Gene, 166 (1995) 237–242

GENE 09281

Isolation and sequence determination of the cDNA encoding DNA polymerase δ from *Drosophila melanogaster*

(DNA replication; PCNA; cloning; PCR; antibody; sequence conservation)

Chuen-Sheue Chiang* and I. R. Lehman

Department of Biochemistry, Beckman Center, Stanford University, Stanford, CA 94305, USA

Received by R. Padmanabhan: 5 July 1995; Accepted: 10 July 1995; Received at publishers: 21 August 1995

SUMMARY

The cDNA encoding the catalytic subunit of *Drosophila melanogaster* (*Dm*) DNA polymerase δ (Polδ) was isolated by a combination of PCR amplification and cDNA library screening. The cDNA is 3457 nucleotides in length and contains an open reading frame (ORF) that encodes a protein of 1092 amino acids (124 799 Da). The ORF contains the sequence that was determined for a peptide from the purified catalytic subunit of *Dm* Polδ. Polyclonal antibodies raised against *Dm* Polδ specifically recognize a protein of the expected size when the cDNA is expressed in either *Escherichia coli* or insect cells. Comparison of the deduced aa sequence with other Polδ sequences demonstrates that Polδ is one of the most highly conserved of the DNA polymerases.

INTRODUCTION

Three DNA polymerases, DNA polymerase α, δ and ε (Polα, δ and ε) are required to replicate eukaryotic chromosomal DNA (for review, see Wang, 1991; Stillman, 1994). Polδ and ε also appear to be involved in DNA repair (Nishida et al., 1988; Budd and Campbell, 1995; Aboussekhra et al., 1995). Recent studies that link mutations of Polδ to the replication error (RER) phenotype of some colorectal tumors are in agreement with this idea (da Costa et al., 1995).

Polδ is highly conserved both structurally and functionally. It consists of two subunits of approx. 120 and 50 kDa. The 120-kDa subunit has the polymerase and 3′-5′ exonuclease activities associated with the two subunit enzyme. The function of the 50-kDa subunit is unknown. Deoxynucleotide polymerization by Polδ is distributive; however, it becomes highly processive in the presence of the accessory protein, proliferating cell nuclear antigen (PCNA). The strong functional conservation of Polδ is supported by the finding that calf thymus PCNA stimulates the processivity of Polδ from *S. cerevisiae* and *S. cerevisiae* PCNA stimulates the processivity of the calf thymus Polδ (Bauer and Burgers, 1988). Polδ from *S. cerevisiae* can also substitute for human Polδ in SV40 replication in vitro (Tsurimoto et al., 1990).

We have been using *Drosophila melanogaster* (*Dm*) embryos as a model for the study of eukaryotic DNA replication and repair (Lehman and Kaguni, 1989). As part of this study, we purified Polδ from 0–2-h embryos to near homogeneity and demonstrated that it consists of only a single 120-kDa polypeptide, that is unresponsive

Correspondence to: Dr. I.R. Lehman, Department of Biochemistry, Beckman Center, Stanford University, Stanford, CA 94305, USA. Tel. (1-415) 723-6164; Fax (1-415) 723-6783.

* Present address: Department of Pathology, Stanford University, Stanford, CA 94305-5324, USA. Tel (1-415) 725-4908.

Abbreviations: aa, amino acid(s); Ab, antibody(ies); bp, base pair(s); cDNA, DNA complementary to RNA; *Dm*, *Drosophila melanogaster*; Exo, exonuclease; HSV-1, herpes simplex virus type 1; kb, kilobase(s) or 1000 bp; nt, nucleotide(s); ORF, open reading frame; PCNA, proliferating cell nuclear antigen; PCR, polymerase chain reaction; Pol, DNA polymerase; *Pol*, gene (DNA) encoding Pol; *S.*, *Saccharomyces*; SDS, sodium dodecyl sulfate; SV40, simian virus 40; *UTR*, untranslated region(s); Zf, zinc finger(s).

SSDI 0378-1119(95)00567-6

to PCNA (Chiang et al., 1993). As a prelude to analyzing the interaction of the *Dm* Polδ with PCNA and determining the function of the 50-kDa subunit, we have cloned the *Polδ* gene for the 120-kDa subunit and determined the nucleotide sequence of the full-length cDNA.

