[Cell Cycle 7:19, 3026-3036; 1 October 2008]; ©2008 Landes Bioscience

Report

X-ray structure of the complex of regulatory subunits of human DNA polymerase δ

Andrey G. Baranovskiy,[1] Nigar D. Babayeva,[1] Victoria G. Liston,[1] Igor B. Rogozin,[2] Eugene V. Koonin,[2] Youri I. Pavlov,[1] Dmitry G. Vassylyev[3] and Tahir H. Tahirov[1,*]

[1]Eppley Institute for Research in Cancer and Allied Diseases; University of Nebraska Medical Center; Omaha, Nebraska USA; [2]National Center for Biotechnology Information; National Library of Medicine; National Institutes of Health; Bethesda, Maryland USA; [3]Department of Biochemistry and Molecular Genetics; Schools of Medicine and Dentistry; University of Alabama at Birmingham; Birmingham, Alabama USA

Abbreviations: Pol δ, DNA polymerase δ; PCNA, proliferating cell nuclear antigen; MIR, multiple isomorphous replacement; PDE, phosphodiesterase; OB domain, oligonucleotide/oligosaccharide binding domain; wHTH, winged helix-turn-helix

Key words: DNA polymerase δ, Pol δ, p50, p66, Pol31, Pol32, OB, Myb, phosphodiesterase, human, yeast

The eukaryotic DNA polymerase δ (Pol δ) participates in genome replication, homologous recombination, DNA repair and damage tolerance. Regulation of the plethora of Pol δ functions depends on the interaction between the second (p50) and third (p66) non-catalytic subunits. We report the crystal structure of p50•p66$_N$ complex featuring oligonucleotide binding and phosphodiesterase domains in p50 and winged helix-turn-helix N-terminal domain in p66. Disruption of the interaction between the yeast orthologs of p50 and p66 by strategic amino acid changes leads to cold-sensitivity, sensitivity to hydroxyurea and to reduced UV mutagenesis, mimicking the phenotypes of strains where the third subunit of Pol δ is absent. The second subunits of all B family replicative DNA polymerases in archaea and eukaryotes, except Pol δ, share a three-domain structure similar to p50•p66$_N$, raising the possibility that a portion of the gene encoding p66 was derived from the second subunit gene relatively late in evolution.

Introduction

Efficient replication and repair of genomic DNA is vitally important for the inheritance in humans and for the prevention of genetic disorders and cancer. Pol δ is one of the three major replicases in eukaryotes.[1,2] In concert with Pol α, it can synthesize both leading and lagging DNA strands in the SV40 system and in yeast strains with no catalytic subunit of Pol ε.[3,4] It belongs to B family DNA polymerases, one of the largest DNA polymerase families, with representatives in many organisms, from prokaryotes, archaea, phages and viruses to multi-cellular eukaryotes.[5] The structure of several B family polymerases from these groups, except eukaryotes, has been solved. Classic example is the single-subunit replicase of bacteriophage RB69, in a complex with DNA and a nucleotide.[6] Because of homology it is likely that catalytic subunits of eukaryotic polymerases possess similar organization. Structural information from lower organisms can be used to model the effects of amino acid substitutions on functions of eukaryotic B family polymerases.[7] However, proper Pol δ function relies on a multi-protein complex of catalytic subunit with tightly associated regulatory subunits, whose structure is unknown. The only information available is a model of yeast Pol ε built on the basis of cryo-electron microscopy.[8] It provides the overview of possible DNA polymerase complex architecture with 20 Å resolution but does not reveal the details of interactions between subunits necessary to understand the mechanism of polymerase transactions.

The best studied genetically and biochemically are Pol δ's from yeast: a three-subunit complex from budding yeast *S. cerevisiae* and a four-subunit complex from fission yeast *S. pombe*. The largest catalytic subunit contains the polymerase and 3'→5' exonuclease active sites domains, as well as regions for protein-protein interactions,[9] for example a PCNA binding motif.[2] The essential second subunit serves as a stabilizer for the catalytic subunit and as a matchmaker with the third subunit. The third subunit plays several important roles. The N-terminal part interacts with the second subunit whereas the C-terminal part contains a conserved PCNA-binding motif and a motif that mediates interaction with Pol α.[10-12] The gene encoding the third subunit, *POL32*, is dispensable for growth in budding yeast, suggesting that the two-subunit Pol δ retains its role in bulk replication.[2] However, other critical biological processes are severely impaired. Deletion of the *POL32* gene renders yeast UV-sensitive and UV-immutable,[13-15] suggesting a role of the Pol32 protein in the regulation of error-prone trans-lesion DNA synthesis (TLS).[5,16] It was recently found that Pol32 is also required for break-induced recombination (BIR) and for telomerase-independent telomere maintenance.[17]

Human Pol δ consists of four subunits: p125, p50, p66 and p12. The first three subunits correspond to the three subunits of *S. cerevisiae* Pol δ. No catalytic activity is associated with the auxiliary

*Correspondence to: Tahir H. Tahirov; Eppley Institute for Research in Cancer and Allied Diseases; University of Nebraska Medical Center; 987696 Nebraska Medical Center; 668 South 41st Street; Lied Transplant Center Bldg.; Room 10737A; Omaha, Nebraska 68198-7696 USA; Tel.: 402.559.7607; Fax: 402.559.3739; Email: ttahirov@unmc.edu

Submitted: 08/01/08; Accepted: 08/04/08

Previously published online as a *Cell Cycle* E-publication: http://www.landesbioscience.com/journals/cc/article/6720

subunits, and they are thought to play a regulatory role, stimulate the polymerase activity of p125 by mediating additional interactions with PCNA, and stabilize the entire Pol δ complex.[18,19] p50 serves as a scaffold for the assembly of Pol δ by interacting simultaneously with all of the other three subunits.[20] In addition, p50 is also involved in the recruitment of several proteins regulating DNA metabolism, including p21,[21] PDIP1, PDIP38, PDIP46 and WRN (reviewed in ref. 21).

The parts of p50 responsible for interactions with p66, p125 and p12 have not been defined. Using two-hybrid screening, the human p66 has been shown to contain p50- and PCNA-binding domains within the 144 N- and 20 C-terminal amino acids, respectively.[22] Interestingly, many essential functions of Pol δ, including the regulation of replication, TLS and BIR, are mediated by the third subunit[13,15,17,23,24] and thus apparently depend on the interaction between its second and third subunits. Here we report the crystal structure of the complex between the second p50 subunit and the 144 amino acids N-terminal domain of the third p66 subunit (p66$_N$) of human Pol δ. The structure revealed a formation of a tight complex between these two subunits of Pol δ, confirmed by two-hybrid analysis of interface mutants. The biological significance of p50•p66 interaction has been demonstrated by genetic experiments in yeast. The results have broad implications for understanding the structure, function and evolution of replicative DNA polymerase complexes.

Results

Overall structure of p50•p66$_N$. The crystal structure of p50•p66$_N$ has been determined using the multiple isomorphous replacement (MIR) method and refined at 3 Å resolution. Four discrete molecules of p50•p66$_N$ without apparent oligomerization were located in an asymmetric unit and each molecule contained all amino acid residues of p66$_N$ and p50, except disordered regions in four loops and the C-terminal tail of p50. The structure revealed an extended cashew-shaped molecule with overall dimensions of 100 x 60 x 50 Å^3 and three domains: phosphodiesterase-like (PDE) and oligonucleotide/oligosaccharide binding (OB) domains in p50, and a winged helix-turn-helix (wHTH) domain in p66 (Figs. 1A and B, and 2). The PDE domain is located in the center of the molecule and bound to an OB domain on one side, and a wHTH domain on the other side. The PDE and OB domains were found in the second subunits of replicative B family DNA polymerases covering all species from archaea to humans.[25,26]

p50 domains. The PDE domain is formed of amino-acid residues 1–41 and 180–469 and comprised of a two-layer β-sheet (β_{15}, β_{16}, β_{14}, β_{13}, β_{12} and β_2 in layer 1; β_{18}, β_{17}, β_8, β_9, β_{10} and β_{11} in layer 2) with α-helices α_7 and α_8 flanking on one side, and α_1, α_3 - α_6 flanking on the other side (Figs. 2 and 3A). The active site of PDEs in archaeal DNA polymerases contains conserved histidine and aspartate residues involved in metal coordination and catalysis.[25] However, catalytic residues are replaced in the second subunits of the eukaryotic B family DNA polymerases, rendering them inactive.[25] It has been proposed that the inactive PDE domain of eukaryotic polymerases is involved in protein-protein interactions. In order to elucidate the structural features enabling the PDE domain of p50 to accommodate multiple protein-protein interactions, we compared it with the structure of a monomeric single subunit PDE protein MJ0936[27] (PDB code 1s3l and Fig. 4A). The 165 aa residue MJ0936

