International Journal of Biochemistry Research & Review
3(3): 206-247, 2013

SCIENCEDOMAIN *international*
www.sciencedomain.org

DNA Polymerases: An Insight into Their Active Sites and Catalytic Mechanism

P. Palanivelu[1*]

[1]*Department of Molecular Microbiology, School of Biotechnology, Madurai Kamaraj University, Madurai – 625 021, India.*

Author's contribution

The only author performed the whole research work. Author PP wrote the first draft of the paper. Author PP read and approved the final manuscript.

Research Article

Received 16th March 2013
Accepted 6th May 2013
Published 31st May 2013

ABSTRACT

Aim: To analyze the active sites of various prokaryotic and eukaryotic DNA polymerases and propose a plausible mechanism of action for the polymerases with the *Escherichia coli* DNA polymerase I as a model system.
Study Design: Bioinformatics, Biochemical and X-ray crystallographic data were analyzed.
Place and Duration of Study: Department of Molecular Microbiology, School of Biotechnology, Madurai Kamaraj University, Madurai – 625 021, India. From 2007 to 2012.
Methodology: The advanced version of T-COFFEE was used to analyze both prokaryotic and eukaryotic DNA polymerase sequences. Along with this bioinformatics data, X-ray crystallographic and biochemical data were used to confirm the possible amino acids in the active sites of different types of polymerases from various sources.
Results: Multiple sequence analyses of various polymerases from different sources show only a few highly conserved motifs among these enzymes except eukaryotic epsilon polymerases where a large number of highly conserved sequences are found. Possible catalytic/active site regions in all these polymerases show a highly conserved catalytic amino acid K/R and the YG/A pair. A distance conservation is also observed between the active sites. Furthermore, two highly conserved Ds and DXD motifs are also observed.
Conclusion: The highly conserved amino acid K/R acts as the proton abstractor in catalysis and the YG/A pair acts as a "steric gate" in selection of only dNTPS for polymerization reactions. The two highly conserved Ds act as the "charge shielder" of dNTPs and orient the

**Corresponding author: Email: ppmkupp@gmail.com; ppalani1953@rediffmail.com;*

alpha phosphate of incoming dNTPs to the 3'-OH end of the growing primer.

Keywords: DNA polymerases; T-COFFEE; conserved motifs; active sites; E. coli DNA polymerase I; polymerase active site; polymerization mechanism; exonuclease active site; proof-reading mechanism.

1. INTRODUCTION

DNA polymerases are cardinal enzymes, which play a vital role in not only preserving but also maintaining the blue print of life in all living cells. Polymerization of deoxynucleoside triphosphates (dNTPs) into DNA by *E. coli* extract was first demonstrated by Arthur Kornberg in 1955 [1]. Since then a large number of DNA polymerases from various sources have been purified, characterized, cloned and sequenced. Complete nucleic acid and protein sequence data are available for many of these enzymes. These data have become valuable tools in analyzing and understanding the highly conserved motifs and their structure-function relations of these enzymes. Many of these prokaryotic enzymes are usually multifunctional enzymes and exhibit three different activities, viz., polymerization, proof-reading and DNA repair. Some of the polymerases like *Thermus aquaticus* and *Pyrococcus furiosus* are extensively used in polymerase chain reaction techniques because of their exceptional thermostability. Except viruses, there is always more than one type of polymerases in a living cell. For example, at least five different polymerases have been characterized in prokaryotes such as *E. coli*, (e.g.), pol I, II, III, IV and V. Eukaryotic DNA polymerase family is a more complex one. At least 18 different polymerases have been reported from eukaryotes. To distinguish them from prokaryotic polymerases, they are called alpha, beta, gamma, delta, epsilon, zeta, eta, theta, iota, tau, etc. Among them, the five key enzymes are; α, β, γ, δ and ε which are also analyzed in this article. Except γ polymerases, all are localized in the nucleus. The DNA polymerases, which undertake replication, are also known as replicases.

1.1 Dynamics of DNA Polymerization and Proof Reading Activities

DNA polymerases synthesize a new strand of DNA on a given template DNA using a primer and the four dNTPs. They add nucleotides one at a time to the 3' OH end of the primer as shown in Fig. 1 and thus the new strand grows from 5'$\rightarrow$ 3'. The overall reaction catalyzed by DNA polymerases may be written as,

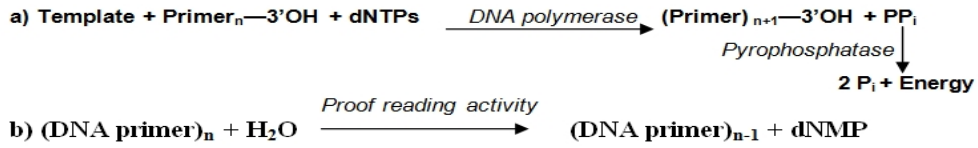


Fig. 1. Dynamics of a) DNA polymerization and b) Poof reading activity

In a dNTP, the strong negative charges on the phosphate groups repel each other and hence weakens the P—O bond. The hydrolysis of P—O bond results in release of large negative free energy, which is utilized in the formation of the phosphodiester bond involving large positive free energy. The inorganic pyrophosphate formed during the reaction is hydrolyzed immediately by the enzyme pyrophosphatase and results in release of more energy which helps in the translocation of the polymerase on the template. Such a coupling of reaction drives the polymerization reaction only in the forward direction, (i.e.), the polymerized nucleotides cannot be depolymerized.

1.2 Crystal Structure of Polymerases

Crystal structures of many polymerases like Klenow polymerase, DNA polymerase I of *E. coli*, *Taq* polymerase, HIV-1 Reverse transcriptase, T_4 RNA polymerase, rat DNA polymerase β, etc., are known. Although their primary structures are markedly different (probably due to convergent evolution), their crystal structures are remarkably similar in overall shape. All DNA polymerases, irrespective of their source, have a common protein fold that resembles the shape of a half opened "right hand" with three distinct domains, “thumb”, “palm” and “fingers” with analogous functions. The catalytic centre is located on the “palm” domain with two conserved Asp residues. The “thumb” and “fingers” domains, although analogous, are not homologous between the pol families. The "thumb," "fingers," and "palm" form a pocket along which the DNA template with the primer can move. The DNA molecule interacts with specific amino acids located in the "palm" region whereas the “thumb” and “finger” are involved in the polymerization reaction. The finger and thumb domains wrap up the DNA around and hold it across the enzyme’s active site, which comprises residues primarily from the palm domain. Among them the palm region is the most conserved [2]. Similar spatial arrangements of the highly conserved carboxylic acids in these enzymes suggest that the polymerases contain similar active sites and therefore, are likely to operate by the same catalytic mechanism.

2. MATERIAL AND METHODS

For multiple sequence analysis of various polymerases the sequences were retrieved from SWISS-PROT and PUBMED sites and analyzed using the advanced version of T-COFFEE available at the website: https://tcoffee.org.

3. RESULTS AND DISCUSSION

3.1 Active Site Analysis of DNA Polymerases

To-date the most well characterized polymerase is the DNA polymerase I of *E. coli.* DNA polymerase I is a multifunctional enzyme and harbours at least 6 different binding/active sites on the single polypeptide viz., One binding site for all four dNTPs, One binding site for template DNA or at nicks on DNAs, One binding site for the primer, One binding site for the 3' OH on the primer, One binding site (exonuclease-1 active site) for proof-reading function and One binding site (exonuclease-2 active site) for DNA repair-function. In addition to the above binding sites, it also contains several binding sites for metal ions as well. The dNMP site (part of proof-reading active site) and dNTP site (part of polymerase active site) bind their respective ligands non-competitively. In eukaryotic polymerases the polymerization and proof reading activities are accomplished by two different polymerases. In order to analyze various active and binding sites of different polymerases a multiple sequence alignment was done.

3.2 Multiple Sequence Analysis of Different Polymerases from Different Sources

Figs. 2A – 2E show the results of multiple sequence alignment of various prokaryotic polymerases (pol I to pol V). (The conserved regions are shaded and the active site regions are underlined and shaded yellow).

Fig. 2F shows the results of multiple sequence alignment of various viral polymerases. (The conserved regions are shaded and the active site regions are underlined and shaded yellow).

Figs. 2G – 2K show the results of multiple sequence alignment of various eukaryotic polymerases (alpha, beta, gamma, delta and epsilon. (The conserved regions are shaded and the active site regions are underlined and shaded yellow).

It is clear from the Figs. 2A-2K, irrespective of the type and origin of the polymerase, marked similarities are observed in their catalytic regions which are summarized in Table 1.