EXPERIMENTAL AND DISCUSSION

(a) Isolation and sequence determination of the *Polδ* cDNA encoding the 120-kDa subunit of *Dm* Polδ

By using PCR amplication of a *Dm* ovarian cDNA library constructed in λgt11 (generous gift from Dr. Laura Kafayen) with oligodeoxynucleotide primers based on the aa sequences MAHNLCYT and VV/IYGDTD in Domains II and I of Polδ, respectively, one positive PCR product of 419 bp was isolated. Sequence analysis demonstrated that it contained an ORF extensively homologous to human and bovine Polδ (67% and 64% aa identity, respectively). This ORF could be aligned with aa 615 to 750 of human Polδ and contained conserved domains VI and III in the correct order and spatial separation (Fig. 1). The ORF had only 10% identity with the catalytic subunit of the *Dm* Polα.

The partial cDNA was used to screen the same cDNA library. One of 23 positive clones that carried the longest insert was further analyzed and sequenced. The cDNA consisted of 3457 nt, including a 5′-*UTR* of 93 nt, an ORF of 1092 aa, a 3′-*UTR* of 68 nt and a poly(A) tail of 20 nt (Fig. 1). The ATG sequence at nt 94 was chosen as the start codon because (*i*) the only preceding ATG sequence was followed by several in-frame stop codons, (*ii*) the sequence surrounding the start codon was in agreement with the Cavener (1987) criteria for a translational start site for *Dm* mRNA and (*iii*) the first 5 deduced aa were identical to those of human, bovine and mouse Polδ. This ORF encodes a protein of 124 799 Da, which is consistent with the 120 kDa for the purified catalytic

Fig. 1. Nucleotide sequence of the *Drosophila melanogaster Polδ* cDNA (EMBL accession No. X88928) and the deduced aa sequence. The directly determined aa sequence is underlined. Also indicated are the oligodeoxynucleotide primers used for PCR (arrows), the sequence of the PCR product (dashed underlining) and the polyadenylation signal (double underlining).

subunit of *Dm* Polδ as judged by SDS-PAGE (Chiang et al., 1993). The 3′-*UTR* had a consensus AATAA polyadenylation signal 20 nt upstream from the polyadenylation site.

(b) The cDNA encodes *Dm* Polδ

Three lines of evidence demonstrate that the ORF encodes *Dm* Polδ: (*i*) The deduced aa sequence contains the aa sequence determined for a peptide generated from the purified 120-kDa *Dm* Polδ, YYADPISTLDFASLYPSI (Fig. 1), that is strictly conserved among members of the Polδ family (Fig. 2). (*ii*) A polyclonal Ab raised in chicken against *Dm* Polδ specifically recognized a 120-kDa protein when the cDNA was expressed in *E. coli*. When the cDNA was constructed in the opposite orientation and transfected into *E. coli*, no Ab-reactive protein could be detected (data not shown). (*iii*) Lysates prepared from Sf9 insect cells transfected with the cDNA-baculovirus construct showed increased DNA polymerase activity as compared with lysates from the insect cells alone or insect cells transfected with wild-type baculovirus, when they were assayed with poly(dA-dT), the preferred template for Polδ. The DNA polymerase activity in the lysates corresponded to a 120-kDa protein as judged by immunoblot analysis (data not shown).

(c) Genomic Southern and RNA transcript analysis

Genomic DNA from *Dm* adults was isolated, digested with *Hin*dIII, transferred to Nytran membrane, and then probed with the cDNA. A single band >12 kb was detected, indicating that there is only a single copy of the gene (data not shown).

When total RNA or mRNA isolated from *Dm* adults was probed with the cDNA, a single band of about 4 kb was detected (data not shown), indicating that only one major mRNA species corresponded to *Polδ*.

(d) Polδ is highly conserved

Polδ belongs to the family of α-like DNA polymerases (reviewed in Wang, 1991). The homology among the various Polδ species is much greater than for Polα. Fig. 2 shows an aa sequence comparison for the catalytic subunit of all known Polδ species. Although the N-terminal region shows the least homology, there are portions of the N-terminal region that are very conserved, e.g., NT-1 and NT-2 (Cullmann et al., 1993). Recent studies have suggested the involvement of this region in the interaction of Polδ with PCNA. Specifically, a truncated yeast Polδ without the N-terminal 220 aa expressed in *E. coli* did not interact with PCNA (Brown and Campbell, 1993), and studies with deletion mutants expressed in insect cells localized the region of human Polδ that binds PCNA to the first 182 aa (Zhang et al., 1995). Furthermore, a peptide containing aa 129 to 149 of human Polδ inhibited PCNA stimulation (Zhang et al,. 1995). Site-directed mutagenesis will be needed to define precisely the aa residues that mediate the interaction of Polδ with PCNA.