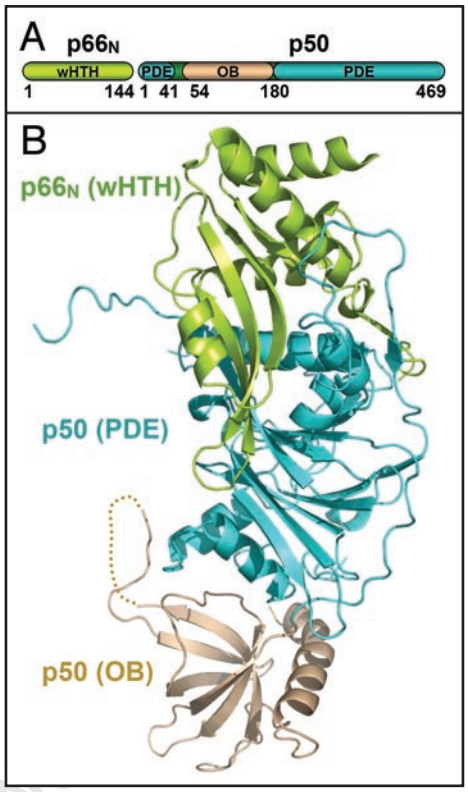

Figure 1. Overall view of p50•p66$_N$. (A) Schematic representation of p50 and p66$_N$ domains in relation to their sequence numbering. (B) Cartoon representation of p50•p66$_N$. The wHTH, PDE and OB domains are color coded. The (B) was prepared with PyMol software (Delano Scientific).

has a compact globular shape. In contrast, the larger 331 aa residue PDE domain of p50 contains many protruding parts and exhibits several notable differences both in the central two-layer β-sheet as well as in the flanking helical regions (Fig. 4A). In particular, β_2 of the N-terminal is added to layer 1 anti-parallel to β_{12}, while the β_{18} from layer 2 is extended to interact with β_3 of p66. The flanking regions of p50 PDE have three additional α-helices, including a significantly longer helix α_1 and longer loops, which together provide a larger surface area (12858 Å^2 in p50 vs 6579 Å^2 in MJ0936) and the flexibility to form many protein-protein interactions.

The OB domain (aa 54–179) is connected with the PDE domain via two linkers. One linker (aa 180–192) is a part of the PDE domain and is well defined in crystal, whereas another linker (aa 42–53) is disordered. The OB domain consists of a five-stranded β-barrel[28] (β_3–β_7) wrapped by helix α_2 on one side (Fig. 4B). The Pro63-introduced bend in this helix shows excellent shape complementarity to the β-barrel. Another side of β-barrel remains fully accessible and, considering the analogy to OB domains of replication protein A (RPA), this side may have a potential for DNA and protein binding.[29,30] The loop $\beta_3\beta_4$ in the OB domain of p50, which is disordered and significantly longer than corresponding DNA-interacting loops in OB domains of RPA, might also contribute to p50 interaction with DNA or protein (Fig. 4B). In addition to two peptide links, the interaction between the OB and PDE domains also involves 17 hydrogen bonds and several scattered hydrophobic contacts (Table S1), and buries a surface area of 2049 Å^2, which is within the range for stable protein-protein interactions.[31]

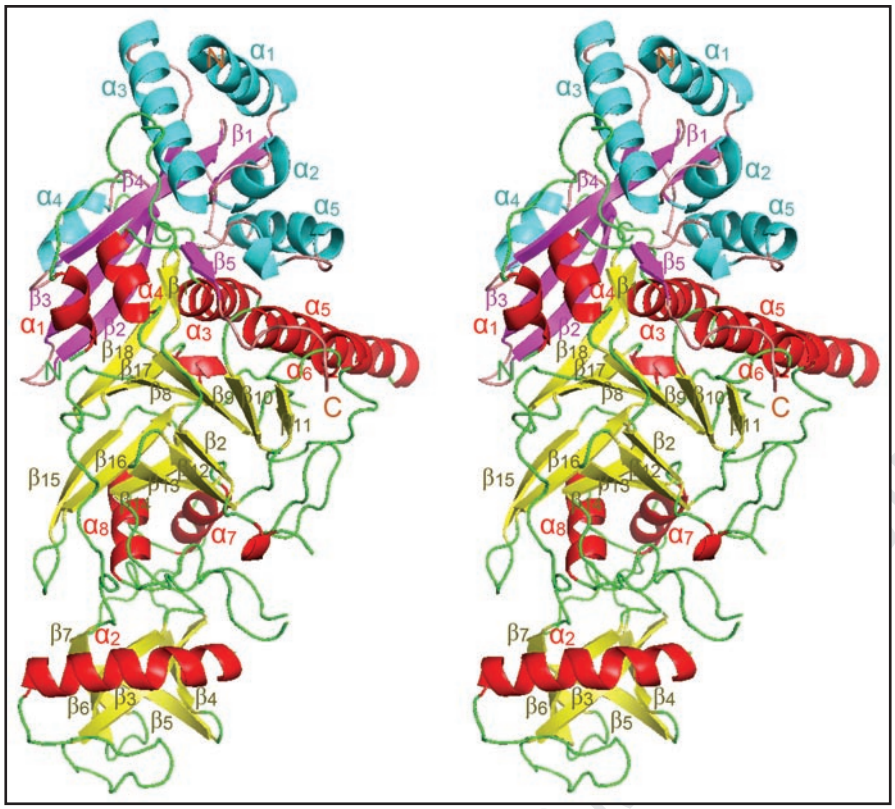

Figure 2. Stereo view of p50•p66$_N$ cartoon. The secondary structure elements are color coded and labeled. α-helices, β-strands and coils are red, yellow and green in p50, and cyan, magenta and light pink in p66$_N$. The figure was prepared with PyMol software (Delano Scientific).

p66$_N$ domain. A p50-bound p66$_N$ is comprised of a four-stranded, anti-parallel β-sheet surrounded by five α-helices and a C-terminal tail, and has a V-shaped structure (Fig. 4C). Search for 3D domains similar to the N-terminal 57 amino acids fragment using the Secondary Structure Matching (SSM) software[32] revealed the closest match with the Repeat 1 (R1) of c-Myb DNA-binding domain (PDB code 1h88 and Fig. 4C) comprised of helix-turn-helix-turn-helix motif. Unlike the other two repeats of c-Myb (R2R3), R1 has not detectable DNA-binding activity but increases the affinity of c-Myb R1R2R3 to DNA five- to ten-fold by providing the long-range electrostatic interaction between the positively charged surface of R1 and the negatively charged surface of DNA.[33] The SSM search with the addition of the next 58 aa revealed the closest match with an NMR structure of the *Pyrococcus furiosus* protein PF0610 (PDB code 2gmg and Fig. 4C) containing a wHTH motif and having a verified non-specific DNA-binding activity.[34] No other domains were found by SSM search with a larger p66$_N$ fragment. Thus, the N-terminal 105 amino acids of p66 represent a novel variation of the wHTH motif that is often found in dsDNA binding domains of transcriptional activators and repressors. We were unable to detect a DNA-binding by p50•p66$_N$ using gel shift assay (data not shown). However, similarly to R1, the distribution of positively charged residues (Lys19, Arg47, Lys48, Lys50, Lys105 and Lys109) on the putative DNA backbone facing surface of p66 suggest that these residues might contribute to the additional electrostatic interactions of Pol δ with dsDNA.

p50•p66 interactions. The p66$_N$ appears to be a natural extension of p50 and has a significantly large (5398 Å^2) surface area buried at the dimer interface. The amino acid residues involved in the p50•p66 interaction are highlighted in Figure 5. Layer 2 of a PDE two-layer β-sheet is joined with the β-sheet of p66 via parallel strands β_{18} (p50) and β_3 (p66), and together they form an extended 10-stranded β-sheet (Fig. 3B). One more intersubunit β-β interaction is observed between the parallel β-strands β_1 (p50) and β_5 (p66). Thirty-seven intersubunit hydrogen bonds, including 12 main-chain to main-chain hydrogen bonds, contribute to p50•p66$_N$ dimer formation (Table S1). Intersubunit hydrophobic interactions covering one larger area and three smaller areas also contribute to the stability of the complex. Furthermore, the presence of a negatively charged p66-interacting surface in p50 (Fig. 6A) and a complementary, positively charged surface of p66 (Fig. 6B) suggests that intersubunit electrostatic interactions also contribute to complex stabilization.