Fig. 2A Prokaryotic DNA polymerases (pol I)

```
sp|P52026|DPO1_BACST  GKTTFRHETFQDYKGGRQQTPPELSEQFPLLRELLKAYRIPAYELDHYEADDIIGTMAAR 120
sp|P00582|DPO1_ECOLI  KGKTFRDELFEHYKSHRPPMPDDLRAQIEPLHAMVKAMGLPLLAVSGVEADDVIGTLARE 124
sp|P0A550|DPO1_MYCTU  SRQTFRLQRYPEYKANRSSTPDEFAGQIDITKEVLGALGITVLSEPGFEADDLIATLATQ 134
sp|O34996|DPO1_BACSU  GKTTFRHGTFKEYKGGRQKTPPELSEQMPFIRELLDAYQISRYELEQYEADDIIGTLAKS 122
sp|Q9F173|DPO1_SALTY  KGKTFRDELFEHYKSHRPPMPDDLRAQIEPLHAMVKAMGLPLLAVSGVEADDVIGTLARE 124
sp|P59200|DPO1_STRR6  GKTTFRTEMYADYKGGRAKTPDEFREQFPFIRELLDHMGIRHYELAQYEADDIIGTLDKL 125
sp|O32801|DPO1_LACLM  GKTTFRTEMFADYKGGRSKTPDEFREQLPFIKEMIEKLGIRHYELANYEADDIIGTLDKM 126
sp|P43741|DPO1_HAEIN  KGKTFRDEMFEQYKSHRPPMPDDLRKQIQPLHDMIRALGIPLLVVEGIEADDVIGTLALQ 124
sp|Q9ZJE9|DPO1_HELPJ  QTKTKRAEKLGEYKQNRKDAPKEMLLQIPIALEWLQKMGFTCVEVGGFEADDVIASLATL 130
sp|Q9CDS1|DPO1_LACLA  GKTTFRTEMFADYKGGRSKTPDEFREQLPFIKEMIEKLGIRHYELANYEADDIIGTLDKM 126
sp|P59199|DPO1_STRPN  GKTTFRTEMYADYKGGRAKTPDEFREQFPFIRELLDHMGIRHYDLAQYEADDIIGTLDKL 125
sp|Q9S1G2|DPO1_RHILE  SAKTFRKDLYDAYKANRSAPPEELIPQFGLIREATRAFNLPCIETEGFEADDIIATYARQ 147
sp|Q92GB7|DPO1_RICCN  GGKNFRHQIYPDYKANRPPPPEDLIIQLPLVRDVASNLNFPILEKNGYEADDIIATFATK 122
sp|P0A551|DPO1_MYCBO  SRQTFRLQRYPEYKANRSSTPDEFAGQIDITKEVLGALGITVLSEPGFEADDLIATLATQ 134
sp|Q9HT80|DPO1_PSEAE  KGPTFRDELFAEYKANRPSMPDDLRVQVEPLHASVRALGLPLLCVEGVEADDVIGTLARS 121
sp|Q04957|DPO1_BACCA  GKTTFRHEAFQEYKGGRQQTPPELSEQFPLLRELLRAYRIPAYELENYEADDIIGTLAAR 120
sp|P46835|DPO1_MYCLE  SRKTFRSECYAGYKANRSSIPAEFHGQIDITKEVLGALGITVFAEAGFEADDLIATLATQ 135
sp|P52027|DPO1_DEIRA  PVKTFRHEQYEGYKSGRAQTPEDLPGQINRIRALVDALGFPRLEEPGYEADDVIASLTRM 163
                         . *      **  *   * ::  *.          :       ****:*.:


sp|P52026|DPO1_BACST  AER--EGFAVKVISGDRDLTQLASPQVTVEITKKGITDIESYTPETVVEKYGLTPEQIVD 178
sp|P00582|DPO1_ECOLI  AEK--AGRPVLISTGDKDMAQLVTPNITLINTMTN----TILGPEEVVNKYGVPPELIID 178
sp|P0A550|DPO1_MYCTU  AEN--EGYRVLVVTGDRDALQLVSDDVTVLYPRKGVSELTRFTPEAVVEKYGLTPRQYPD 192
sp|O34996|DPO1_BACSU  AEK--DGFEVKVFSGDKDLTQLATDKTTVAITRKGITDVEFYTPEHVKEKYGLTPEQIID 180
sp|Q9F173|DPO1_SALTY  AEK--VGRPVLISTGDKDMAQLVTPNITLINTMTN----TILGPDEVVNKYGVPPELIID 178
sp|P59200|DPO1_STRR6  AEQ--DGFDITIVSGDKDLIQLTDEHTVVEISKKGVAEFEAFTPDYLMEEMGLTPAQFID 183
sp|O32801|DPO1_LACLM  AEAPDVNFDVTIVTGDKDMIQLVDGNTRVEISKKGVAEFEEFTPDYLLEKMGLTPSQFID 186
sp|P43741|DPO1_HAEIN  ASS--LGKKVLISTGDKDMAQLVDDNIMLINTMNN----SLLDRKGVIEKYGIPPELIID 178
sp|Q9ZJE9|DPO1_HELPJ  SP-----YKTRIYSKDKDFNQLLSDKIALFDGK------TEFLAKDCVEKYGILPSQFTD 179
sp|Q9CDS1|DPO1_LACLA  AEAPNVNFDVTIVTGDKDMIQLVDGNTRVEISKKGVAEFEEFTPDYLLEKMGLTPAQFID 186
sp|P59199|DPO1_STRPN  AEQ--DGFDITIVSGDKDLIQLTDEHTVVEISKKGVAEFEAFTPDYLMEEMGLTPAQFID 183
sp|Q9S1G2|DPO1_RHILE  AEA--TGADVTIVSSDKDLMQLVSPNVHMYDSMKD----KQIGIPDVIEKWGVPPEKMID 201
sp|Q92GB7|DPO1_RICCN  TAA--LGAHVVIISSDKDLLQLMTENIKIYDPLKG----KYITEDDVVKKFGTTSDKLRE 176
sp|P0A551|DPO1_MYCBO  AEN--EGYRVLVVTGDRDALQLVSDDVTVLYPRKGVSELTRFTPEAVVEKYGLTPRQYPD 192
sp|Q9HT80|DPO1_PSEAE  SAA--ADRPVVISTGDKDMAQLVDGHITLVNTMTG----SRLDVDGVKEKFGVGPELIID 175
sp|Q04957|DPO1_BACCA  AEQ--EGFEVKVISGDRDLTQLASPHVTVDITKKGITDIEPYTPEAVREKYGLTPEQIVD 178
sp|P46835|DPO1_MYCLE  AEN--EGYRVLVVTGDRDALQLVSNDVTVLYPRKGVSELTRFTPEAVIEKYGVTPAQYPD 193
sp|P52027|DPO1_DEIRA  AEG--KGYEVRIVTSDRDAYQLLDEHVKVIAN-----DFSLIGPAQVEEKYGVTVRQWVD 216
                      :         : : *:*  **   .  :                 :: *       :
```

```
sp|P52026|DPO1_BACST  NMEFTGVKVDTKRLEQMGAELTEQLQAVERRIYELAGQEFNINSPKQLGTVLFDKLQLPV 546
sp|P00582|DPO1_ECOLI  RIERNGVKIDPKVLHNHSEELTLRLAELEKKAHEIAGEEFNLSSTKQLQTILFEKQGIKP 598
sp|P0A550|DPO1_MYCTU  KMESAGIAVDLPMLTELQSQFGDQIRDAAEAAYGVIGKQINLQSPKQLQVVLFDELGMPK 573
sp|O34996|DPO1_BACSU  EMESTGVKVDVDRLKRMGEELGAKLKEYEEKIHEIAGEPFNINSPKQLGVILFEKIGLPV 550
sp|Q9F173|DPO1_SALTY  RVERNGVKIDPAVLHKHSEEITLRLAELEKKAHDIAGEAFNLSSTKQLQTILFEKQGIKP 598
sp|P59200|DPO1_STRR6  KMEIAGIVVKKETLLEMQAENELVIEKLTQEIYELAGEEFNVNSPKQLGVLLFEKLGLPL 545
sp|O32801|DPO1_LACLM  KMEITGISVSQNTLEEIGAENEEKLASLTREIYDLAGEEFNINSPKQLGVILFEKLQLPV 548
sp|P43741|DPO1_HAEIN  RMERTGVLIDSDALFMQSNEIASRLTALEKQAYALAGQPFNLASTKQLQEILFDKLELPV 601
sp|Q9ZJE9|DPO1_HELPJ  GMEFQGFKIDAPYFKRLEQEFKNELNVLERQILDLIGVDFNLNSPKQLGEVLYDKLGLPK 575
sp|Q9CDS1|DPO1_LACLA  KMEITGIAVSQNTLEEIGAENEEKLASLTREIYDLAGEEFNINSPKQLGVILFEKLQLPV 548
sp|P59199|DPO1_STRPN  KMEIAGIMVKKETLLEMQAENELVIEKLTQEIYELAGEEFNVNSPKQLGVLLFEKLGLPL 545
sp|Q9S1G2|DPO1_RHILE  RMEARGITVDRQILSRLSGELAQGAARLEDEIYVLAGERFNIGSPKQLGDILFGKMGLSG 679
sp|Q92GB7|DPO1_RICCN  KMEKVGITVDANYLKQLSAEFGTEILKLEEEIFALSGTKFNIGSPKQLGEILFEKMQLPF 545
sp|P0A551|DPO1_MYCBO  KMESAGIAVDLPMLTELQSQFGDQIRDAAEAAYGVIGKQINLQSPKQLQVVLFDELGMPK 573
sp|Q9HT80|DPO1_PSEAE  RIERNGALVDANLLGIQSRELGEKMVALERQAYDLAGQEFNLGSPKQLGAILYDKLGLPV 582
sp|Q04957|DPO1_BACCA  EMEFAGVKVDTKRLEQMGEELAEQLRTVEQRIYELAGQEFNINSPKQLGVILFEKLQLPV 546
sp|P46835|DPO1_MYCLE  DMEKAGIAADLRLLTELQSQFGDQIRDAAEAAYAVIGKQINLSSPKQLQVVLFEELGMPK 577
sp|P52027|DPO1_DEIRA  RMEVRGVQVDSDFLQTLSIQAGVRLADLESQIHEYAGEEFHIRSPKQLETVLYDKLELAS 627
                       :*  *   .   :     :                *  ::: *.***  :*: :  :

sp|P52026|DPO1_BACST  TGKVHTMFNQALTQTGRLSSVEPNLQNIPIRLEEGRKIRQAFVPSEP--DW-LIFAADYS 655
sp|P00582|DPO1_ECOLI  TGRVHTSYHQAVTATGRLSSTDPNLQNIPVRNEEGRRIRQAFIAPE---DY-VIVSADYS 707
sp|P0A550|DPO1_MYCTU  DGRIHTTFNQTIAATGRLSSTEPNLQNIPIRTDAGRRIRDAFVVGD---GYAELMTADYS 682
sp|O34996|DPO1_BACSU  SHKVHTRFNQALTQTGRLSSTDPNLQNIPIRLEEGRKIRQAFVPSEK--DW-LIFAADYS 659
sp|Q9F173|DPO1_SALTY  TGRVHTSYHQAVTATGRLSSTDPNLQNIPVRNEEGRRIRQAFIAPE---DY-LIVSADYS 707
sp|P59200|DPO1_STRR6  DGKIHTRYVQDLTQTGRLSSVDPNLQNIPARLEQGRLIRKAFVPEWE--DS-VLLSSDYS 655
sp|O32801|DPO1_LACLM  DGKIHTRYVQDLTQTGRLSSVDPNLQNIPVRLEEGRKIRKAFVPS-Q--DS-LLLSSDYS 655
sp|P43741|DPO1_HAEIN  TGRVHTSYHQAVTATGRLSSSDPNLQNIPIRNEEGRHIRQAFIARE---GY-SIVAADYS 710
sp|Q9ZJE9|DPO1_HELPJ  DDKIHTTFIQTGTATGRLSSHSPNLQNIPVRSPKGLLIRKGFIASSK--EY-CLLGVDYS 679
sp|Q9CDS1|DPO1_LACLA  DGKIHTRYVQDLTQTGRLSSVDPNLQNIPVRLEEGRKIRKAFVPS-K--DS-LLLSSDYS 655
sp|P59199|DPO1_STRPN  DGKIHTRYVQDLTQTGRLSSVDPNLQNIPARLEQGRLIRKAFVPEWE--DS-VLLSSDYS 655
sp|Q9S1G2|DPO1_RHILE  TKRVHTSYSLASTTTGRLSSSEPNLQNIPVRTAEGRKIRTAFISTP---GH-KLISADYS 789
sp|Q92GB7|DPO1_RICCN  THRVHTTFLQTSTTTGRLSSQEPNLQNVPIRSSEGNQIRKAFIAEE---GY-KLISADYS 656
sp|P0A551|DPO1_MYCBO  DGRIHTTFNQTIAATGRLSSTEPNLQNIPIRTDAGRRIRDAFVVGD---GYAELMTADYS 682
sp|Q9HT80|DPO1_PSEAE  TGRIHTSYHQAVAATGRLSSSDPNLQNIPIRTAEGRRIRQAFVAPQ---GY-KLLAADYS 692
sp|Q04957|DPO1_BACCA  TKKVHTIFNQALTQTGRLSSTEPNLQNIPIRLEEGRKIRQAFVPSES--DW-LIFAADYS 656
sp|P46835|DPO1_MYCLE  DGRIHTTFNQTIATTGRLSSTEPNLQNIPIRTNAGRQIRDAFVVGSENNGYTELMTADYS 689
sp|P52027|DPO1_DEIRA  TGRLHTTFAQTAVATGRLSSLNPNLQNIPIRSELGREIRKGFIAED---GF-TLIAADYS 737
                        ::** :    . ****** .*****:* *   *  ** .*:          :.  ***

sp|P52026|DPO1_BACST  QIELRVLAHIAEDDNLIEAFRRGLDIHTKTAMDIFHVSE-EDVTANMRRQAKAVNFGIVY 714
sp|P00582|DPO1_ECOLI  QIELRIMAHLSRDKGLLTAFAEGKDIHRATAAEVFGLPL-ETVTSEQRRSAKAINFGLIY 766
sp|P0A550|DPO1_MYCTU  QIEMRIMAHLSGDEGLIEAFNTGEDLHSFVASRAFGVPI-DEVTGELRRRVKAMSYGLAY 741
sp|O34996|DPO1_BACSU  QIELRVLAHISKDENLIEAFTNDMDIHTKTAMDVFHVAK-DEVTSAMRRQAKAVNFGIVY 718
sp|Q9F173|DPO1_SALTY  QIELRIMAHLSRDKGLLTAFAEGKDIHRATAAEVFGLPL-DSVTGEQRRSAKAINFGLIY 766
sp|P59200|DPO1_STRR6  QIELRVLAHISKDEHLIKAFQEGADIHTSTAMRVFGIERPDNVTANDRRNAKAVNFGVVY 715
sp|O32801|DPO1_LACLM  QIELRVLAHISADEHLIDAFKHGADIHTSTAMRVFGIEKAEDVTANDRRNAKAVNFGVVY 715
sp|P43741|DPO1_HAEIN  QIELRIMAHLSGDQGLINAFSQGKDIHRSTAAEIFGVSL-DEVTSEQRRNAKAINFGLIY 769
sp|Q9ZJE9|DPO1_HELPJ  QIELRLLAHFSQDKDLMEAFLKGRDIHLETSKALFG----EDLAKEKRSIAKSINFGLVY 735
sp|Q9CDS1|DPO1_LACLA  QIELRVLAHISGDEHLIDAFKHGADIHTSTAMRVFGIEKAEDVTANDRRNAKAVNFGVVY 715
sp|P59199|DPO1_STRPN  QIELRVLAHISKDEHLIKAFQEGADIHTSTAMRVFGIERPDDVTANDRRNAKAVNFGVVY 715
sp|Q9S1G2|DPO1_RHILE  QIELRVLAHVAEIPQLTKAFEDGVDIHAMTASEMFGVPV-EGMPGEVRRRAKAINFGIIY 848
sp|Q92GB7|DPO1_RICCN  QIELRILSHIANIDALKQAFINKDDIHTQTACQIFNLQK-HELTSEHRRKAKAINFGIIY 715
sp|P0A551|DPO1_MYCBO  QIEMRIMAHLSGDEGLIEAFNTGEDLHSFVASRAFGVPI-DEVTGELRRRVKAMSYGLAY 741
sp|Q9HT80|DPO1_PSEAE  QIELRIMAHLAKDDGLLDAFRHDLDVHRATAAEVFGVPL-EDVSGDQRRSAKAINFGLIY 751
sp|Q04957|DPO1_BACCA  QIELRVLAHIAEDDNLMEAFRRDLDIHTKTAMDIFQVSE-DEVTPNMRRQAKAVNFGIVY 715
sp|P46835|DPO1_MYCLE  QIEMRIMAHLSRDEGLIEAFHTGEDLHSFVASRAFGIPI-EDITPELRRRVKAMSYGLAY 748
sp|P52027|DPO1_DEIRA  QIELRLLAHIADDPLMQQAFVEGADIHRRTAAQVLGLDE-ATVDANQRRAAKTVNFGVLY 796
                      ***:*:::*.:    :  **    *:*  .:   :       :    *  .*::.:*: *
```