The central region of the protein is the most conserved among all DNA polymerases. It contains all six α-like homology boxes (Pol boxes I to VI with box I the most and box VI the least conserved) and three Exo boxes (Fig. 2). The Pol boxes constitute parts of the polymerase active site. Studies with Pol from HSV-1 and the *Bacillus* phage φ29 have demonstrated the necessity of box I for DNA polymerase activity (Dorsky and Crumpacker, 1990; Bernad et al., 1990) and substrate recognition (Marcy et al., 1990). Studies with human Polα have also shown that box I is involved in metal-specific catalysis (Copeland and Wang, 1993). Mutations in box II/III of HSV-1 Pol and box II of human Polα indicate that these boxes participate in dNTP (Larder et al., 1987; Tsurumi et al., 1987; Gibbs et al., 1988; Dong et al., 1993a) and primer binding (Dong et al., 1993b).

On the basis of aa sequence homology to DNA polymerase I (Pol I) of *E. coli*, it was suggested that three regions (Exo I, II and III) are related to the 3′-5′ Exo activity associated with eukaryotic DNA polymerases (Bernad et al., 1989). These regions also contain conserved aa sequences that are necessary for metal, and single-stranded DNA binding (Ollis et al., 1985; Derbyshire et al., 1991). Point mutations in the predicted Exo boxes of φ29 Pol resulted in a loss of 3′-5′ Exo activity (Bernad et al., 1989). Recently, a new Exo I box has been proposed, and subsequently confirmed by mutational analysis of *S. cerevisiae* Polδ (Simon et al., 1991). The original Exo I box was renamed Exo I′.

The C-terminal region of Polδ has five conserved regions, termed CT-1, CT-2, CT-3, ZnF1 and ZnF2 (Yang et al., 1992; Cullmann et al., 1993). The function of these regions is unknown. However, the C-terminal 227 aa of the HSV-1 Pol, including CT-2 and CT-3, are necessary for interaction with the UL 42 protein (Digard and Coen, 1990), a protein that is analogous in function to PCNA (Hernandez and Lehman, 1990). No significant homology downstream from the CT-3 box could be detected between the eukaryotic Pol and the HSV-1 Pol. ZnF1 and ZnF2 are two potential zinc finger (Zf) domains that were first identified in human Polα (Wong et al., 1988). Although the Zf regions could be good candidates for the interaction of Polδ with PCNA (Yang et al., 1992; Cullmann et al., 1993), the N-terminal region of the protein appears to mediate this interaction as noted above. The Zf regions might therefore be involved in the interaction of the 120-kDa and 50-kDa subunits. Site-directed mutagenesis in combination with biochemical studies is clearly necessary to elucidate the function of this region.