For the functional studies of p50•p66 interaction it would be appropriate to generate point mutations capable of disrupting the complex without affecting the fold of its subunits. One approach to achieve this result would involve elimination of intersubunit interactions by removing the side chains involved in interactions. However, these mutations will not disrupt the main-chain to main-chain hydrogen bonds. Due to the aforementioned extensive inter-subunit interactions, the impact of one or two of interacting side chains truncations (e.g., by mutation to alanine) is expected to be low. The second approach would involve generation of amino acid changes capable of disrupting the p50•p66 interactions by sterical hindrance with a bulky side chain or by repulsion of side chains with the same charge. To increase the impact of sterical hindrance or repulsion

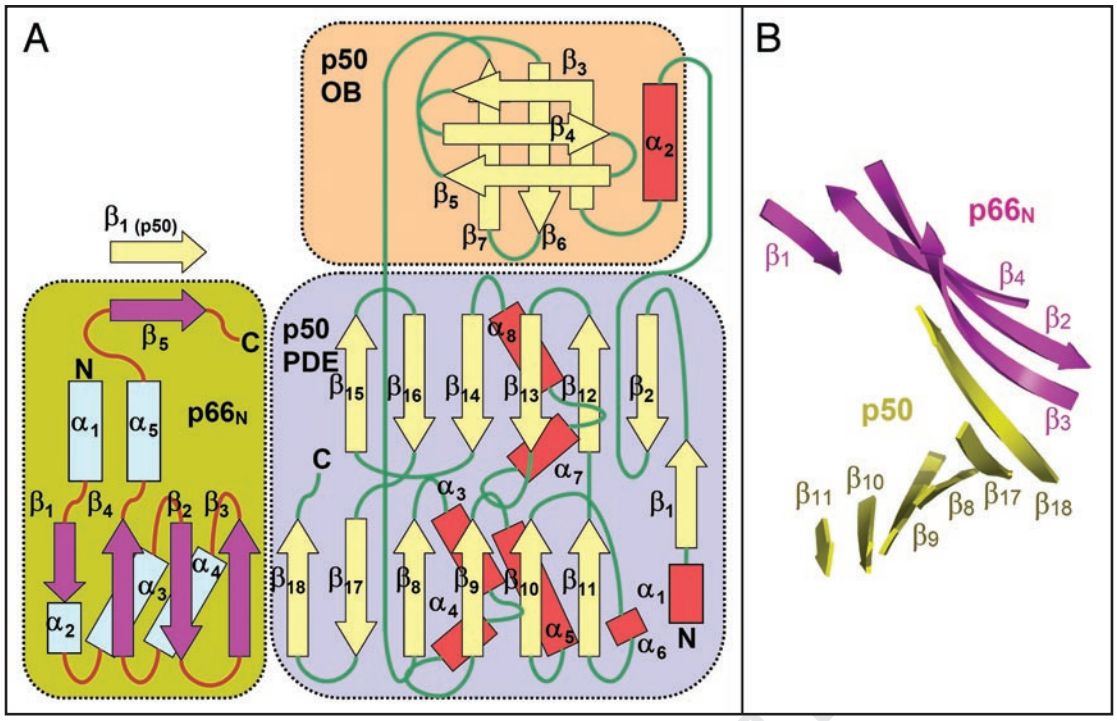

Figure 3. The topology of p50•p66$_N$. (A) The topology diagram of p50 and p66$_N$ subunits with labeling of β-strands and α-helices. (B) Cartoon representation of a 10-stranded β-sheet formed by joining of β_3 from p66 β-sheet (magenta) and β_{18} from a layer 2 (yellow) of a PDE two-layer β-sheet. The (B) was prepared with PyMol software (Delano Scientific).

by amino acid substitution, it is important that the introduced side chain did not turn away from the facing molecule. Guided by computer modeling, we selected three sites in p50 (Leu217, Gly224 and Glu231) to generate single and double tryptophan substitutions (Fig. 5), aiming to disrupt the p50•p66 interactions by sterical hindrance. The effects of the mutations encoding selected amino acid changes were examined by two-hybrid assay (Fig. 6C). p50 and p66$_N$ exhibit strong interactions in the yeast Matchmaker GAL4 two-hybrid system for two reporters in the two combinations of fusions of these proteins to activating or DNA binding domains (Fig. 6C, upper row of plates). The described above single and double amino acid changes, depending on the nature of the two-hybrid construct, abolished interaction either for both reporters (middle row of plates), or precluded the interaction for β-galactosidase reporter only (bottom row). Thus the generated mutations lead to disruption of p50•p66 interaction.

Functional studies using yeast orthologs of p50 and p66. The main disadvantage of functional studies with deletion mutants is the high probability of protein fold disruption. For example, β_{18} of p50 interacts with p66, but β_{18} is also a part of the two-layer β-sheet of the PDE domain. Thus, the deletion of β_{18}, in addition to the disruption of p50•p66 interaction, also might affect the stability of p50. In the absence of a three-dimensional structure, residues located, mainly, in the highly conserved regions were previously selected for amino acid changes and genetic studies of yeast p50 orthologs, Pol31 (in *S. cerevisiae*)[35] and Cdc1 (in *Schizosaccharomyces pombe*).[36] Using the structure of p50•p66$_N$, we mapped the temperature-sensitive Pol31 mutants[35] to the secondary structure elements of p50 that are necessary for its proper folding (Fig. 5). Thus, the temperature-sensitive phenotypes were, probably, due to deformations in the overall folding of Pol31. In contrast, replacement of a conserved threonine (Thr381 in p50) located in close proximity to the exposed and disordered loop $\alpha_7\beta_{14}$ (Fig. 5) produced a cold-sensitive allele.[35] The latter mutation is unlikely to induce significant changes in the folding of Pol31; instead, it might disrupt the protein-protein interactions, e.g., with the Pol3 or Pol δ-interacting factor.

Predictions of the secondary structures of bakers yeast Pol31 and Pol32 (yeast ortholog of p66) made by Phyre software[37] suggest that they are folded similarly to p50 and p66, respectively. Consequently, in spite of the differences in fine map of protein-protein interaction interfaces, the overall mode of Pol31•Pol32 complex assembly is expected to be similar to that of p50•p66. Thus, we used the structure of p50•p66$_N$ for the design of amino acid changes that will disrupt the Pol31•Pol32 interaction. In case of yeast proteins we applied a combination of sterical hindrance and charge repulsion. We introduced an Arg-Arg-Gly insertion after Agr243 in the loop $\alpha_3\alpha_4$ located in the center of the Pol31•Pol32 interface. The addition of glycine is expected to reduce the main-chain tension introduced by the insertion of arginines and will not disrupt the overall fold of Pol31. In the second Pol31 variant, an Arg-Arg-Gly insertion was combined with two point mutations Gly242Trp, Asp249Trp. We introduced the appropriate mutations (*pol31-RRG* and *pol31-WRRGW*) into the yeast strain allowing for concomitant detection of induced and spontaneous mutations in three genetic loci.[38] Neither of the mutations led to a growth defect or temperature sensitivity, indicating that the mutations in this essential gene do not affect the general properties of the second subunit (Fig. 7A). Both mutations, however, led to mild cold-sensitivity, increased sensitivity to hydroxyurea (Fig. 7A), elevated sensitivity to UV light irradiation (Fig. 7B) and reduced UV mutagenesis (Fig. 7C), in comparison to the parent strain. The

