction with the accessory subunit. *J Biol Chem* **280**: 31341–31346
- Double S, Tabor S, Long AM, Richardson CC, Ellenberger T (1998) Crystal structure of a bacteriophage T7 DNA replication complex at 2.2 Å resolution. *Nature* **391**: 251–258
- Fan L, Kim S, Farr CL, Schaefer KT, Randolph KM, Tainer JA, Kaguni LS (2006) A novel processive mechanism for DNA synthesis revealed by structure, modeling and mutagenesis of the accessory subunit of human mitochondrial DNA polymerase. *J Mol Biol* **358**: 1229–1243
- Frank J, Radermacher M, Penczek P, Zhu J, Li Y, Ladjadi M, Leith A (1996) SPIDER and WEB: processing and visualization of images in 3D electron microscopy and related fields. *J Struct Biol* **116**: 190–199
- Graziewicz MA, Longley MJ, Bienstock RJ, Zeviani M, Copeland WC (2004) Structure–function defects of human mitochondrial DNA polymerase in autosomal dominant progressive external ophthalmoplegia. *Nat Struct Mol Biol* **11**: 770–776
- Graziewicz M, Longley M, Copeland W (2006) DNA polymerase γ in mitochondrial DNA replication and repair. *Chem Rev* **106**: 383–405
- Griffith J, Michalowski S, Makhov A (1999) Electron microscopy of DNA–protein complexes and chromatin. *Meth Enzymol* **304**: 214–230
- Hainfeld J (1987) A small gold-conjugated antibody label: improved resolution for electron microscopy. *Science* **236**: 450–453
- Henderson R (2004) Realizing the potential of electron cryo-microscopy. *Q Rev Biophys* **37**: 3–13
- Ito J, Braithwaite D (1990) Yeast mitochondrial DNA polymerase is related to the family A DNA polymerases. *Nucleic Acids Res* **18**: 6716
- Johnson AA, Tsai Y-c, Graves SW, Johnson KA (2000) Human mitochondrial DNA polymerase holoenzyme: reconstitution and characterization. *Biochemistry* **39**: 1702–1708
- Kaguni LS (2004) DNA polymerase γ , the mitochondrial replicase. *Annu Rev Biochem* **73**: 293–320
- Kujoth GC, Hiona A, Pugh TD, Someya S, Panzer K, Wohlgemuth SE, Hofer T, Seo AY, Sullivan R, Jobling WA, Morrow JD, Van Remmen H, Sedivy JM, Yamasoba T, Tanokura M, Weindruch R, Leeuwenburgh C, Prolla T (2005) Mitochondrial DNA mutations, oxidative stress, and apoptosis in mammalian aging. *Science* **309**: 481–484
- Longley MJ, Graziewicz MA, Bienstock RJ, Copeland WC (2005) Consequences of mutations in human DNA polymerase γ . *Gene* **354**: 125–131
- Longley MJ, Nguyen D, Kunkel TA, Copeland WC (2001) The fidelity of human DNA polymerase γ with and without exonucleolytic proofreading and the p55 accessory subunit. *J Biol Chem* **276**: 38555–38562
- Ludtke SJ, Baldwin PR, Chiu W (1999) EMAN: semiautomated software for high-resolution single-particle reconstructions. *J Struct Biol* **128**: 82–97
- Miyata T, Oyama T, Mayanagi K, Ishino S, Ishino Y, Morikawa K (2004) The clamp-loading complex for processive DNA replication. *Nat Struct Mol Biol* **11**: 632–636
- Miyata T, Suzuki H, Oyama T, Mayanagi K, Ishino Y, Morikawa K (2005) Open clamp structure in the clamp-loading complex visualized by electron microscopic image analysis. *Proc Natl Acad Sci USA* **102**: 13795–13800
- Obradovic Z, Peng K, Vucetic S, Radivojac P, Dunker AK (2005) Exploiting heterogeneous sequence properties improves prediction of protein disorder. *Proteins* **61**: 176–182
- Olson M, Wang Y, Elder R, Kaguni L (1995) Subunit structure of mitochondrial DNA polymerase from *Drosophila* embryos. *J Biol Chem* **270**: 28932–28937
- Petersen EF, Goddard TD, Huang CC, Couch GS, Greenblatt DM, Meng EC, Ferrin TE (2004) UCSF Chimera—A visualization system for exploratory research and analysis. *J Comput Chem* **25**: 1605–1612
- Pinz KG, Bogenhagen DF (2006) The influence of the DNA polymerase γ accessory subunit on base excision repair by the catalytic subunit. *DNA Repair* **5**: 121–128
- Scheres SHW, Gao H, Valle M, Herman GT, Eggermont PPB, Frank J, Carazo J-M (2007) Disentangling conformational states of macromolecules in 3D-EM through likelihood optimization. *Nat Methods* **4**: 27–29
- Sickmeier M, Hamilton JA, LeGall T, Vacic V, Cortese MS, Tantos A, Szabo B, Tompa P, Chen J, Uversky VN, Obradovic Z, Dunker AK (2007) DisProt: the database of disordered proteins. *Nucleic Acids Res* **35**: D786–D793
- Trifunovic A, Wredenberg A, Falkenberg M, Spelbrink J, Rovio A, Bruder C, Bohlooly-Y M, Gidlof S, Oldfors A, Wibom R, Tornell J, Jacobs HT, Larsson N-G (2004) Premature ageing in mice expressing defective mitochondrial DNA polymerase. *Nature* **429**: 417–423
- Wall J, Hainfeld J, Simon M (1998) Scanning transmission electron microscopy of nuclear structures. *Methods Cell Biol* **53**: 139–164
- Yakubovskaya E, Chen Z, Carrodegua JA, Kisker C, Bogenhagen DF (2006) Functional human mitochondrial DNA polymerase γ forms a heterotrimer. *J Biol Chem* **281**: 374–382
- Yang S, Yu X, Galkin VE, Egelman EH (2003) Issues of resolution and polymorphism in single-particle reconstruction. *J Struct Biol* **144**: 162–171
- Ye F, Carrodegua JA, Bogenhagen DF (1996) The γ subfamily of DNA polymerases: cloning of a developmentally regulated cDNA encoding *Xenopus laevis* mitochondrial DNA polymerase γ . *Nucleic Acids Res* **24**: 1481–1488
- Zheng W, Khrapko K, Collier HA, Thilly WG, Copeland WC (2006) Origins of human mitochondrial point mutations as DNA polymerase γ -mediated errors. *Mutation Research/Fundamental and Molecular Mechanisms of Mutagenesis* **599**: 11–20