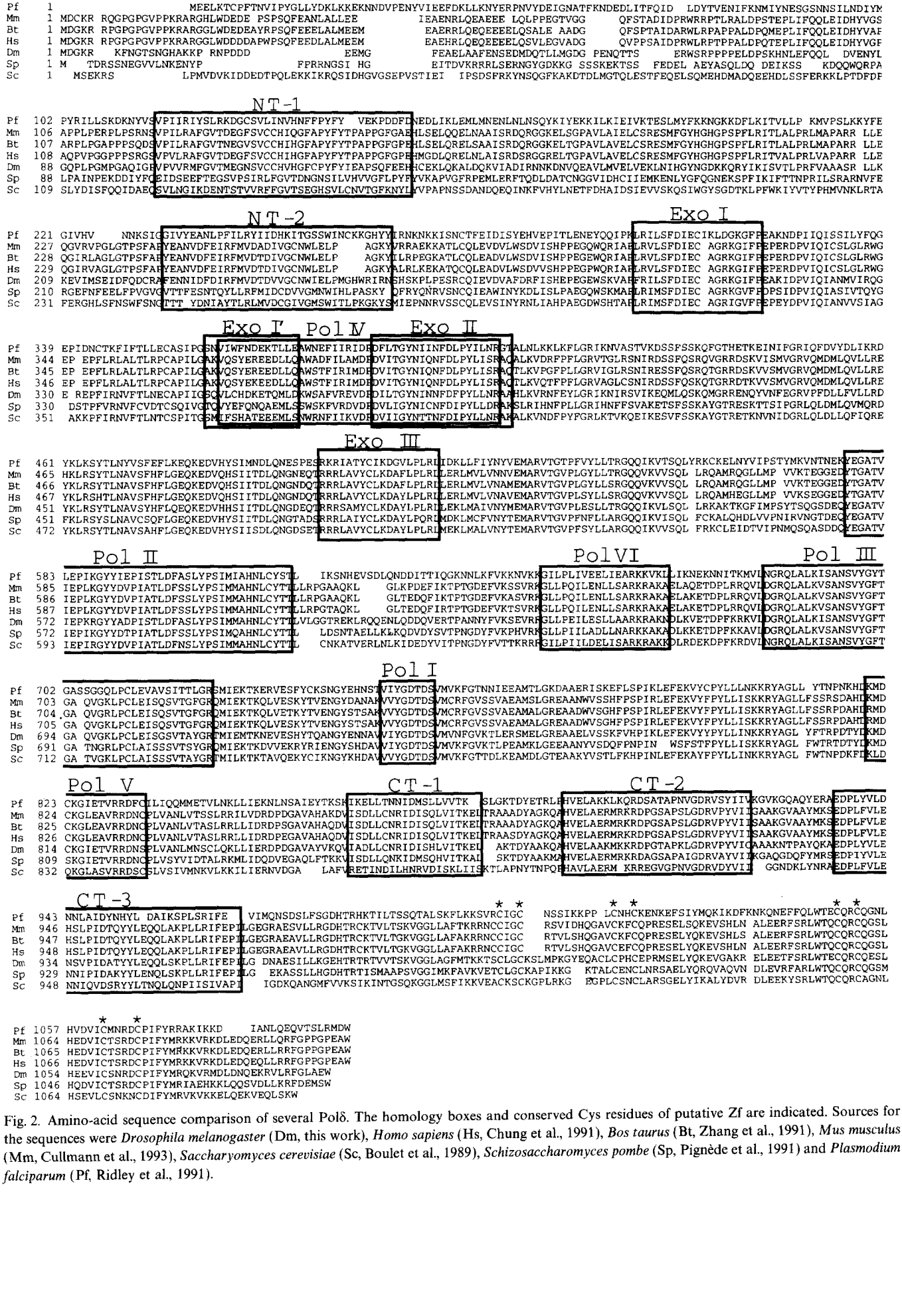

Fig. 2. Amino-acid sequence comparison of several Polδ. The homology boxes and conserved Cys residues of putative Zf are indicated. Sources for the sequences were *Drosophila melanogaster* (Dm, this work), *Homo sapiens* (Hs, Chung et al., 1991), *Bos taurus* (Bt, Zhang et al., 1991), *Mus musculus* (Mm, Cullmann et al., 1993), *Saccharyomyces cerevisiae* (Sc, Boulet et al., 1989), *Schizosaccharomyces pombe* (Sp, Pignède et al., 1991) and *Plasmodium falciparum* (Pf, Ridley et al., 1991).

(e) Conclusions

(*1*) A full-length cDNA encoding *Dm* Polδ has been isolated using a combination of PCR and cDNA library screening.

(*2*) The deduced ORF contains all the homology boxes for Polδ.

(*3*) A comparison of the deduced aa sequence with those from other sources indicate that Polδ is very highly conserved.

ACKNOWLEDGEMENT

This work was supported by Grant GM-06196 from the National Institutes of Health.

REFERENCES

Aboussekhra, A., Biggerstaff, M., Shivji, M.K.K., Vilpo, J.A., Moncollin, V., Podust, V.N., Protic, M., Hübscher, U., Egly, J.-M. and Wood, R.D.: Mammalian DNA nucleotide excision repair reconstituted with purified protein components. Cell 80 (1995) 859–868.