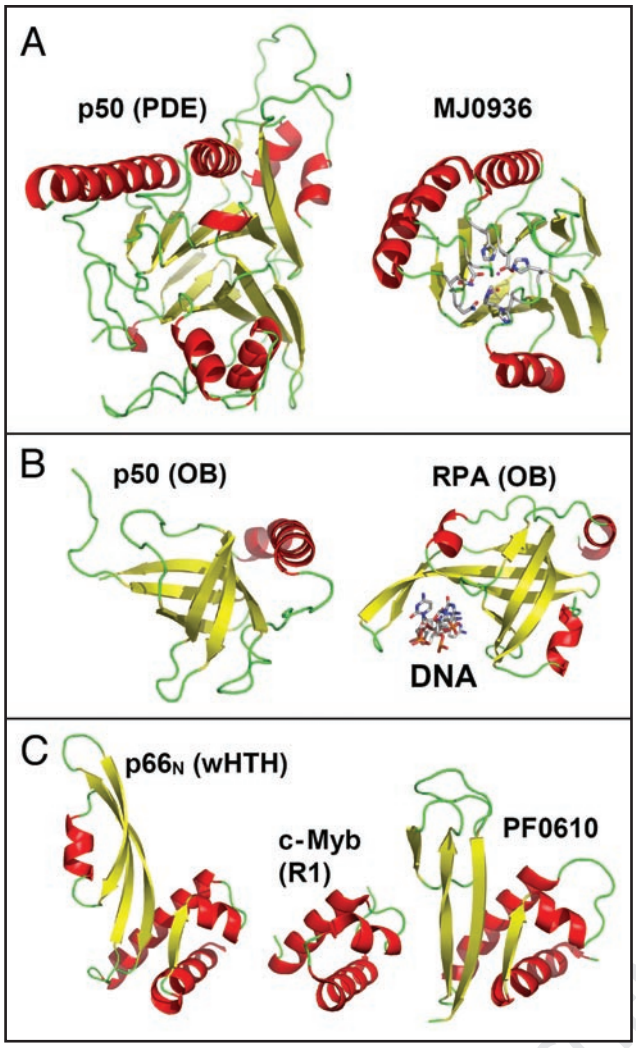

Figure 4. Comparison of p50•p66$_N$ domains with known structures. Comparisons of (A) p50 PDE domain with MJ0936[27] (PDB code 1s3l), (B) p50 OB domain with OB domain (aa 302–420) of DNA bound RPA70 subunit[29] (PDB code 1gmc), and (C) p66 wHTH domain residues 2–106 with c-Myb R1[33] (PDB code 1h88) and PF0610[34] (PDB code 2gmg). The folding of domains is shown by cartoons. The MJ0936 active site residues and ions are shown in sticks and balls, respectively. ssDNA bound to the OB domain of RPA70 is shown in sticks. All panels were prepared with PyMol software (Delano Scientific).

properties of these mutants parallel the properties of strains with the complete deletion of the third subunit,[13,15] although all phenotypes are somewhat milder as expected from the partial defect of the interaction between the subunits. The *pol31-WRRGW* mutation effects are stronger than the effects of the *pol31-RRG*, consistent with the cumulative effect of two types of amino acid changes. Thus, the mutations seem to specifically affect the interaction between the second and third subunits, which is consistent with the structure-predicted functional consequences of the amino acid changes.

Interestingly, in spite of the formation of a very tight complex between the p50 and p66 subunits, and the functional importance of p50•p66 interaction, the residues from their dimerization interface were mapped to the least conserved regions (Fig. 5). Nevertheless, despite variations in the amino acid sequence compositions in interacting areas, the secondary structure elements of p50 and p66$_N$ derived from the crystal structure coincide well with their counterparts from other species predicted by Phyre software,[37] including the orthologs from the fission and budding yeasts. Thus, p50•p66$_N$ complex and its orthologs are likely to share the same overall fold, which means a significant number of direct inter-subunit backbone to backbone hydrogen bonds (Table S1). The structure of p50•p66$_N$ provides a plausible explanation for the previously reported results of two-hybrid assays of interactions between Cdc1 deletion mutants an Cdc27, the fission yeast orthologs of p50 and p66, and highlights the importance of inter-subunit backbone interactions.[36] In particular, deletion of C-terminal 20 residues of Cdc1 abolished its interaction with Cdc27 in yeast two-hybrid assays.[36] Indeed, these residues of Cdc1 correspond to a β_{18} and the N-terminal tail of p50 (Fig. 5), which provide seven out of the 12 inter-subunit backbone interactions with p66 (Table S1). Similarly, the deletion of up to 100 N-terminal residues of the bakers yeast third subunit Pol32 abolished the interaction between the Pol31 and Pol 32 subunits.[11]

Discussion

The crystal structure of Pol δ p50•p66$_N$ reveals, for the first time, the three-dimensional organization of the essential second subunits of the B family DNA polymerases and provides a platform for their structure-based rational genetic and functional studies. Three distinct folds comprising the p50•p66$_N$ complex represent novel variations of the wHTH, OB and PDE domains. The PDE domain is in the center of the molecule and serves for a variety of protein-protein interactions. The wHTH and OB domains flanking the PDE domain represent the folds found in dsDNA- and ssDNA-binding proteins, respectively. These two domains have a potential to augment the substrate binding by the catalytic subunit of Pol δ. For example, the ssDNA-binding affinity of a single OB domain of RPA is low; however, the addition of a second OB domain linked to a first OB domain increases the ssDNA-binding affinity of RPA approximately 100-fold.[39] The wHTH and OB domains might also perform a regulatory role by detecting certain types of DNA distortions due to damage and switching the Pol δ from processive replication to transactions required for DNA repair.

The second subunits of B family DNA polymerases, including both archaeal and eukaryotic polymerases, are three-domain proteins: an approximately one third of the protein consists of the distinct "unknown" domain in the N-terminal end, and the remaining part is comprised of the OB and PDE domains (Fig. 8). However, the second subunits of eukaryotic Pol δ appear to be an exception because they do not contain such special domain at N-terminal part. Thus it is not surprising that previous bioinformatics studies of B family DNA polymerase second subunits were able to predict the presence of OB and PDE domains but have not detected the similarities in their N-terminal sequences.[25,26] The structure of p50•p66$_N$ presented here revealed tight complex of OB and PDE-type domains in p50, and the wHTH-type domain in p66$_N$. The high degree of structural complementarity between the PDE and wHTH domains pointed to the possibility that the "unknown" domain in N-terminal parts of the second subunits of B family DNA polymerases, except for Pol δ, might be the wHTH-type domain and interacts with the PDE domain similarly to what observed in structure of p50•p66$_N$. Indeed, predictions of the secondary structure elements in N-terminal parts of human pola2 and pole2, their yeast orthologs, as well as second