Fig. 2A. CLUSTAL format for T-COFFEE version of Pol I DNA polymerases (Prokaryotes)

The possible catalytic region is underlined and shaded yellow.
Only regions showing highly conserved blocks are shown.

Fig. 2B Prokaryotic DNA polymerases (pol II)

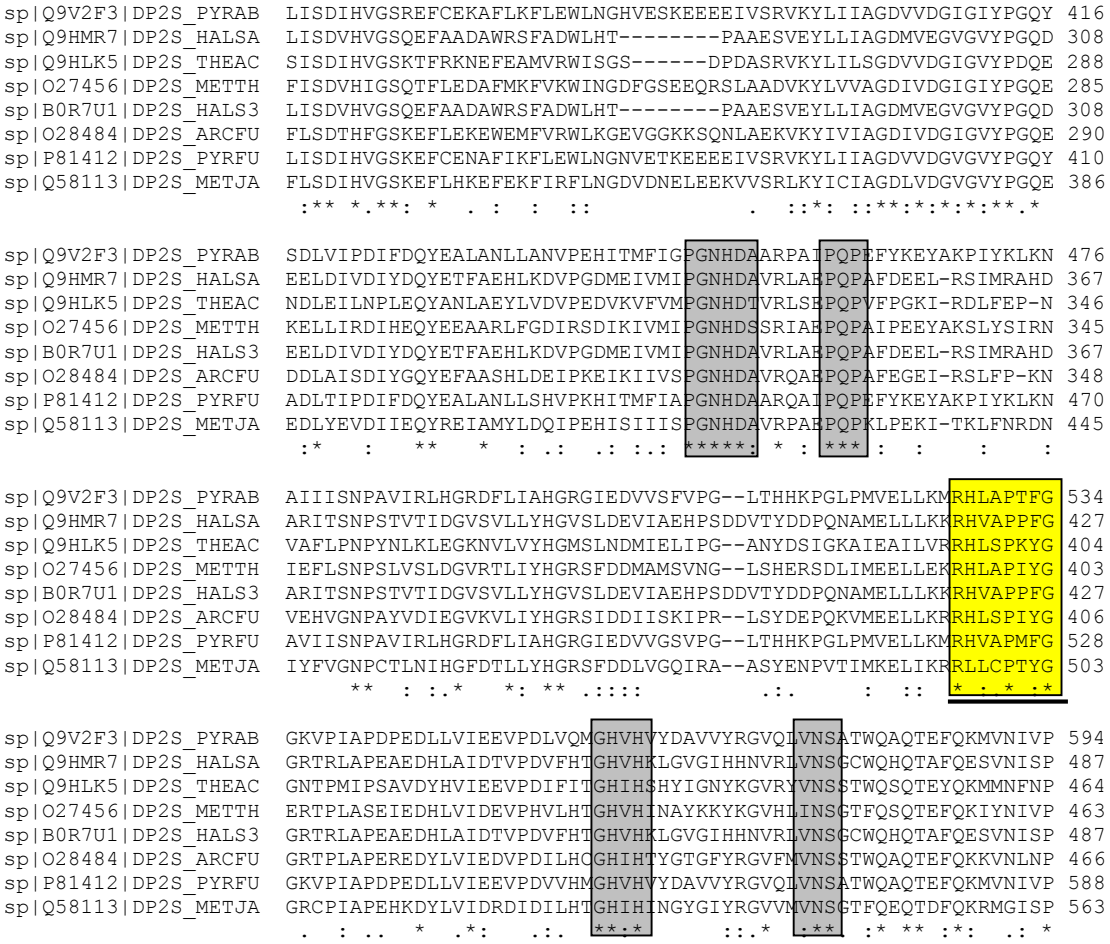
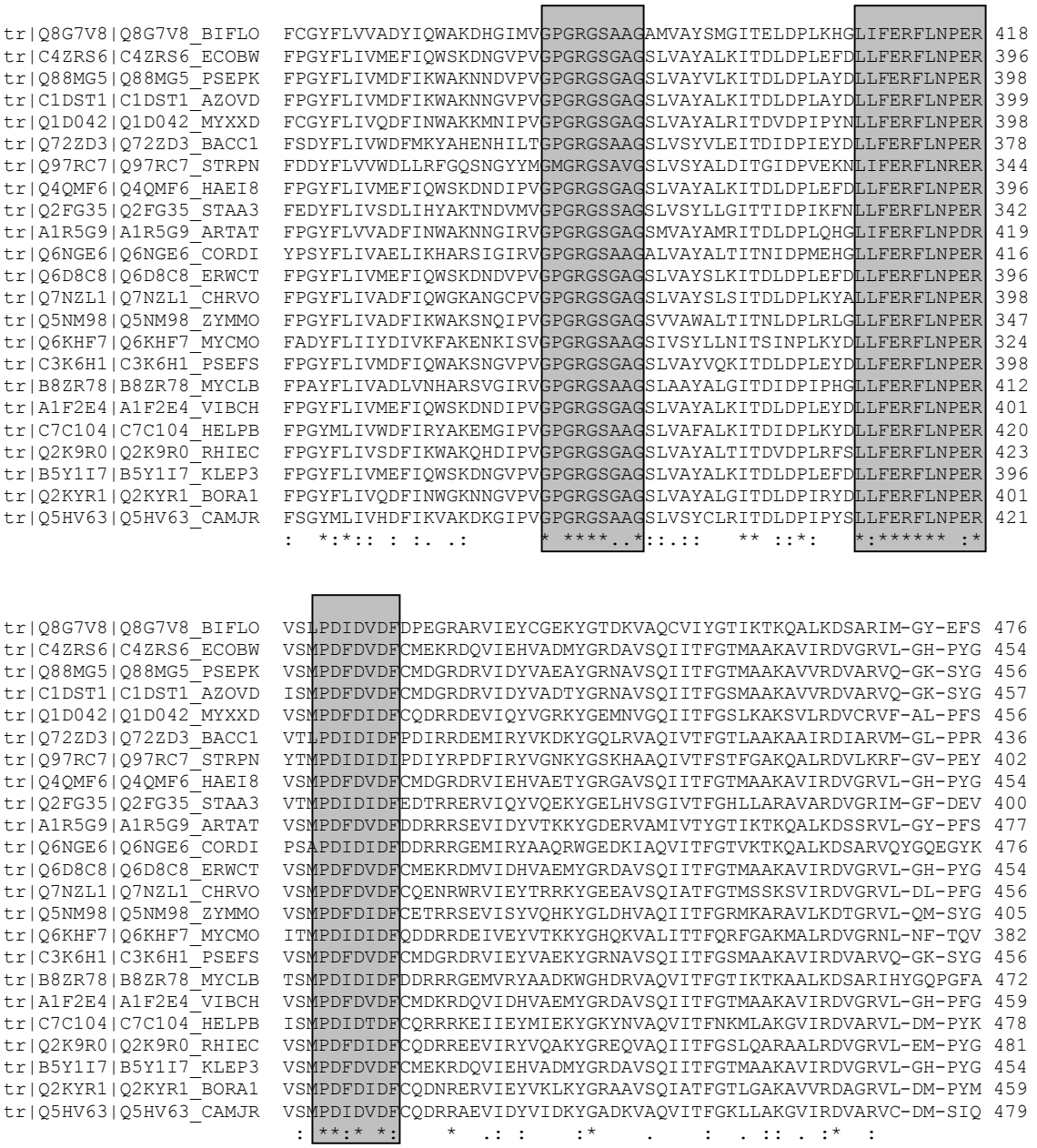


Fig. 2B. CLUSTAL format for T-COFFEE version of Pol II polymerases (Prokaryotes)

The possible catalytic region is underlined and shaded yellow.
Only regions showing highly conserved blocks are shown.