Bauer, G.A. and Burgers, P.M.J.: The yeast analog of mammalian cyclin/proliferating cell nuclear antigen interacts with mammalian DNA polymerase δ. Proc. Natl. Acad. Sci. USA 85 (1988) 7506–7510.

Bernad, A., Blanco, L., Lázaro, J.M., Martin, G. and Salas, M.: A conserved 3′-5′ exonuclease active site in prokaryotic and eukaryotic DNA polymerases. Cell 59 (1989) 219–228.

Bernad, A., Lázaro, J.M., Salas, M. and Blanco, L.: The highly conserved amino acid sequence motif Tyr-Gly-Asp-Thr-Asp-Ser in α-like DNA polymerases is required by ϕ29 DNA polymerase for protein-primed initiation and polymerization. Proc. Natl. Acad. Sci. USA 87 (1990) 4610–4614.

Boulet, A., Simon, M., Faye, G., Bauer, G.A. and Burgers, P.M.J.: Structure and function of the *Saccharomyces cerevisiae CDC2* gene encoding the large subunit of DNA polymerase III. EMBO J. 8 (1989) 1849–1854.

Brown, W.C. and Campbell, J.L.: Interaction of proliferating cell nuclear antigen with yeast DNA polymerase δ. J. Biol. Chem. 268 (1993) 21706–21710.

Budd, M.E. and Campbell, J.L.: DNA polymerase required for repair of UV-induced damage in *Saccharomyces cerevisiae*. Mol. Cell. Biol. 15 (1995) 2173–2179.

Cavener, D.R.: Comparison of the consensus sequence flanking translational start sites in *Drosophila* and vertebrates. Nucleic Acids Res. 15 (1987) 1353–1361.

Chiang, C.-S., Mitsis, P.G. and Lehman, I.R.: DNA polymerase δ from embryos of *Drosophila melanogaster*. Proc. Natl. Acad. Sci USA 90 (1993) 9105–9109.

Chung, D.W., Zhang, J., Tan, C.-K., Davie, E.W., So, A.G. and Downey, K.M.: Primary structure of the catalytic subunit of human DNA polymerase δ and chromosomal location of the gene. Proc. Natl. Acad. Sci. USA 88 (1991) 11197–11201.

Copeland, W.C. and Wang, T.S.-F.: Mutational analysis of the human DNA polymerase α. The most conserved region in α-like DNA polymerases is involved in metal-specific catalysis. J. Biol. Chem. 268 (1993) 11028–11040.

Cullmann, G., Hindges, R., Berchtold, M.W. and Hübscher, U.: Cloning of a mouse cDNA encoding DNA polymerase δ: refinement of the homology boxes. Gene 134 (1993) 191–200.

da Costa, L.T., Liu, B., El-Deiry, W.S., Hamilton S. R., Kinzler, K.W., Vogelstein, B., Markowitz, S., Willson, J.K.V., de la Chapelle, A., Downey, K.M. and So, A.G.: Polymerase δ variants in RER colorectal tumours. Nature Genet. 9 (1995) 10–11.

Derbyshire, V., Grindley, N.D.F. and Joyce, C.M.: The 3′-5′ exonuclease of DNA polymerase I of *Escherichia coli*: contribution of each amino acid at the active site to the reaction. EMBO J. 10 (1991) 17–24.

Digard, P. and Coen, D.M.: A novel functional domain of an α-like DNA polymerase. J. Biol. Chem. 265 (1990) 17393–17396.

Dong, Q., Copeland, W.C. and Wang, T.S.-F.: Mutational studies of human DNA polymerase α. Identification of residues critical for deoxynucleotide binding and misinsertion fidelity of DNA synthesis. J. Biol. Chem. 268 (1993a) 24163–24174.

Dong, Q., Copeland, W.C. and Wang, T.S.-F.: Mutational studies of human DNA polymerase a. Serine 867 in the second most conserved region among α-like DNA polymerases is involved in primer binding and mispair primer extension. J. Biol. Chem. 268 (1993b) 24175–24182.

Dorsky, D.I. and Crumpacker, C.S.: Site-specific mutagenesis of a highly conserved region of the herpes simplex virus type 1 DNA polymerase gene. J. Virol. 64 (1990) 1394–1397.