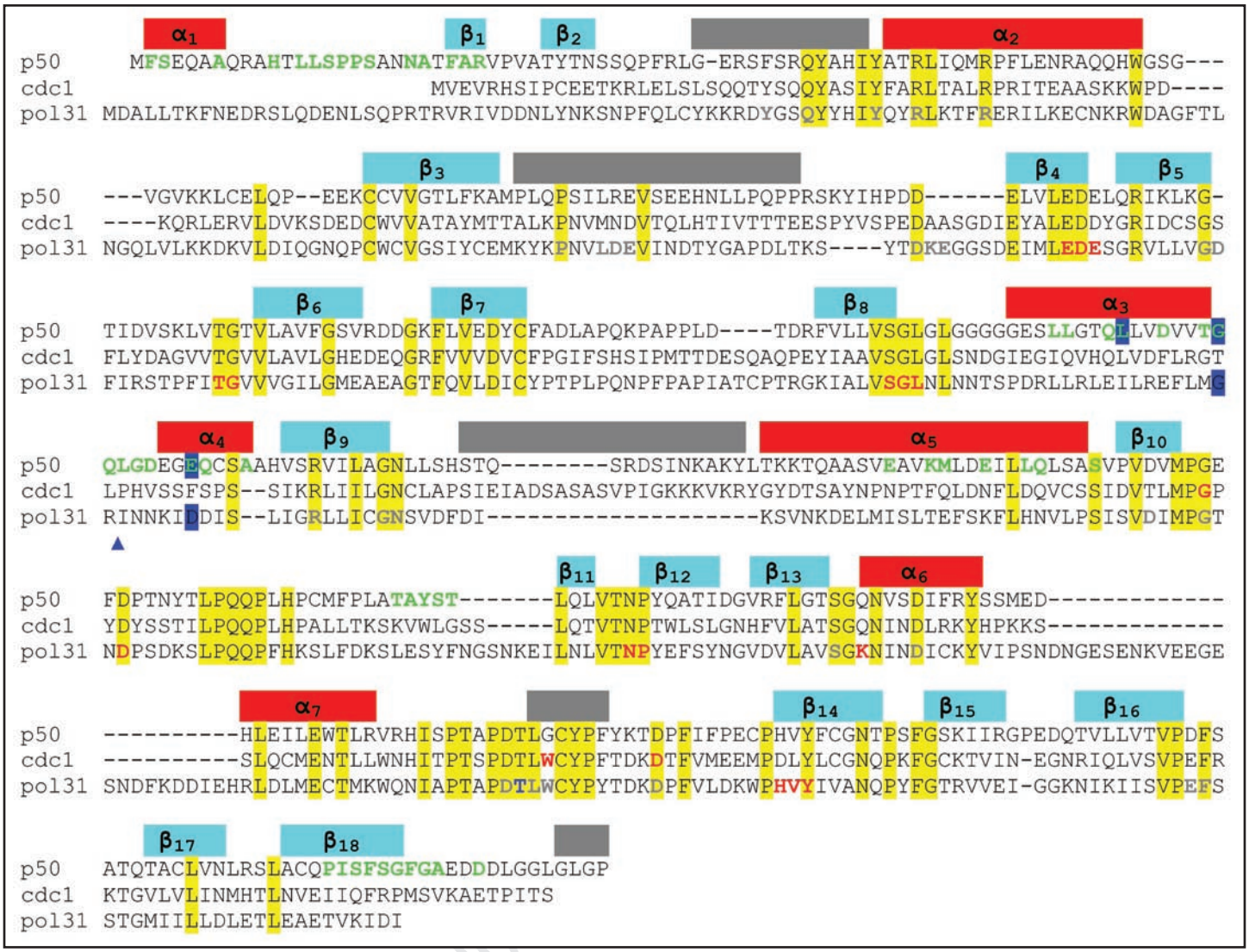

Figure 5. Alignment of human p50, *S. pombe* Cdc1 and *S. cerevisiae* Pol31 sequences. The red, cyan and grey bars correspond to positions of α-helices, β-strands and disordered regions in p50. The identical residues are highlighted in yellow. The p50 amino acid residues involved in interactions with p66$_N$ are in bold and green. The amino acid residues mutated in Cdc1,[36] and Pol31,[35] are in bold and red (temperature sensitive), blue (cold sensitive) and grey (no temperature sensitivity). The p50 and Pol31 residues mutated in current study are highlighted in blue and the position of RRG insertion in Pol31 is indicated by triangle.

subunits of archaeal Pol II (e.g., *Methanothermobacter thermautotrophicus* DNA polymerase II) using Phyre software[37] revealed the presence of αααββ motifs usually found in wHTH domains. This allows us to hypothesize that p50•p66$_N$ complex, instead of p50 alone, should be considered as the structural counterpart of the second subunits of B family DNA polymerases. This hypothesis is further supported by the statistically significant similarity between multiple alignments of N-terminal parts of p66 sequences and N-terminal parts of the Pol α second subunit (Fig. S1A). Taking into account that N-terminal parts of the Pol α, Pol ε and archaeal Pol II second subunits appear to be homologous (Fig. S1B–D), it seems likely that the coding sequence for the N-terminal region of p66 was derived from the second subunit genes subsequent to the divergence of Pol α and Pol δ from their last common ancestor (Fig. 8). To test these predictions, further structural studies of at least one of the second subunits of B family DNA polymerases, in addition to Pol δ, will be required.

Experimental Procedures

Structure determination and refinement. Preparations of p50•p66$_N$ samples, crystallization, cryoprotection of crystals, X-ray diffraction data collection and processing are described elsewhere.[40] Heavy-atom derivatives were obtained by soaking of crystals in reservoir solutions with dissolved Hg-, Pt-, Au-, Pb- and Sm-containing compounds selected from the Hampton Research heavy atom screens. Cryoprotection of the derivative crystals and diffraction data collection were performed as described for native crystals.[40] Complete diffraction data sets were collected and processed using 37 cryoprotected derivative crystals. The best two derivatives were obtained by soaking for 10 and 30 minutes in solutions containing approximately 0.1 mM of $HgCl_2$ and 0.5 mM of $K_2Pt(NO_2)_4$, respectively. Hg and Pt derivative crystals diffracted up to 4.0 Å and 3.3 Å resolutions, respectively. The heavy atom positions were obtained from 5 Å difference Patterson maps using CNS software.[41] The Pt derivative

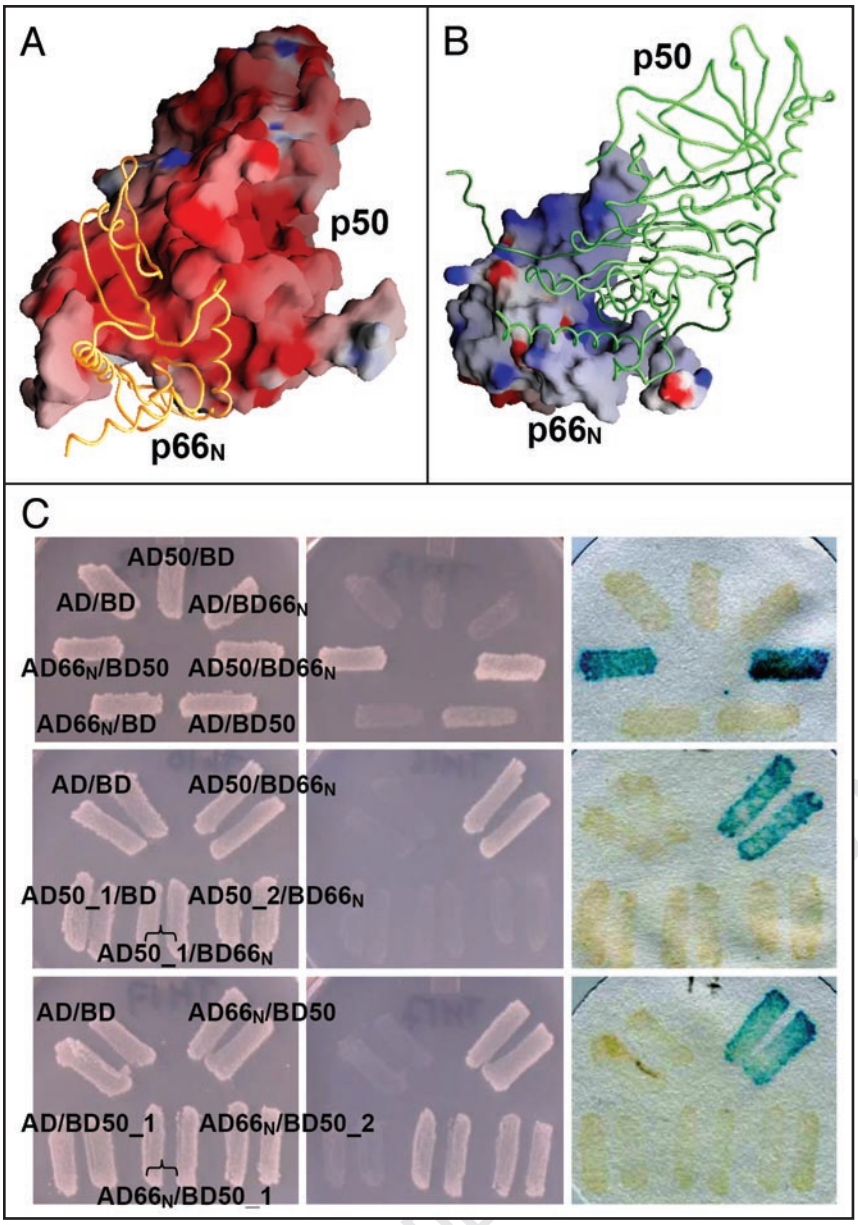

Figure 6. p50•p66$_N$ interactions. Electrostatic surface potential of (A) p66$_N$-bound p50, and (B) p50-bound p66$_N$; positively and negatively charged areas are colored blue and red, respectively. To assess the location of binding sites, the bound p66$_N$ (orange) and p50 (green) are shown as tubes. The panels a and b were prepared with GRASP software.[48] (C) Yeast two-hybrid analysis of the interaction of human p50, its variants with double amino acid changes and p66$_N$. Left column represents the growth of yeast patches on leucine and tryptophan omission plates selective for the two plasmids. Central column plates are without leucine, tryptophan and histidine, but supplemented with aminotriazole. Growth is allowed when two proteins interact. Third column represents paper filters with yeast patches stained for β-galactosidase activity (development of blue color when proteins interact). 50_1 and 50_2 are mutations encoding for changes Leu217Trp + Gly224Trp and Gly224Trp + Glu231Trp, respectively. Mutations encoding single amino acids substitutions Leu217Trp, Gly224Trp or Glu231Trp had effects similar to double mutations. AD stands for the pGAD424 plasmid and BD stands for pGBT9 plasmids. AD50 and AD66$_N$ stand for the constructs with pGAD424 with insert of the *POLD2* and the *POLD3*$_{1-432}$ genes respectively. Similarly, BD50 and BD66$_N$ stand for respective inserts into pGBT9.

as well as all other successfully derivatized but poorly diffracting crystals shared the same four major heavy atom sites. The only exception was an Hg derivative with two different major heavy atom sites.