Fig. 2C Prokaryotic DNA polymerases (PolI III)



```
tr|Q8G7V8|Q8G7V8_BIFLO  TLGLVKMDFLGLSNLTIIHDTLKNIEANG--------K-PAIDYTKIPLDDAETYKLMSR 630
tr|C4ZRS6|C4ZRS6_ECOBW  YAGLVKFDFLGLRTLTIINWALEMINKRRAK----NGE-PPLDIAAIPLDDKKSFDMLQR 602
tr|Q88MG5|Q88MG5_PSEPK  AAGLVKFDFLGLRTLTIIKWAMEIINREQAK----KNL-PDLNIDFIPLDDRKTYELLQK 604
tr|C1DST1|C1DST1_AZOVD  AAGLVKFDFLGLRTLTIIKWAMETIHREQRR----RGETELVDIDRIALDDKATYALLQR 606
tr|Q1D042|Q1D042_MYXXD  AAGLVKFDFLGLKTLTVIQHALDLVKRNH--------G-KDIPRHEIPLNDEKTWELMAK 610
tr|Q72ZD3|Q72ZD3_BACC1  ELGLLKMDFLGLRNLTLLENIIKFIVQKT--------G-KEIDIRNLPLQDEKTFQLLGR 579
tr|Q97RC7|Q97RC7_STRPN  ASGLLKMDFLGLRNLTFVQKMQELLAETE--------G-IHLKIEEIDLEDKETLALFAS 542
tr|Q4QMF6|Q4QMF6_HAEI8  YAGLVKFDFLGLRTLTIIKWALDIINVRMVR----EGK-PRVDIAAIPLDDPESFELLKR 602
tr|Q2FG35|Q2FG35_STAA3  RIGLLKIDFLGLRNLSIIHQILTQVKKDLGI----------NIDIEKIPFDDQKVFELLSQ 541
tr|A1R5G9|A1R5G9_ARTAT  GLGLIKMDFLGLRNLTIISDALENIKMNR--------G-VDLDLENLELDDAASYELLAR 631
tr|Q6NGE6|Q6NGE6_CORDI  AIGLLKMDFLGLRNLTVIGDCLENIKINR--------N-EELDLEALQTDDPNVYKLLSS 630
tr|Q6D8C8|Q6D8C8_ERWCT  YAGLVKFDFLGLRTLTIIDWALEMINARRAK----LGQ-EPIDIATIPLGDKKSFDMLQR 602
tr|Q7NZL1|Q7NZL1_CHRVO  QIGLVKFDFLGLRNLTIIELAQKYIKDLT--------G-EDVDVAHLPLDDKPAYKVFAT 602
tr|Q5NM98|Q5NM98_ZYMMO  KTGLVKFDFLGLKTLSVLNKAVQMLKAR---------G-VEVDLDHIPLDDEAVFELLKQ 550
tr|Q6KHF7|Q6KHF7_MYCMO  NFSLLKIDLLGLKALSTIKLIEQKIREAYKIKYNFSNEDEYTEQLKKDLDPKKAFQLLSD 531
tr|C3K6H1|C3K6H1_PSEFS  AAGLVKFDFLGLRTLTIIDWALKTINRDRAK----VNE-EPLDIAFIPLDDKPTYSLLQK 604
tr|B8ZR78|B8ZR78_MYCLB  AIGLLKMDFLGLRNLTIIGDAIENIKTNR--------G-IDLDLESVPLDDQATYELLGR 626
tr|A1F2E4|A1F2E4_VIBCH  TAGLVKFDFLGLRTLTIIDWALGLVNPRLKK----AGK-PPVRIEAIPLDDARSFRNLQD 607
tr|C7C104|C7C104_HELPB  PVDLIKFDFLGLKTLTVIDDALKIIKTQH--------N-IDVDFLSLDMDDPKVYKTIQS 633
tr|Q2K9R0|Q2K9R0_RHIEC  QAGLVKFDFLGLKTLTVLKVAVDFVAKR---------G-INVDLAAIPLDDKPTYEMLSR 626
tr|B5Y1I7|B5Y1I7_KLEP3  YAGLVKFDFLGLRTLTIINWALEMINKRRAK----NGE-GPLDIAAIPLDDKKSFDMLQR 602
tr|Q2KYR1|Q2KYR1_BORA1  AAGLVKFDFLGLRNLTILDWAVRYVRQFN------PDQ-RDFDIMALSLDDPAAYKVLCD 608
tr|Q5HV63|Q5HV63_CAMJR  DVDLIKFDFLGLKTLTVINNAIKLIKKRY--------N-KDIIWETIDVNDSKVYKTIQS 625
                          .*:*:*:***  *: :      :                              :

tr|Q8G7V8|Q8G7V8_BIFLO  GDTLGVFQLDSDGMRSLLKTLKPDNFNDISALIALYRPGPMSMDSHTNYAKRKNGLQKIT 690
tr|C4ZRS6|C4ZRS6_ECOBW  SETTAVFQLESRGMKDLIKRLQPDCFEDMIALVALFRPGPLQSGMVDNFIDRKHGREEIS 662
tr|Q88MG5|Q88MG5_PSEPK  AETTAVFQLESRGMKELIKKLKPDCLEDLIALVALFRPGPLQSGMVDDFINRKHGRAELA 664
tr|C1DST1|C1DST1_AZOVD  AETTAVFQLESRGMKELIKKLKPDNIEDMIALVALFRPGPLQSGMVDDFINRKHGRAELS 666
tr|Q1D042|Q1D042_MYXXD  GDTAGIFQMESSGFTEMVVKLKPNCFEDVIAAGALYRPGPLDSGMVDVFINRKHGREKVS 670
tr|Q72ZD3|Q72ZD3_BACC1  GDTTGVFQLESGGMRNVLRGLKPNEFEDIVAVNSLYRPGPME--QIPTFIESKHGKRKIE 637
tr|Q97RC7|Q97RC7_STRPN  GNTKGIFQFEQPGAIRLLKRVQPVCFEDVVATTSLNRPGASD--YINNFVARKHGQEEVT 600
tr|Q4QMF6|Q4QMF6_HAEI8  SETTAVFQLESRGMKDLIKRLQPDCFEDIIALVALFRPGPLQSGMVDNFIDRKHGREEVS 662
tr|Q2FG35|Q2FG35_STAA3  GDTTGIFQLESDGVRSVLKKLKPEHFEDIVAVTSLYRPGPMEE--IPTYITRRHDPSKVQ 599
tr|A1R5G9|A1R5G9_ARTAT  GDTLGVFQLDGGPMRSLLKLMKPDNFEDISAVLALYRPGPMGANAHTDYALRKNGIQEVI 691
tr|Q6NGE6|Q6NGE6_CORDI  GDTLGVFQLDSGGMQELLKRMQPTGFNDIVASLALYRPGPMGVNAHWDYADRKNGRKPIE 690
tr|Q6D8C8|Q6D8C8_ERWCT  SETTAVFQLESRGMKDLIKRLKPDCFEDMIALVALFRPGPLQSGMVDNFIDRKHGREAIS 662
tr|Q7NZL1|Q7NZL1_CHRVO  ANTTAVFQFESTGMKKMLVEAKPSKFEEIIAFVALYRPGPMD--LIPDFIQRMHGA-KFE 659
tr|Q5NM98|Q5NM98_ZYMMO  GDTVGVFQLESEGMRRTLAAVKPTRFEDIIALVALYRPGPMDN--IPSFGARKNGGEEIV 608
tr|Q6KHF7|Q6KHF7_MYCMO  GNTAGVFQLESPGMIQTIKKVKISKFSDIVDVISLYRPGPMQY--IQEYIDNKNDASKIK 589
tr|C3K6H1|C3K6H1_PSEFS  AETTAVFQLESRGMKELIKKLKPDCLEDLIALVALFRPGPLQSGMVDDFINRKHGRAELA 664
tr|B8ZR78|B8ZR78_MYCLB  GDTLGVFQLDGGPMRDLLRRMQPTGFEDIVAVLALYRPGPMGMNAHNDYADRKNNRQVIK 686
tr|A1F2E4|A1F2E4_VIBCH  AKTTAVFQLESRGMKELIKRLQPDCFEDIIALVALFRPGPLQSGMVDNFIDRKHGREAIS 667
tr|C7C104|C7C104_HELPB  GDTVGIFQIESGMFQGLNKRLRPSSFEDIIAIIALGRPGPMESGMVDDFVNRKHGVEPIA 693
tr|Q2K9R0|Q2K9R0_RHIEC  GETVGVFQVESAGMRKALIGMKPDCIEDIIALVALYRPGPMEN--IPVYNARKHGEEEVE 684
tr|B5Y1I7|B5Y1I7_KLEP3  SETTAVFQLESRGMKDLIKRLQPDCFEDMIALVALFRPGPLQSGMVDNFIDRKHGREEIS 662
tr|Q2KYR1|Q2KYR1_BORA1  ANTTAVFQLESRGMKELLKKLRPNTFEDIIAMLALYRPGPLESGMVDDFVNRKHGRAAVD 668
tr|Q5HV63|Q5HV63_CAMJR  GNTLGIFQIESGGMQSLNARLKPERFEDIIAVLALYRPGPMESGMLDDFIDRKHGLKNIE 685
                         ..* .:**.:          :   :.::    :* ***.        :    :.   .
```