Gibbs, J.S., Chiou, H.C., Bastow, K.F., Cheng, Y.-C. and Coen, D.M.: Identification of amino acids in herpes simplex virus DNA polymerase involved in substrate and drug recognition. Proc. Natl. Acad. Sci. USA 85 (1988) 6672–6676.

Hernandez, T.R. and Lehman, I.R.: Functional interaction between the herpes simplex-1 DNA polymerase and UL42 protein. J. Biol. Chem. 265 (1990) 11227–11232.

Larder, B.A., Kemp, S.D. and Darby, G.: Related functional domains in virus DNA polymerases. EMBO J. 6 (1987) 169–175.

Lehman, I.R. and Kaguni, L.S.: DNA polymerase α. J. Biol. Chem. 264 (1989) 4265–4268.

Marcy, A.I., Hwang, C.B.C., Ruffner, K.L. and Coen, D.M.: Engineered herpes simplex virus DNA polymerase point mutants: the most highly conserved region shared among α-like DNA polymerases is involved in substrate recognition. J. Virol. 64 (1990) 5883–5890.

Nishida, C., Reinhard, P. and Linn, S.: DNA repair synthesis in human fibroblasts requires DNA polymerase δ. J. Biol. Chem. 263 (1988) 501–510.

Ollis, D.L., Brick, P., Hamlin, R., Xuong, N.G. and Teitz, T.A.: Structure of large fragment of *Escherichia coli* DNA polymerase I complexed with dTMP. Nature 313 (1985) 762–766.

Pignède, G., Bouvier, D., de Recondo, A.-M. and Baldacci, G.: Characterization of the *POL3* gene product from *Schizosaccharomyces pombe* indicates inter-species conservation of the catalytic subunit of DNA polymerase δ. J. Mol. Biol. 222 (1991) 209–218.

Ridley, R.G., White, J.H., McAleese, S.M., Goman, M., Alano, P., de Vries, E. and Kilbey, B.J.: DNA polymerase δ: gene sequence from *Plasmodium falciparum* indicates that this enzyme is more highly conserved than DNA polymerase α. Nucleic Acids Res. 19 (1991) 6731–6736.

Simon, M., Giot, L. and Faye, G.: The 3′ to 5′ exonuclease activity located in the DNA polymerase δ subunit of *Saccharoymyces cerevisiae* is required for accurate replication. EMBO J. 10 (1991) 2165–2170.

Stillman, B.: Smart machines at the DNA replication fork. Cell 78 (1994) 725–728.

Tsurimoto, T., Melendy, T. and Stillman, B.: Sequential initiation of lagging and leading strand synthesis by two different polymerase complexes at the SV40 DNA replication origin. Nature 346 (1990) 534–539.

Tsurumi, T., Maeno, K. and Nishiyama, Y.: A single-base change within the DNA polymerase locus of herpes simplex virus type 2 confers resistance to aphidicolin. J. Virol. 61 (1987) 388–394.

Wang, T.S.-F.: Eukaryotic DNA polymerases. Annu. Rev. Biochem. 60 (1991) 513–552.

Wong, S.W., Wahl, A.F., Yuan, P.M., Arai, N., Pearson, B.E., Arai, K., Korn, D., Hunkapiller, M.W. and Wang, T.S.-F.: Human DNA polymerase α gene expression is cell proliferation dependent and its primary structure is similar to both prokaryotic and eukaryotic replicative DNA polymerases. EMBO J. 7 (1988) 37–47.

Yang, C.-L., Chang, L.-S., Zhang, P., Hao, H., Zhu, L., Toomey, N.L. and Lee, M.Y.W.T.: Molecular cloning of the cDNA for the catalytic subunit of human DNA polymerase δ. Nucleic Acids Res. 20 (1992) 735–745.

Zhang, J., Chung, D.W., Tan, C.-K., Downey, K.M., Davie, E.W. and So, A.G.: Primary structure of the catalytic subunit of calf thymus DNA polymerase δ: sequence similarities with other DNA polymerases. Biochemistry 30 (1991) 11742–11750.

Zhang, S.-J., Zeng, X.-R., Zhang, P., Toomey, N.L., Chuang, R.-Y., Chang, L.-S. and Lee, M.Y.W.T.: A conserved region in the N terminus of DNA polymerase δ is involved in proliferation cell nuclear antigen binding. J. Biol. Chem. 270 (1995) 7988–7992.