The structure of p50•p66$_N$ was solved by MIR method using Pt and Hg derivatives. Programs from the CCP4 program suite[42] were used for all calculations related to phase determination and density modification. Single isomorphous replacement phases, computed with Hg derivative using the program MLPHARE at 5 Å resolution, were used for the verification of metal sites in Pt derivative with difference Fourier maps, and vise versa. R_{iso} for both crystals were too high beyond the 4.2 Å resolution, limiting the use of higher resolution data (Table 1). That is why the phases for both derivatives were refined at 4.2 Å resolution until convergence with the program MLPHARE, giving an overall figure of merit of 0.361 (Table 1). The electron density maps calculated with MIR phases were not interpretable. Phase extension from 5 Å to 4.0 Å resolution combined with solvent flattening and histogram matching with the program DM removed some noise, but the maps still were difficult to interpret. Our preliminary characterization of p50•p66$_N$ crystals pointed to the presence of four molecules in an asymmetric unit. DM calculations with the addition of four-fold density averaging and automatic mask calculation produced an interpretable map with clear boundaries of each molecule (Fig. S2A). The non-crystallographic symmetry (ncs) matrices were calculated using the positions of heavy atoms. We assumed that each complex contains one Pt site, and one Hg site is located between two complexes related by non-crystallographic two-fold symmetry. A total of 420 main chain residues belonging to 11 separate peptides were built into the averaged DM map using the program TURBO-FRODO. The phases and mask calculated from the initial model were used for DM with phase extension from 5 Å to 3 Å. The resulting map was of high quality and allowed the correction and addition of main chain residues, as well as the identification and addition of bulky side chains (Fig. S2B). The model was multiplied using ncs matrices, subjected to rigid-body refinement with CNS software, and new ncs matrices and a mask were obtained from the refined structure. This model building procedure was repeated 11 times and resulted in R_{free} of 44%. The correct assignment of residues was confirmed with positions of Pt and Hg atoms bound to p50 Cys81 and Cys314, respectively, and the formation of disulfide bond between p50 Cys438 and Cys447. Further refinement was performed with CNS software[41] using conjugate gradient minimization, simulated annealing and temperature factor refinement protocols. Finally, 61 water molecules were added with the water-pick protocol of CNS. Application of zonal scaling,[43] bulk solvent correction, as well as ncs and

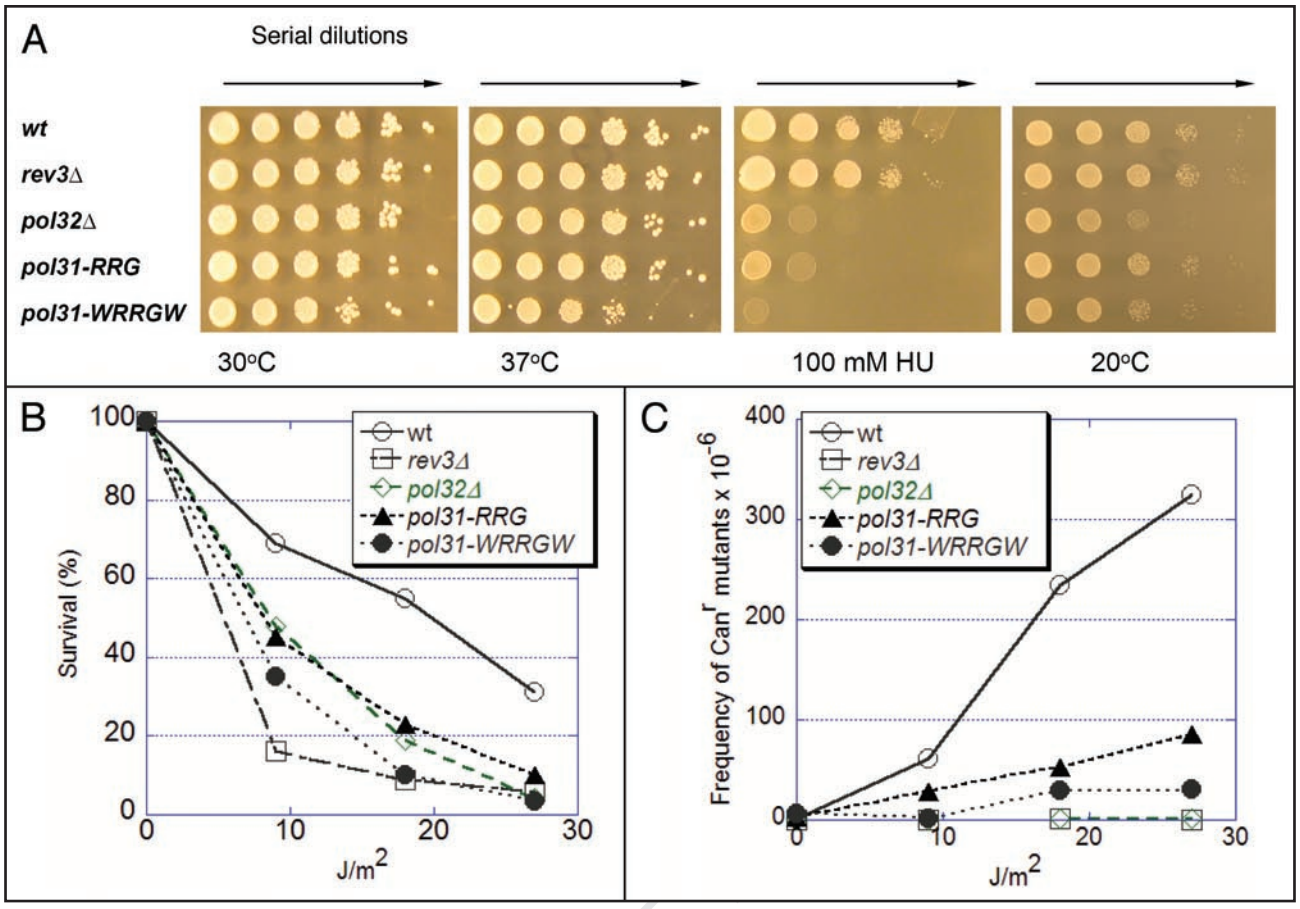

Figure 7. Phenotypes of yeast strain with mutations in the *POL31*. The results for strains with deletions of the *REV3* and *POL32* genes are presented for comparison. (A) Studies of sensitivity to elevated temperature, hydroxyurea (HU) and lowered temperature. Fresh stationary cultures were serially 10-fold diluted in microtiter plates and plated with multi-prong device onto YPD plates, which were incubated at different temperatures or onto YPD with 100 mM HU. (B) Survival after UV -irradiation. (C) UV-Mutagenesis. Appropriate dilutions were plated on selective and synthetic complete plates and UV irradiated. Colonies were counted after 3 days of incubation at 30°C.

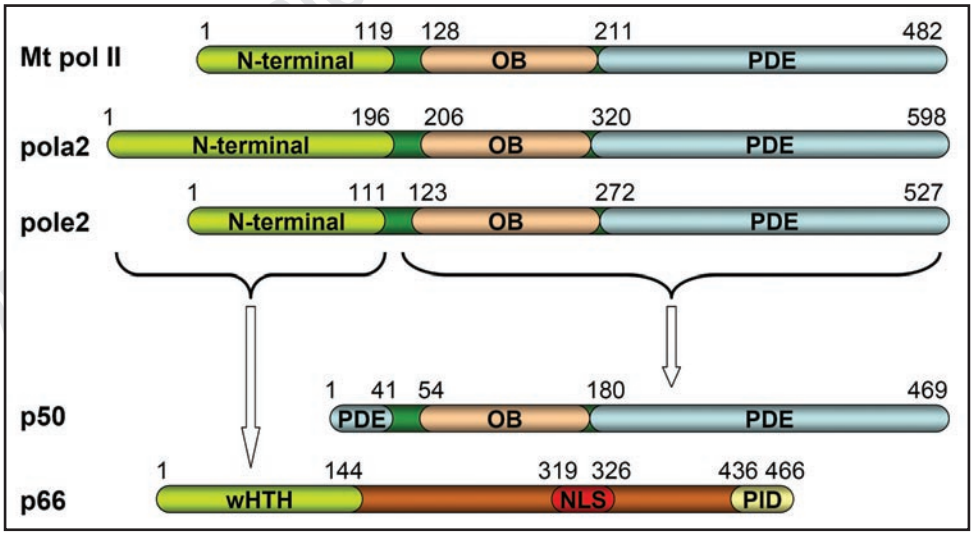

Figure 8. Domains of the B family DNA polymerase second subunits. The domain boundaries in the second subunits of archaeal *Methanothermobacter thermautotrophicus* DNA polymerase II (Mt pol II), human Pol α (pola2), human Pol ε (pole2) and Pol δ (p50), and third subunit of human Pol δ (p66).