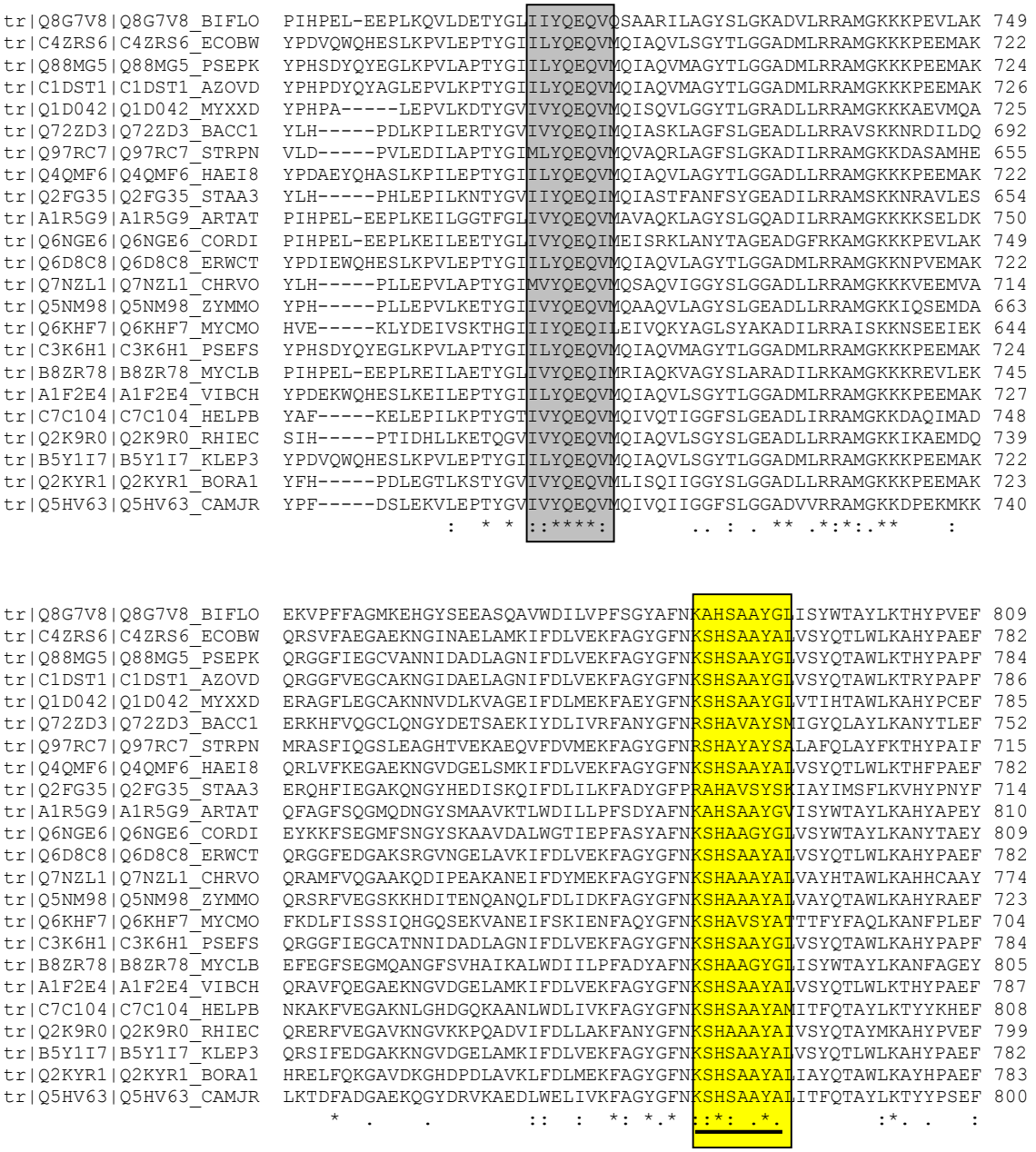
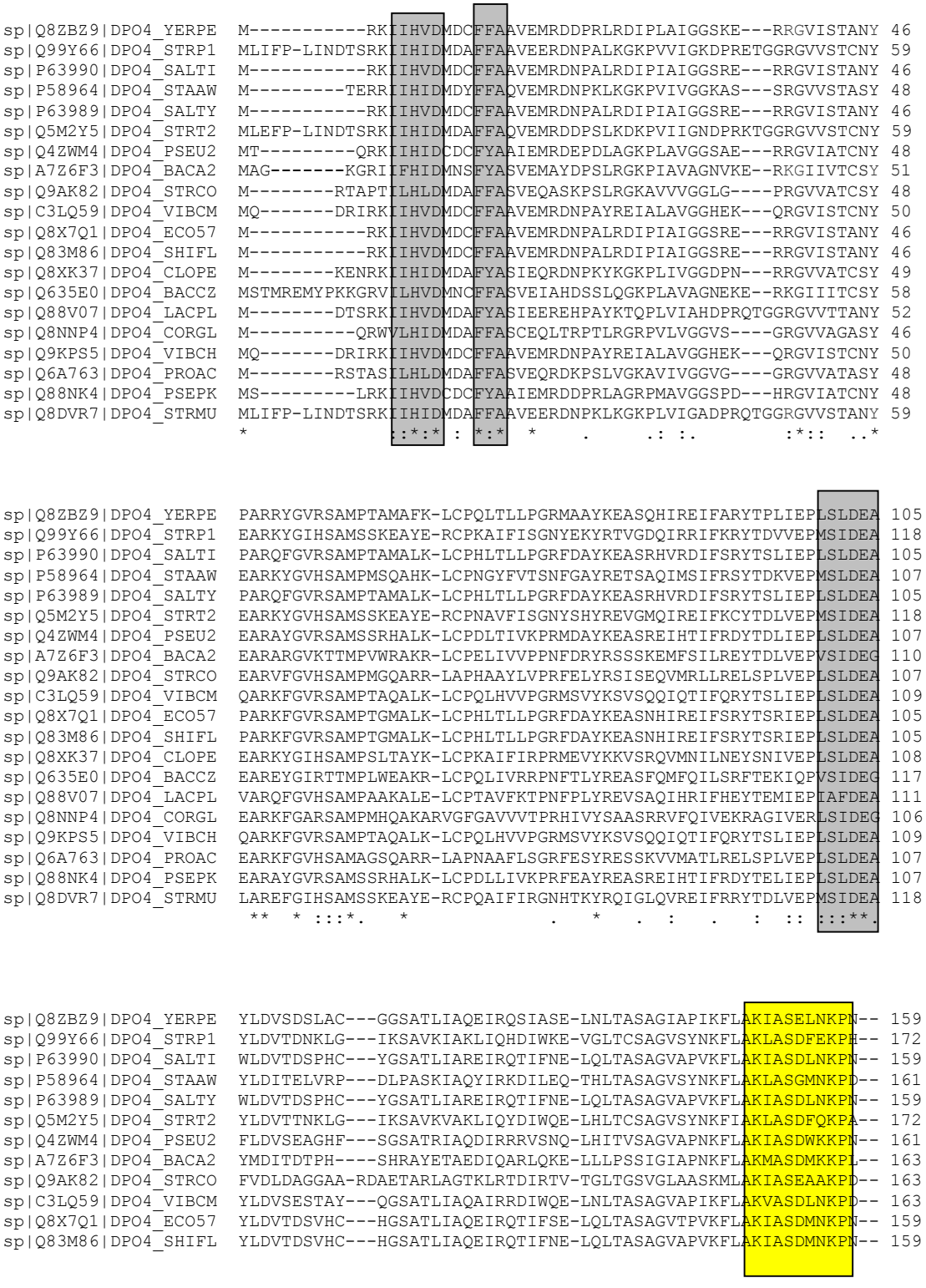


Fig. 2C. CLUSTAL format for T-COFFEE Version of polymerase III (alpha, catalytic subunit) (Prokaryotes)

The possible catalytic region is underlined and shaded yellow.
Only regions showing highly conserved blocks are shown.

Fig. 2D Prokaryotic DNA polymerases (pol IV)



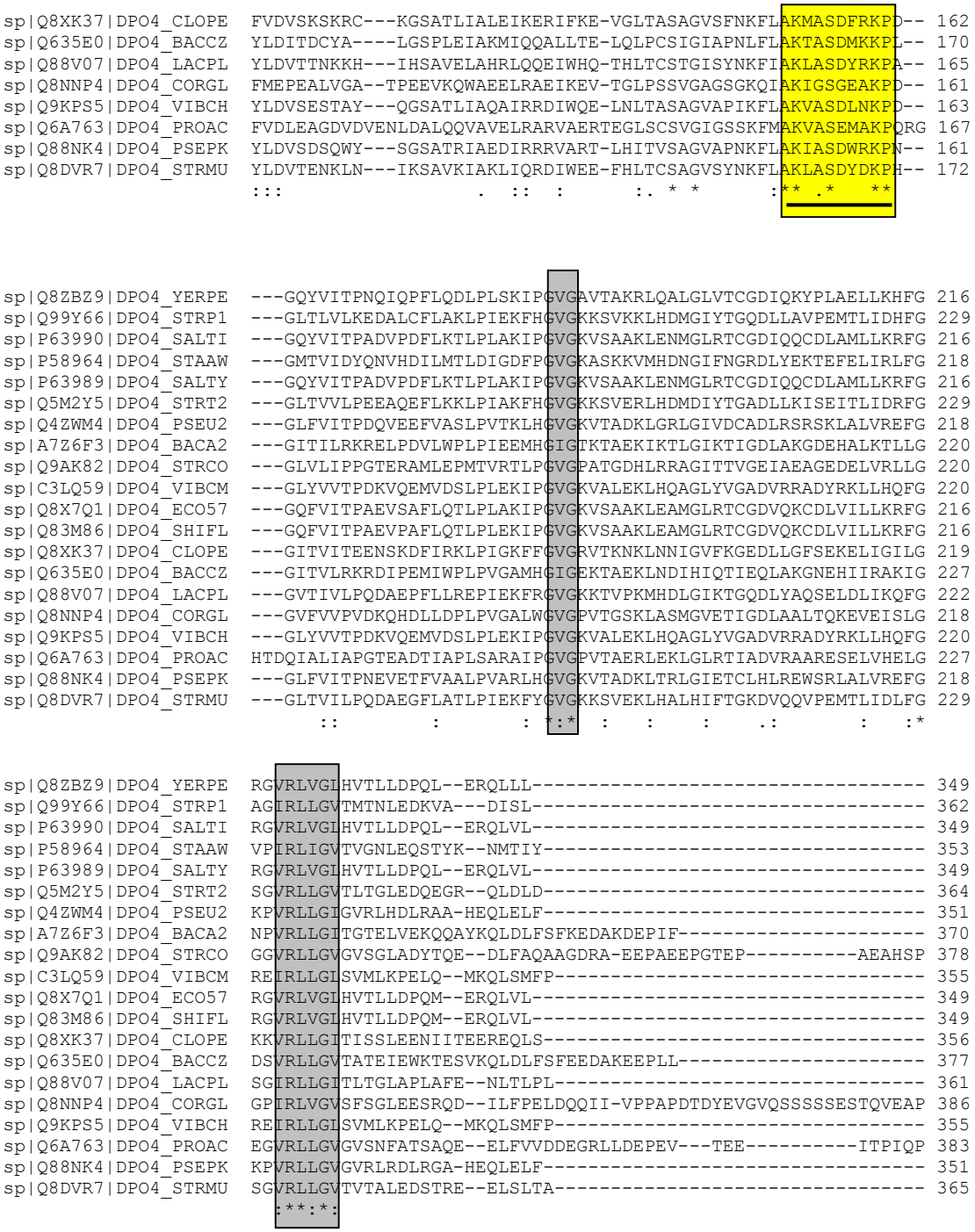


Fig. 2D. CLUSTAL format for T-COFFEE version of polymerase IV (Family Y- prokaryotes)

The possible catalytic region is underlined and shaded yellow.
Only regions showing highly conserved blocks are shown.

Fig. 2E Prokaryotic DNA polymerases (pol V- umuC, catalytic subunit)

```
gi|291617317|ref|YP_003520059.1|     ---------MSLDTFGREIKTRIKKEVHLNVCVGIAPTKTLAKLANHAAKKWS-KTGGVL 163
gi|290792075|gb|ADD63401.1|          ---------MGLEDFGRQLRQHVYDCTRLTIGVGAGPTKTLAKSAQWASKEWK-QFRGVL 162
gi|52841464|ref|YP_095263.1|         ---------DSYDSFCEQLQKKILKHTGIPTSIGIGPTKTLAKAANHLCKKVY-KI-PVF 166
gi|242376964|emb|CAQ31686.1|         ---------RDLTDFGREIRATVLQRTHLTVGVGIAQTKTLAKLANHAAKKWQRQTGGVV 164
gi|15672570|ref|NP_266744.1|         FYPDIRNRYEQMNLITLDLQREILDKLGLYVTVGMGDNPLLAKLAMDNYAKHNQNMRALI 232
gi|6491876|gb|AAF14056.1|AF039838_1  ----------------------------------------AKLANHAAKKWQEQTGGVV 19
gi|292488580|ref|YP_003531464.1|     ---------VSLADFGHEMRNQVLRNTGLTVGVGIAQTKTLAKLANHAAKKWP-ATGGVV 163
gi|150024494|ref|YP_001295320.1|     ---------YDLNDYCLTIRQRVKKVVGMPVSIGVAPTKALSKVANKIAKKFTDRTNGVY 162
gi|291554011|emb|CBA21065.1|         ---------VSLADFGHEMRNQVLRNTGLTVGVGIAQTKTLAKLANHAAKKWP-ATGGVV 163
                                                                  .* *      :       :

gi|291617317|ref|YP_003520059.1|     DLS--NRERQKKLLSLVPVADIWGVGRRLSQRLEEMGIKTADDLACQSTSLIRKNFSVVL 221
gi|290792075|gb|ADD63401.1|          ALTRGNPQRTRKLLSLQPVEEIWGVGNRIARKLNVLGIKTALDLALTNPAFIRKNFSVVL 222
gi|52841464|ref|YP_095263.1|         NIT----SNRGRLLQQISVGDIWGVGRQWANKLISRGIHTAYDLAMTNPHLLKKCFNAVL 222
gi|242376964|emb|CAQ31686.1|         DLS--NLERQRKLMSALPVDDVWGIGRRISKKLDAMGIKTVLDLADTDIRFIRKHFNVVL 222
gi|15672570|ref|NP_266744.1|         R--YE--DVPSKLWTLPKMTDFWGIGKRTEKRLNKLGISSIKELTNADPLLLKQKLGTIG 288
gi|6491876|gb|AAF14056.1|AF039838_1  DLS--NIERQRKLMAALPVDEVWGIGRRIGKKLDVMGIKTVLDLADTDIRFIRKHFSVVL 77
gi|292488580|ref|YP_003531464.1|     DLT--HIDRQRKLLALVAVEDVWGVGRRISKKLNLMGIETALDLAESSLWVIRKHFNVVL 221
gi|150024494|ref|YP_001295320.1|     VID--TEEKRIKALKWTKIEDVWGIGRQISKRLQAIKVFTAYDFVNLNDNYVKTNFSVVE 220
gi|291554011|emb|CBA21065.1|         DLT--HIDRQRKLLALVAVEDVWGVGRRISKKLNLMGIETALDLAESSLWVIRKHFNVVL 221
                                           .   :      : :.**:*.:  .:*    : :  ::.  .   ::  :..:
```

Fig. 2E. CLUSTAL format for T-COFFEE Version of polymerase V (Family Y- prokaryotes)

The possible catalytic region is underlined and shaded yellow.
Only regions showing highly conserved blocks are shown.

Fig. 2F Viral polymerases

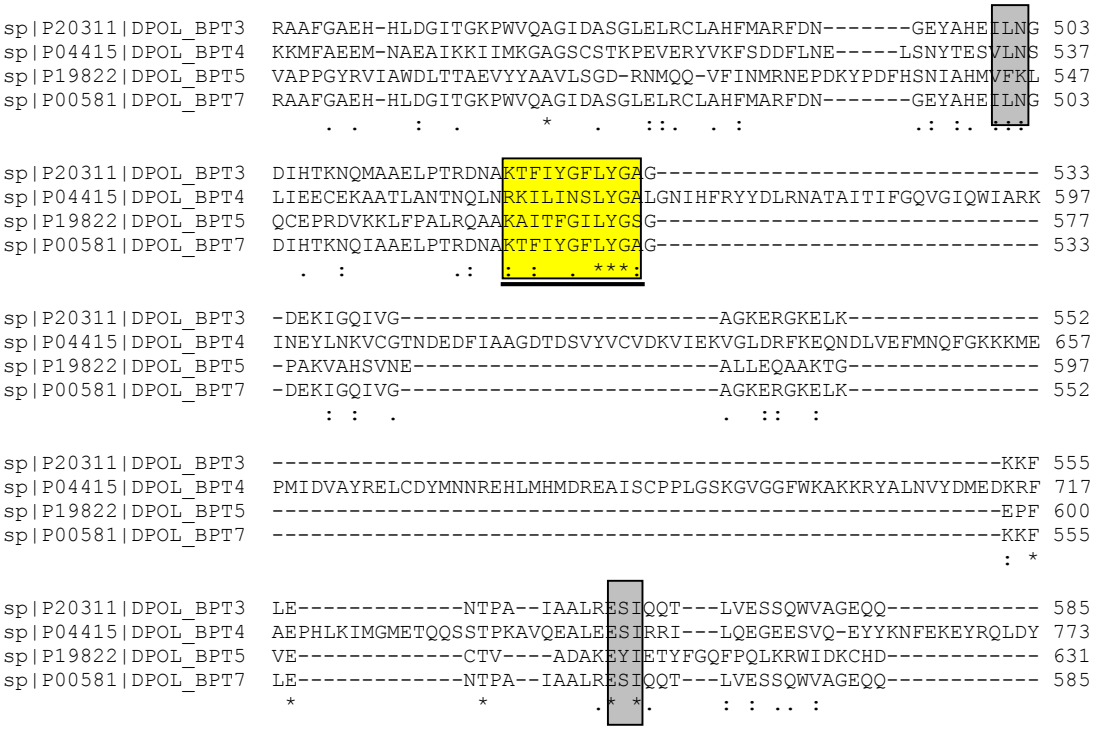
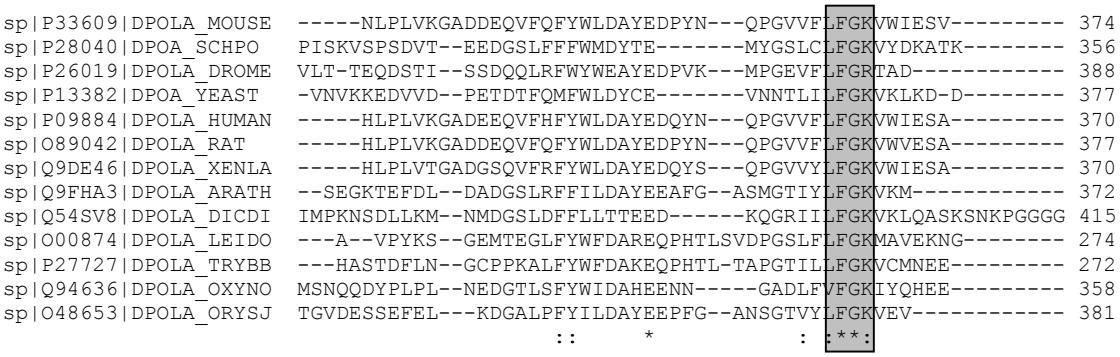


Fig. 2F. CLUSTAL format for T-COFFEE Version of polymerases of viral origin

The possible catalytic region is underlined and shaded yellow.
Only regions showing highly conserved blocks are shown.