Table 1 **Data collection, phasing and refinement statistics**

	native	Hg derivative	Pt derivative
Data collection			
Space group	$P2_1$	$P2_1$	$P2_1$
Cell dimensions			
a, *b*, *c* (Å)	95.13, 248.54, 103.46	95.04, 248.99, 102.34	94.86, 248.76, 103.74
β (°)	106.94	106.59	107.17
Resolution (Å)*	40–3 (3.11–3)	40–4.2 (4.35–4.2)	40–4.2 (4.35–4.2)
Unique reflections	85645	29275	301365
R_{merge}	0.079 (0.430)	0.148 (0.289)	0.065 (0.126)
$I/\sigma(I)$	21.2 (3.8)	10.3 (5.5)	29.2 (14.9)
Completeness (%)	93.0 (90.4)	88.4 (86.2)	93.9 (91.1)
Redundancy	2.6 (2.1)	2.1 (2.0)	
MIR phasing			
Resolution (Å)	25–4	25–4.2	25–4.2
R_{iso}*		0.245	0.144
R_{cullis}* (acentric/centric)*		0.84/0.79	0.87/0.89
Phasing power* (acentric/centric)		1.22/0.96	0.96/0.87
Combined figure of merit*	0.361		
Refinement			
Resolution (Å)	30–3		
No. reflections	83574		
R_{work}/R_{free}	0.257/0.281		
No. atoms			
Protein	17096		
Water	61		
B-factors (Å^2)	60.4		
R.m.s. deviations			
Bond lengths (Å)	0.011		
Bond angles (°)	1.6		
Ramachandran plot			
Favored (%)	80.6		
Allowed (%)	18.9		
Generous (%)	0.6		
Disallowed (%)	0		

*Number of shells is 10 for each structure. Values in parentheses are for highest-resolution shell. $R_{iso} = \Sigma_{hkl}|F_{PH} - F_P|/\Sigma_{hkl}F_P$; $R_{cullis} = \Sigma_{hkl}|F_{PH} - |F_P \pm F_H(calc)||/\Sigma_{hkl}F_{PH} - F_P|$; Phasing power $= \langle F_H(calc)\rangle/\langle F_{PH} - |F_P \pm F_H(calc)|\rangle$, where F_{PH} and F_P are the structure-factor amplitudes of the derivative and native crystals, respectively, and $F_H(calc)$ is the calculated heavy atoms structure-factor amplitude.

main chain restraints, improved the quality of maps and resulted in excellent geometry of molecule with Ramachandran plot showing 80.6%, 18.9% and 0.6% of residues in the core, additionally allowed and generously allowed regions, respectively (Table 1). The density was not observed for the total of 56 residues, including p50 amino acid residues 42–53, 102–122, 251–253, 382–387 and 466–469.

Yeast two-hybrid studies. *POLD2* and its alleles with site-directed mutations and *POLD3*$_{1-432}$ were cloned into the *Eco*RI-*Sal*I sites of the Matchmaker pGAD424 and pGBT9 vectors from Clontech. We have used primers (AGCCA**GAATTC**ATGTTTTCTGAGCAGGCTGCC), (AGTAGAT**GTCGAC**TCAGGGTCCCAGCCC) for cloning and (ACATACACCCAGATGACGAGC) for sequencing of the mutation region of the *POLD2*. The primers for the *POLD3* were: (AGCTA**GAATTC**GCGGACCAGCTTTATCTGG), (AGAGAT**GTCGAC**TTAAGCTCTAGGGACGGCAG) for cloning and (AGTGCTATACAATGTGCAGC) for sequencing. Host strain CG1945 was transformed to Leu$^+$ Trp$^+$ by all combinations of empty vectors and fusion plasmids. Four independent transformants from each variant were tested on histidine-less medium and for blue color formation filter assay as suggested by the manufacturer.

Functional studies. Yeast integrative plasmid pRS306 was constructed by cloning the PCR fragment corresponding to the C-terminal two thirds of the *POL31* gene and a terminator region (nucleotides 160–1664 counting from ATG) with artificial *Hind*III-*Xho*I flanking sites using ACTAAGCTTACCACATTTATCAATACCG and TATCTCGAGTTCATCAGATTTGTGCTT. Nucleotide changes to encode the mutations discussed above were generated by Stratagene Quick mutagenesis kit. Plasmids were linearized by *Mfe*I and used for the transformation of yeast strains to the Ura$^+$

phenotype and the subsequent excision of the *URA3* marker for the analysis of phenotypes and mutation rates as done previously.[44,45] The analysis was done for two strains with similar results. One strain was yo699 (*MAT*a *CAN1 his7-2 leu2-Δ::kanMX ura3-Δ trp1-289 ade2-1 lys2-ΔGG2899-2900*, ψ^- derivative of the strain Δl(-2)l-7B-YUNI300[38] and the other 9A-YPOM206 (*MATα CAN1 his7-2 leu2Δ -0 ura3-Δ trp1-289 ade2-1 lys2-ΔGG2899-2900*) derived from a diploid resulting from cross of yo600 to BY4742.

Protein sequences analysis. The non-redundant (nr) database of National Center for Biotechnology Information (http://ncbi.nlm.nih.gov/) was searched using the PSI-BLASTP program.[46] Iterative sequence similarity searches were performed using single sequences of human p66$_N$, Pol α and Pol ε second subunits (NP_006582, NP_002680, and NP_002683) and archaeal Pol II second subunit from *Methanothermobacter thermautotrophicus* (NP_276521) as queries with default parameters.[46] Each search was run for a minimum of ten iterations or to convergence. Multiple alignments produced by PSI-BLAST were compared using the COMPASS program with default parameters.[47]

Accession Numbers

Coordinates and structure factors have been deposited in the RCSB Protein Data Bank with accession code 3E0J.

Acknowledgements

We thank P.V. Shcherbakova, D. Bugreev and E.I. Stepchenkova for critical reading the manuscript, K. Spitler and E.I. Stepchenkova for the construction of the pRS306-POL31C vector, D. Bugreev for discussion of protein expression and purification approaches, J. Lovelace and G.E. Borgstahl for the maintenance and management of the Eppley Institute's X-ray Crystallography facility. The primers were synthesized in the Eppley Institute's Molecular Biology Core facility. Both facilities are supported by the Cancer Center Support Grant P30CA036727. This work is funded in part by UNMC Eppley Cancer Center Pilot Projects LB595 to Y.I.P. and T.H.T., in part by the 2008 NE DHHS grant LB506 to Y.I.P., and in part by the NIH grants GM74252 and GM74840 to D.G.V. I.B.R. and E.V.K. were supported by the Intramural Research Program of the National Library of Medicine at National Institutes of Health/DHHS.

Note

Supplementary materials can be found at:

www.landesbioscience.com/supplement/BaranovskiyCC7-19-Sup.pdf

References

1. Downey KM, Tan CK, So AG. DNA polymerase delta: a second eukaryotic DNA replicase. Bioessays 1990; 12:231-6.
2. Garg P, Burgers PM. DNA polymerases that propagate the eukaryotic DNA replication fork. Crit Rev Biochem Mol Biol 2005; 40:115-28.
3. Waga S, Stillman B. Anatomy of a DNA replication fork revealed by reconstitution of SV40 DNA replication in vitro. Nature 1994; 369:207-12.
4. Kesti T, Flick K, Keranen S, Syvaoja JE, Wittenberg C. DNA polymerase ε catalytic domains are dispensable for DNA replication, DNA repair and cell viability. Mol Cell 1999; 3:679-85.
5. Pavlov YI, Shcherbakova PV, Rogozin IB. Roles of DNA polymerases in replication, repair and recombination in eukaryotes. Int Rev Cytol 2006; 255:41-132.
6. Franklin MC, Wang J, Steitz TA. Structure of the replicating complex of a pol α family DNA polymerase. Cell 2001; 105:657-67.
7. McCulloch SD, Kunkel TA. The fidelity of DNA synthesis by eukaryotic replicative and translesion synthesis polymerases. Cell Res 2008; 18:148-61.
8. Asturias FJ, Cheung IK, Sabouri N, Chilkova O, Wepplo D, Johansson E. Structure of *Saccharomyces cerevisiae* DNA polymerase epsilon by cryo-electron microscopy. Nat Struct Mol Biol 2006; 13:35-43.
9. Sanchez Garcia J, Ciufo LF, Yang X, Kearsey SE, MacNeill SA. The C-terminal zinc finger of the catalytic subunit of DNA polymerase δ is responsible for direct interaction with the B-subunit. Nucleic Acids Res 2004; 32:3005-16.
10. Gray FC, Pohler JR, Warbrick E, MacNeill SA. Mapping and mutation of the conserved DNA polymerase interaction motif (DPIM) located in the C-terminal domain of fission yeast DNA polymerase δ subunit Cdc27. BMC Mol Biol 2004; 5:21.
11. Johansson E, Garg P, Burgers PM. The Pol32 subunit of DNA polymerase δ contains separable domains for processive replication and proliferating cell nuclear antigen (PCNA) binding. J Biol Chem 2004; 279:1907-15.
12. Bermudez VP, MacNeill SA, Tappin I, Hurwitz J. The influence of the Cdc27 subunit on the properties of the *Schizosaccharomyces pombe* DNA polymerase delta. J Biol Chem 2002; 277:36853-62.
13. Gerik KJ, Li X, Pautz A, Burgers PM. Characterization of the two small subunits of *Saccharomyces cerevisiae* DNA polymerase delta. J Biol Chem 1998; 273:19747-55.
14. Huang ME, Le Douarin B, Henry C, Galibert F. The *Saccharomyces cerevisiae* protein YJR043C (Pol32) interacts with the catalytic subunit of DNA polymerase alpha and is required for cell cycle progression in G_2/M. Mol Gen Genet 1999; 260:541-50.
15. Huang ME, de Calignon A, Nicolas A, Galibert F. Pol32, a subunit of the *Saccharomyces cerevisiae* DNA polymerase δ, defines a link between DNA replication and the mutagenic bypass repair pathway. Curr Genet 2000; 38:178-87.
16. Lawrence CW. Cellular roles of DNA polymerase ζ and Rev1 protein. DNA Repair (Amst) 2002; 1:425-35.
17. Lydeard JR, Jain S, Yamaguchi M, Haber JE. Break-induced replication and telomerase-independent telomere maintenance require Pol32. Nature 2007; 448:820-3.
18. Shikata K, Ohta S, Yamada K, Obuse C, Yoshikawa H, Tsurimoto T. The human homologue of fission Yeast cdc27, p66, is a component of active human DNA polymerase delta. J Biochem 2001; 129:699-708.
19. Podust VN, Chang LS, Ott R, Dianov GL, Fanning E. Reconstitution of human DNA polymerase delta using recombinant baculoviruses: the p12 subunit potentiates DNA polymerizing activity of the four-subunit enzyme. J Biol Chem 2002; 277:3894-901.
20. Li H, Xie B, Zhou Y, Rahmeh A, Trusa S, Zhang S, et al. Functional roles of p12, the fourth subunit of human DNA polymerase delta. J Biol Chem 2006; 281:14748-55.
21. Li H, Xie B, Rahmeh A, Zhou Y, Lee MY. Direct interaction of p21 with p50, the small subunit of human DNA polymerase delta. Cell Cycle 2006; 5:428-36.
22. Pohler JR, Otterlei M, Warbrick E. An in vivo analysis of the localisation and interactions of human p66 DNA polymerase delta subunit. BMC Mol Biol 2005; 6:17.
23. Huang ME, Rio AG, Galibert MD, Galibert F. Pol32, a subunit of *Saccharomyces cerevisiae* DNA polymerase delta, suppresses genomic deletions and is involved in the mutagenic bypass pathway. Genetics 2002; 160:1409-22.
24. Gibbs PE, McDonald J, Woodgate R, Lawrence CW. The relative roles in vivo of *Saccharomyces cerevisiae* Pol eta, Pol zeta, Rev1 protein and Pol32 in the bypass and mutation induction of an abasic site, T-T (6–4) photoadduct and T-T cis-syn cyclobutane dimer. Genetics 2005; 169:575-82.
25. Aravind L, Koonin EV. Phosphoesterase domains associated with DNA polymerases of diverse origins. Nucleic Acids Res 1998; 26:3746-52.
26. Makiniemi M, Pospiech H, Kilpelainen S, Jokela M, Vihinen M, Syvaoja JE. A novel family of DNA-polymerase-associated B subunits. Trends Biochem Sci 1999; 24:14-6.
27. Chen S, Yakunin AF, Kuznetsova E, Busso D, Pufan R, Proudfoot M, et al. Structural and functional characterization of a novel phosphodiesterase from *Methanococcus jannaschii*. J Biol Chem 2004; 279:31854-62.
28. Murzin AG. OB (oligonucleotide/oligosaccharide binding)-fold: common structural and functional solution for non-homologous sequences. EMBO J 1993; 12:861-7.
29. Bochkarev A, Pfuetzner RA, Edwards AM, Frappier L. Structure of the single-stranded-DNA-binding domain of replication protein A bound to DNA. Nature 1997; 385:176-81.
30. Bochkareva E, Kaustov L, Ayed A, Yi GS, Lu Y, Pineda-Lucena A, et al. Single-stranded DNA mimicry in the p53 transactivation domain interaction with replication protein A. Proc Natl Acad Sci USA 2005; 102:15412-7.
31. Janin J. Specific versus non-specific contacts in protein crystals. Nat Struct Biol 1997; 4:973-4.
32. Krissinel E, Henrick K. Secondary-structure matching (SSM), a new tool for fast protein structure alignment in three dimensions. Acta Crystallogr D Biol Crystallogr 2004; 60:2256-68.
33. Tahirov TH, Sato K, Ichikawa-Iwata E, Sasaki M, Inoue-Bungo T, Shiina M, et al. Mechanism of c-Myb-C/EBP beta cooperation from separated sites on a promoter. Cell 2002; 108:57-70.
34. Wang X, Lee HS, Sugar FJ, Jenney FE Jr, Adams MW, Prestegard JH. PF0610, a novel winged helix-turn-helix variant possessing a rubredoxin-like Zn ribbon motif from the hyperthermophilic archaeon, Pyrococcus furiosus. Biochemistry 2007; 46:752-61.
35. Vijeh Motlagh ND, Seki M, Branzei D, Enomoto T. Mgs1 and Rad18/Rad5/Mms2 are required for survival of *Saccharomyces cerevisiae* mutants with novel temperature/cold sensitive alleles of the DNA polymerase δ subunit, Pol31. DNA Repair (Amst) 2006; 5:1459-74.
36. MacNeill SA, Moreno S, Reynolds N, Nurse P, Fantes PA. The fission yeast Cdc1 protein, a homologue of the small subunit of DNA polymerase delta, binds to Pol3 and Cdc27. EMBO J 1996; 15:4613-28.

37. Bennett-Lovsey RM, Herbert AD, Sternberg MJ, Kelley LA. Exploring the extremes of sequence/structure space with ensemble fold recognition in the program Phyre. Proteins 2008; 70:611-25.
38. Pavlov YI, Newlon CS, Kunkel TA. Yeast origins establish a strand bias for replicational mutagenesis. Mol Cell 2002; 10:207-13.
39. Arunkumar AI, Stauffer ME, Bochkareva E, Bochkarev A, Chazin WJ. Independent and coordinated functions of replication protein A tandem high affinity single-stranded DNA binding domains. J Biol Chem 2003; 278:41077-82.
40. Baranovskiy AG, Babayeva ND, Tahirov TH. Crystallization and preliminary crystallographic analysis of the complex of second and third regulatory subunits of human Pol delta. Acta Crystallogr F Struct Biol Cryst Commun 2008; in press.
41. Brunger AT, Adams PD, Clore GM, DeLano WL, Gros P, Grosse-Kunstleve RW, et al. Crystallography & NMR system: A new software suite for macromolecular structure determination. Acta Crystallogr D Biol Crystallogr 1998; 54:905-21.
42. Collaborative Computational Project Number 4. The CCP4 suite: programs for protein crystallography. Acta Crystallogr D Biol Crystallogr 1994; 50:760-3.
43. Vassylyev DG, Vassylyeva MN, Perederina A, Tahirov TH, Artsimovitch I. Structural basis for transcription elongation by bacterial RNA polymerase. Nature 2007; 448:157-62.
44. Pavlov YI, Shcherbakova PV, Kunkel TA. In vivo consequences of putative active site missense mutations in yeast replicative DNA polymerases α, ε, δ and ζ. Genetics 2001; 159:47-64.
45. Kozmin SG, Pavlov YI, Kunkel TA, Sage E. Roles of *Saccharomyces cerevisiae* DNA polymerases Pol η and Pol ζ in response to irradiation by simulated sunlight. Nucleic Acids Res 2003; 31:4541-52.
46. Altschul SF, Madden TL, Schaffer AA, Zhang J, Zhang Z, Miller W, et al. Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res 1997; 25:3389-402.
47. Sadreyev RI, Grishin NV. Accurate statistical model of comparison between multiple sequence alignments. Nucleic Acids Res 2008; 36:2240-8.
48. Nicholls A, Sharp KA, Honig B. Protein folding and association: insights from the interfacial and thermodynamic properties of hydrocarbons. Proteins 1991; 11:281-96.