Fig. 2G Alpha polymerases (Eukaryotes)



```
sp|P33609|DPOLA_MOUSE  S-EYLEVRYSAEV-----PQLPQNLKGETFSHVFGTNTSSLELFLMNRKIKGPCWLEVKN 507
sp|P28040|DPOA_SCHPO   A-DYLEVIYSYSYP-----ALPTDLTGSSFSHVFGTNTALFEQFVLSRRVMGPCWLKIQQ 484
sp|P26019|DPOLA_DROME  C-DYLEVHYDGKKPL---PNLSADKKYNSIAHIFGATTNALERFLLDRKIKGPCWLQVTG 523
sp|P13382|DPOA_YEAST   S-DYLKVLLPYQTPKSSRDTIPSDLSSDTFYHVFGGNSNIFESFVIQNRIMGPCWLDIKG 501
sp|P09884|DPOLA_HUMAN  S-EYLEVKYSAEM-----PQLPQDLKGETFSHVFGTNTSSLELFLMNRKIKGPCWLEVKS 503
sp|O89042|DPOLA_RAT    S-EYLEVRYSAEV-----PQLPQNLKGETFSHVFGTNTSSLELFLMNRKIKGPCWLEVKN 510
sp|Q9DE46|DPOLA_XENLA  S-EYLEVRYSADS-----PQLPQDLKGETFSHVFGTNTSSLELFLLSRKIKGPSWLEIKS 503
sp|Q9FHA3|DPOLA_ARATH  EQYVLKINYSFKD-----RPLPEDLKGESFSALLGSHTSALEHFILKRKIMGPCWLKISS 522
sp|Q54SV8|DPOLA_DICDI  QHYVWKLSYPSNQM-----VFPNDIKGSTFRCAYGITSSPVELFLIKRKIMGPTWLTVSG 569
sp|O00874|DPOLA_LEIDO  KTQWAKLRYPGRY-----PPLNAKGPFRHILIIMGASSSLLELFLIKRKLKGPSFLRISG 395
sp|P27727|DPOLA_TRYBB  KNRWAKLVYPGRY-----PPFPNKGGLTHVQVVVGASRSLLELFLIKKRLMGPSYLEIEH 393
sp|Q94636|DPOLA_OXYNO  H-QFVKIKYDATF-----PSLPSTVQGNTFECIFGANQSMLESFILKRKIRGPCWMTIRN 491
sp|O48653|DPOLA_ORYSJ  EQYVLKINYPYKD-----PALPTDLRGQHFHALLGTNNSALELLLIKRKIKGPSWLSISK 530
                            ::            :        .    *     .* :::.:: ** :: :

sp|P33609|DPOLA_MOUSE  VLPLALQITNIAGNIMSRTLMGGRSERNEFLLLHAFYENNYIVPDKQIF-RKPQQKLG-- 821
sp|P28040|DPOA_SCHPO   MLQLSKNLTNIAGNSWARTLTGTRAERNEYILLHEFKKNGYIVPDKQQSIRRHAEAFG-- 808
sp|P26019|DPOLA_DROME  IMPLALQITNICGNTMTRTLQGGRSERNEFLLLHAFHEKNYIVPDKKPV-SKRSGAG--- 838
sp|P13382|DPOA_YEAST   LLTLTKQLTNLAGNAWAQTLGGTRAGRNEYILLHEFSRNGFIVPDKEGN-RSRAQKQR-- 826
sp|P09884|DPOLA_HUMAN  VLPLALQITNIAGNIMSRTLMGGRSERNEFLLLHAFYENNYIVPDKQIF-RKPQQKLG-- 817
sp|O89042|DPOLA_RAT    VLPLALQITNIAGNIMSRTLMGGRSERNEFLLLHAFYENNYIVPDKQIF-RKPQQKPG-- 824
sp|Q9DE46|DPOLA_XENLA  VLPLALQITNIAGNVMSRTLMGGRSERNEYILLHAFTENNFIVPDKPVF-KKMQQTTV-- 817
sp|Q9FHA3|DPOLA_ARATH  VLPLTLQLTNISGNLWGKTLQGARAQRIEYYLLHTFHSKKFILPDKISQ-RMKEIKSSKR 861
sp|Q54SV8|DPOLA_DICDI  VFPLTKQLTNLAGNQWDKSLKSNRAERIEYILLHNFHEKKYLLPDKIYQ-KSSSSG-G-- 876
sp|O00874|DPOLA_LEIDO  VIRLTKRLTTIAGNLWSRTLFGARSERIEYILLHTFHDLKFITPDRYVQ-NFKRGRD--- 719
sp|P27727|DPOLA_TRYBB  VIPLTKRLTLLAGNLWSRTLYGARSERIEYILLHAFHNLKFVTPDKKKR-DLKRGRE--- 717
sp|Q94636|DPOLA_OXYNO  VIPLTKQLTNIAGNLWFRSLQNARAERNEMILLHEFKKKKFVLPDKKQLN-AKDLKKNM- 818
sp|O48653|DPOLA_ORYSJ  VLPLTRQLTNISGNLWGKTLQGSRAQRVEYILLHAFHARKFIVPDKFAR-S-KEFNSTKR 869
                       :: *: .:* :.**   ::* . *: * *  *** *    :: **:

sp|P33609|DPOLA_MOUSE  --------------DEDEEIDGDTNKYKKGRKKATYAGGLVLDPKVGFYDKFILLLDFNS 867
sp|P28040|DPOA_SCHPO   --------------AED----G-LQEESLGKKKDKYKGGLVFEPQKGLYETCILVMDFNS 849
sp|P26019|DPOLA_DROME  --------------DTDATLSGA-DATMQTKKKAAYAGGLVLEPMRGLYEKYVLLMDFNS 883
sp|P13382|DPOA_YEAST   --------------QNE----E-NADAPVNSKKAKYQGGLVFEPEKGLHKNYVLVMDFNS 867
sp|P09884|DPOLA_HUMAN  --------------DEDEEIDGDTNKYKKGRKKAAYAGGLVLDPKVGFYDKFILLLDFNS 863
sp|O89042|DPOLA_RAT    --------------DEDEEIDGDTNKYKKGRKKAAYAGGLVLDPKVGFYDKFILLLDFNS 870
sp|Q9DE46|DPOLA_XENLA  --------------EDNDDMGTDQN-KNKSRKKAAYAGGLVLEPKVGFYDKFILLLDFNS 862
sp|Q9FHA3|DPOLA_ARATH  RMDYAPEDRNVDELDADLTLENDPSKGSKTKKGPAYAGGLVLEPKRGLYDKYVLLLDFNS 921
sp|Q54SV8|DPOLA_DICDI  --------------G---------GGAKDKDNHAAYSGGLVLDPKIDFYDRYVVLLDFNS 913
sp|O00874|DPOLA_LEIDO  --------------D---------EEEEDGKRKAKYQGGMVLDPKCGLYSDYILLLDFNS 756
sp|P27727|DPOLA_TRYBB  --------------D---------D-DDEGKRKTKYQGGMVLEPKSGLYSEYILLLDFNS 753
sp|Q94636|DPOLA_OXYNO  --------------FADEYEEGDGKTKGGKRKKAAYAGGLVIEPKAGFYDNIILLLDFNS 864
sp|O48653|DPOLA_ORYSJ  KMNPDTEAARPDEADPSIDDEGHHVDQGKTKKGPSYAGGLVLEPKKGLYDKYVLLLDFNS 929
                                                .   * **:*::*  .::.  ::::****

sp|P33609|DPOLA_MOUSE  LYPSIIQEFNICFTTVQRVTSEVQKAT---------E-DEE--QEQIP------------ 903
sp|P28040|DPOA_SCHPO   LYPSIIQEYNICFTTVDRSPSNS----------------DS--DDQIP------------ 879
sp|P26019|DPOLA_DROME  LYPSIIQEYNICFTTVQQPVDA---------------------DELP------------ 909
sp|P13382|DPOA_YEAST   LYPSIIQEFNICFTTVDRNKE-------------------D--IDELP------------ 894
sp|P09884|DPOLA_HUMAN  LYPSIIQEFNICFTTVQRVASEAQKVT---------E-DGE--QEQIP------------ 899
sp|O89042|DPOLA_RAT    LYPSIIQEFNICFTTVQRVASETLKAT---------E-DEE--QEQIP------------ 906
sp|Q9DE46|DPOLA_XENLA  LYPSIIQEYNICFTTVHREAPSTQK-------------GED--QDEIP------------ 895
sp|Q9FHA3|DPOLA_ARATH  LYPSIIQEYNICFTTIPRSED-G-----------------------VP------------ 945
sp|Q54SV8|DPOLA_DICDI  LYPSIIQEYNVCFTTINRVKRDD----------------GK--WEE-A------------ 942
sp|O00874|DPOLA_LEIDO  LYPSLIQEFNICFTTVDRESGSE--------------------IDVPPPENLICASCAA 795
sp|P27727|DPOLA_TRYBB  LYPSLIQEFNVCYTTIDRDENTV----------------S----AEVPPPESLICLSCRA 793
sp|Q94636|DPOLA_OXYNO  LYPSIIQEYNLCFTTVNRRPTKNFDGSEMKNQYKKGENGEEEVDIEEA------------ 912
sp|O48653|DPOLA_ORYSJ  LYPSIIQEYNICFTTVDRSADGN----------------------VP------------ 954
                       ****:***:*:*:**: :                           .
```

```
sp|P33609|DPOLA_MOUSE  -ELPDPN--LEMGILPREIRKLVERRKQVKQLMKQQDLNPDLVLQYDIRQKALKLTANSM 960
sp|P28040|DPOA_SCHPO   -DTPSAS--ANQGIFPRLIANLVERRRQIKGLLKDNSATPTQRLQWDIQQQALKLTANSM 936
sp|P26019|DPOLA_DROME  -TLPDSK--TEPGILPLQLKRLVESRKEVKKLMAAPDLSPELQMQYHIRQMALKLTANSM 966
sp|P13382|DPOA_YEAST   -SVPPSE--VDQGVLPRLLANLVDRRREVKKVMKTET-DPHKRVQCDIRQQALKLTANSM 950
sp|P09884|DPOLA_HUMAN  -ELPDPS--LEMGILPREIRKLVERRKQVKQLMKQQDLNPDLILQYDIRQKALKLTANSM 956
sp|O89042|DPOLA_RAT    -ELPDPN--LDMGILPREIRKLVERRKQVKQLMKQQDLNPDLVLQYDIRQKALKLTANSM 963
sp|Q9DE46|DPOLA_XENLA  -ELPHSD--LEMGILPREIRKLVERRRHVKQLMKQPDLNPDLYLQYDIRQKALKLTANSM 952
sp|Q9FHA3|DPOLA_ARATH  -RLPSS---QTPGILPKLMEHLVSIRKSVKLKMKKET--GLKYWELDIRQQALKLTANSM 999
sp|Q54SV8|DPOLA_DICDI  -MPPPSS--IEKGILPKVLHGLVSKRREIKKRMEQEK-NKIIKAQYDIQQQAVKLIANSM 998
sp|O00874|DPOLA_LEIDO  AGLSAPC--LHKCVLPKVIKSLVDSRREVKRLMKIEK-DANNLALLEIRQKALKLTANSM 852
sp|P27727|DPOLA_TRYBB  AGLPSPC--LHKCILPKVIRGLVDSRREIKRMMKSEK-DPGNLAMLEIRQLALKLTANSM 850
sp|Q94636|DPOLA_OXYNO  -DLPDKNVNLKDAVLPMVLRDLVQKRKAVKDKMKTEK-DHVKLSQLEIRQKAIKLTANSM 970
sp|O48653|DPOLA_ORYSJ  -NLPAS---KTTGVLPELLKSLVERRRMVKSWLKTAS--GLKRQQFDIQQQALKLTANSM 1008
                          .         ::*  :  **. *: :*  :            .*:* *:** ****

sp|P33609|DPOLA_MOUSE  YGCLGFSYSRFYAKPLAALVTYKGREILMHTKDMVQKM--NLEVIYGDTDSIMINTNST- 1017
sp|P28040|DPOA_SCHPO   YGCLGYTKSRFYARPLAVLITYKGREALMNTKELADQM--GLQVIYGDTDSVMLNTNVT- 993
sp|P26019|DPOLA_DROME  YGCLGFAHSRFFAQHLAALVTHKGREILTNTQQLVQKM--NYDVVYGDTDSLMINTNIT- 1023
sp|P13382|DPOA_YEAST   YGCLGYVNSRFYAKPLAMLVTNKGREILMNTRQLAESM--NLLVVYGDTDSVMIDTGCD- 1007
sp|P09884|DPOLA_HUMAN  YGCLGFSYSRFYAKPLAALVTYKGREILMHTKEMVQKM--NLEVIYGDTDSIMINTNST- 1013
sp|O89042|DPOLA_RAT    YGCLGFSYSRFYAKPLAALVTYKGREILMHTKEMVQKM--NLEVIYGDTDSIMINTNST- 1020
sp|Q9DE46|DPOLA_XENLA  YGCLGFSYSRFYAKPLAALVTHQGREILLHTKEMVQKM--NLEVIYGDTDSIMINTNCN- 1009
sp|Q9FHA3|DPOLA_ARATH  YGCLGFSNSRFYAKPLAELITLQGRDILQRTVDLVQN-HLNLEVIYGDTDSIMIHSGLD- 1057
sp|Q54SV8|DPOLA_DICDI  YGCLGFSHSRFYALPLAELVTRKGRENLQKGASIVNKM--CYDVIYGDTDSLMIYTGVG- 1055
sp|O00874|DPOLA_LEIDO  YGCLGFEYSRFHAQPLAELVTRQGRLALQSTVDLIPQLNPSLRVIYGDTDSVMIQTGIKN 912
sp|P27727|DPOLA_TRYBB  YGCLGFEYSRFYAQPLAELVTRQGRLALQNTVELIPQISPSIRVIYGDTDSVMIQTGIKD 910
sp|Q94636|DPOLA_OXYNO  YGCLGFGSSRFHAQAIAALITRTGRETLLRTKDIAEN-KLGFNVVYGDTDSIMINTGSN- 1028
sp|O48653|DPOLA_ORYSJ  YGCLGFSNSRFYAKPLAELITLQGREILQNTVDLVQN-NLNLEVIYGDTDSIMIHTGLD- 1066
                       *****:  ***.*  :* *:*  **  *    .:  .      *:******:*: :.

sp|P33609|DPOLA_MOUSE  NLEEVFKLGNKVKSEVNKLYK--LLEIDIDAVFKSLLLLKKKYAALVVEPT-SDGNYIT 1074
sp|P28040|DPOA_SCHPO   DKNHALRIGNEFKEKVNERYS--KLEIDIDNVYQRMLLHAKKKYAALQLDSQ---G--KP 1046
sp|P26019|DPOLA_DROME  DYDQVYKIGHNIKQSVNKLYK--QLELDIDGVFGCLLLLKKKYAAIKLSKD-SKGNLRR 1080
sp|P13382|DPOA_YEAST   NYADAIKIGLGFKRLVNERYR--LLEIDIDNVFKKLLLHAKKKYAALTVNLD-KNGNGTT 1064
sp|P09884|DPOLA_HUMAN  NLEEVFKLGNKVKSEVNKLYK--LLEIDIDGVFKSLLLLKKKYAALVVEPT-SDGNYVT 1070
sp|O89042|DPOLA_RAT    NLEEVFKLGNKVKNEVNKLYK--LLEIDIDGVFKSLLLLKKKYAALVVEPT-SDGNYIT 1077
sp|Q9DE46|DPOLA_XENLA  NLEEVFKLGNRVKSEINKSYK--LLEIDIDGIFKSLLLLKKKYAALTVEPT-GDGKYVT 1066
sp|Q9FHA3|DPOLA_ARATH  DIEEVKAIKSKVIQEVNKKYR--CLKIDCDGIYKRMLLLRKKKYAAVKLQ-F-KDGKPCE 1113
sp|Q54SV8|DPOLA_DICDI  TFNEAETIGKEIQKKINDQYRGSVMEIGLDGIFKRLLLFKKKKYACLKEFRI-DSTTTKC 1114
sp|O00874|DPOLA_LEIDO  DIKAVRDLGLDLKAKINKRYQ--SLEIDIDGVFRAILLLKKKKYAALTVTDW-QGEGKTY 969
sp|P27727|DPOLA_TRYBB  DIVKVRNLGFEIKGKVNQRYQ--SLELDIDGVFRAMLLLRKKKYAALSVVDW-QGEGKVY 967
sp|Q94636|DPOLA_OXYNO  QLQQALEMGKRLKGEVNCLYK--CLEIEIDGVFKSLLLLKKKYAALKYENFLSPAEVKV 1086
sp|O48653|DPOLA_ORYSJ  DISRAKGIAGKVIQEVNKKYR--CLEIDLDGIYKRMLLLKKKYAAIKVA-L--DGSLRE 1121
                          .  :   .   :*  *    :::  * ::  :**  *****.:
```

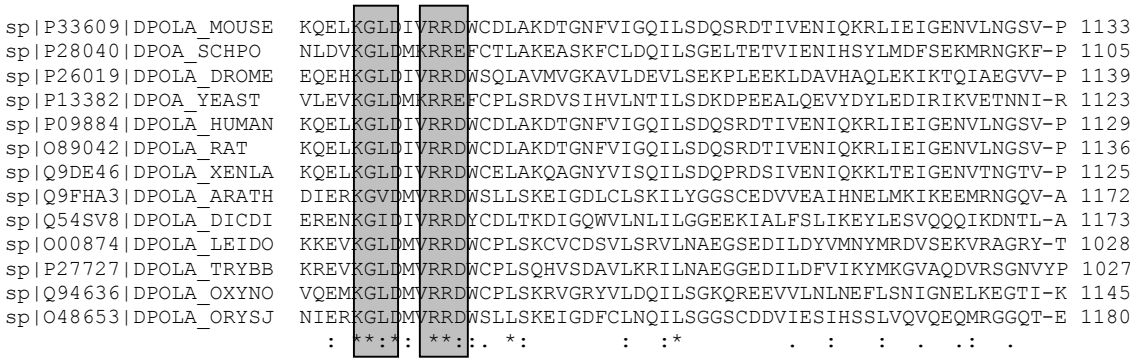


Fig. 2G. CLUSTAL format for T-COFFEE version of alpha polymerases (Eukaryotes)

The possible catalytic region is underlined and shaded yellow.
Only regions showing highly conserved blocks are shown.

Fig. 2H Beta polymerases (Eukaryotes)

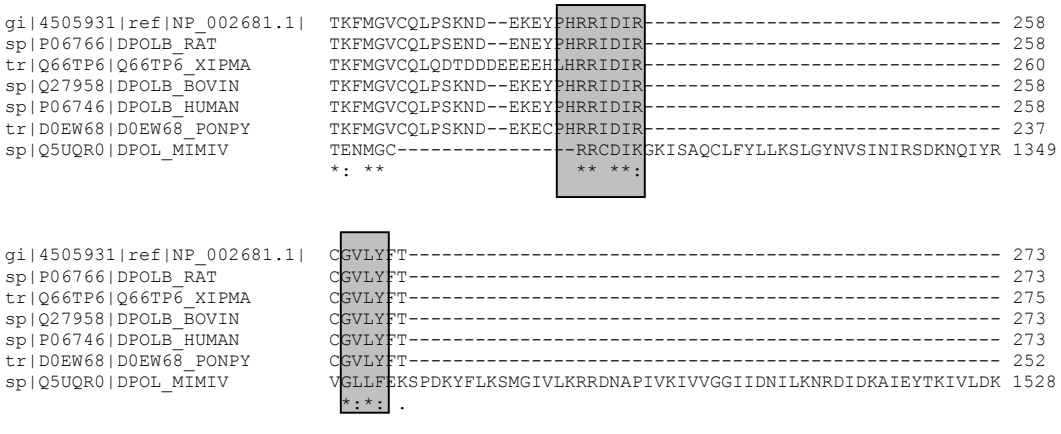
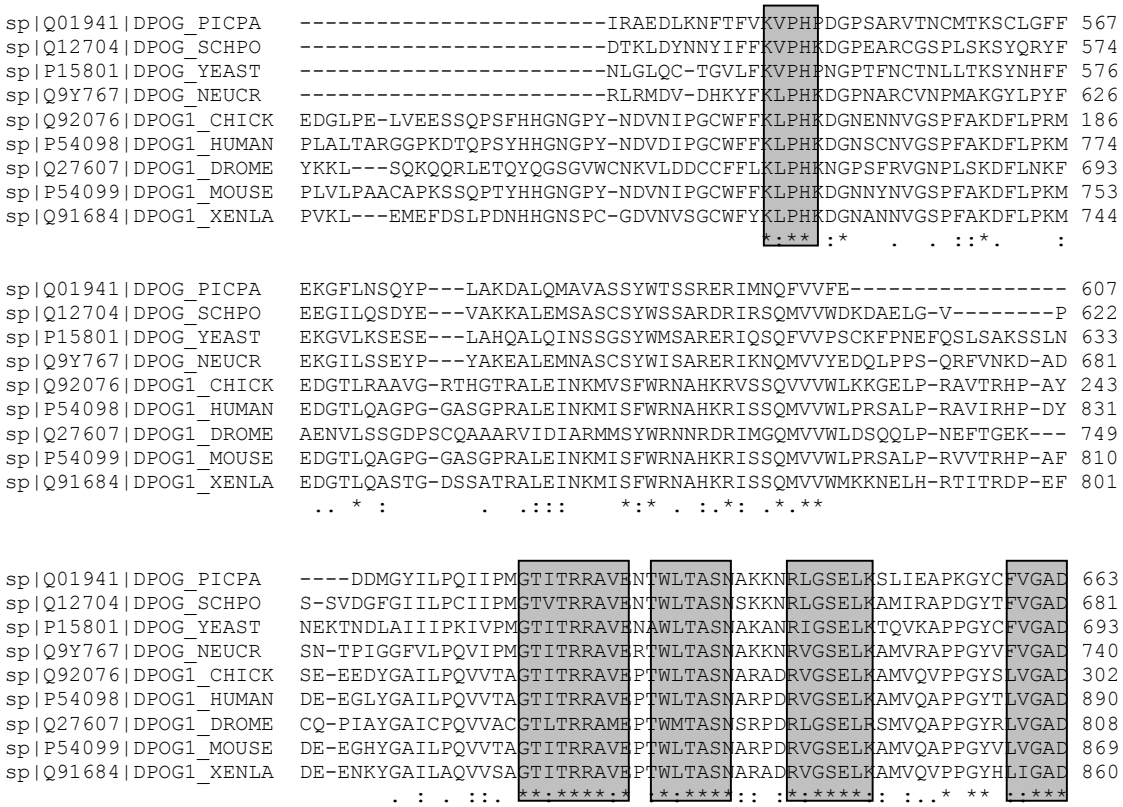


Fig. 2H. CLUSTAL format for T-COFFEE version of beta polymerases (Eukaryotes)

The possible catalytic region is underlined and shaded yellow.
Only regions showing highly conserved blocks are shown.

Fig. 2I Gamma polymerases (Eukaryotes, Mitochondrial)



```
sp|Q01941|DPOG_PICPA   VDSEELWIASLIGDSVF-KIHGGTAIGWMTLEGTKNEGTDLHSKTAKILGISRNEAKIFN 722
sp|Q12704|DPOG_SCHPO   VDSEELWIVALMGDSQF-RLHGATALGMMTLEGKKSEGTDLHSKTAAILGVSRDSAKVFN 740
sp|P15801|DPOG_YEAST   VDSEELWIASLVGDSIF-NVHGGTAIGWMCLEGTKNEGTDLHTKTAQILGCSRNEAKIFN 752
sp|Q9Y767|DPOG_NEUCR   VDSEELWIASVVGDATF-KLHGGNAIGFMTLEGTKSQGTDLHSRTASILGITRNDAKVFN 799
sp|Q92076|DPOG1_CHICK  VDSQELWIAAVLGEAHFAGMHGCTAFGWMTLQGKKSDGTDLHSKTAATVGISREHAKVFN 362
sp|P54098|DPOG1_HUMAN  VDSQELWIAAVLGDAHFAGMHGCTAFGWMTLQGRKSRGTDLHSKTATTVGISREHAKIFN 950
sp|Q27607|DPOG1_DROME  VDSQELWIASVLGDAYACGEHGATPLGWMTLSGSKSNGSDMHSITAKAVGISRDHAKVIN 868
sp|P54099|DPOG1_MOUSE  VDSQELWIAAVLGDAHFAGMHGCTAFGWMTLQGRKSRGTDLHSKTAATVGISREHAKIFN 929
sp|Q91684|DPOG1_XENLA  VDSQELWIAAILGEAHFAGIHGCTAFGWMTLQGKKSSGTDLHSKTASTVGISREHAKVFN 920
                       ***:****.:::*::     ** ..:* * *.* *. *:*:*: **  :* :*: **::*

sp|Q01941|DPOG_PICPA   YGRIYGAGIKFTTTLLKKFNPALSDAEAKATANALYTATKGIS-GR-------------- 767
sp|Q12704|DPOG_SCHPO   YGRLYGAGLKHTTLLLMQMNPTLKTAEAKELAKKLYASTKGVKSKM-------------- 786
sp|P15801|DPOG_YEAST   YGRIYGAGAKFASQLLKRFNPSLTDEETKKIANKLYENTKGKT-KR-------------- 797
sp|Q9Y767|DPOG_NEUCR   YGRIYGAGLKFASQLLRQFNPSLTEAETTAIATKLYDATKGAKTNR-------------- 845
sp|Q92076|DPOG1_CHICK  YGRIYGAGQPFAERLLMQFNHRLTQQQAREKAQQMYAVTKGIRRFHLSEEGEWLVKELEL 422
sp|P54098|DPOG1_HUMAN  YGRIYGAGQPFAERLLMQFNHRLTQQEAAEKAQQMYAATKGLRWYRLSDEGEWLVRELNL 1010
sp|Q27607|DPOG1_DROME  YARIYGAGQLFAETLLRQFNPTFSASEAKAKAMKMFSITKGKRVYRLREEFHDELED--- 925
sp|P54099|DPOG1_MOUSE  YGRIYGAGQSFAERLLMQFNHRLTRQEAAEKAQQMYAVTKGLRRYRLSADGEWLVKQLNL 989
sp|Q91684|DPOG1_XENLA  YGRIYGAGQPFAERLLMQFNHRLTQEQAAEKAKQMYAVTKGIRRYILSKEGEWLVEELGI 980
                       *.*:****  .:  ** ::*  :.  ::   *  ::  ***

sp|Q01941|DPOG_PICPA   ------------------------Y---------DKKSIWYGGSESIIFNRLEAIAEMAH 794
sp|Q12704|DPOG_SCHPO   ------------------------SK-RLQEMGLPKLTFWSQGTESFVFNKLEAMAQLPS 821
sp|P15801|DPOG_YEAST   ------------------------SK--------LFKKFWYGGSESILFNKLESIAEQET 825
sp|Q9Y767|DPOG_NEUCR   ------------------------KS--------LYKRPFWRGGTESFVFNMLEEFAEQER 874
sp|Q92076|DPOG1_CHICK  AVDKAEDGTVSAQDVQKIQREAMRKSRRKKKWDVVAHRMWAGGTESEMFNKLESIALSAS 482
sp|P54098|DPOG1_HUMAN  PVDRTEGGWISLQDLRKVQRETARKSQW-KKWEVVAERAWKGGTESEMFNKLESIATSDI 1069
sp|Q27607|DPOG1_DROME  -------RAYSSYEA---SRLAIQRN--RTLAEVFHRPNWQGGTESAMFNRLEEIATGSQ 973
sp|P54099|DPOG1_MOUSE  PVDRTEDGWVSLQDLRMIRREASRKSRW-KKWEVAAERAWTGGTESEMFNKLESIAMSDT 1048
sp|Q91684|DPOG1_XENLA  SVERGEENSVNLQDLRKIQKDATKRSR--RKWNLVSRRIWTGGTESQMFNKLETIAMSPS 1038
                                                  *  *:** :** ** :*

sp|Q01941|DPOG_PICPA   PKTPVLGAGITAPLQKANL--STNNFLTSRINWAIQSSGVDYLHLLIISMDYLIKLFDID 852
sp|Q12704|DPOG_SCHPO   PRTPVLDAGITQALSSKNL--SKNSFMTSRVNWAIQSSAVDYLHLLLVSMNHLIKKYYLE 879
sp|P15801|DPOG_YEAST   PKTPVLGCGITYSLMKKNL--RANSFLPSRINWAIQSSGVDYLHLLCCSMEYIIKKYNLE 883
sp|Q9Y767|DPOG_NEUCR   PRTPVLGAGITEALMSRWV--SKGGFLTSRINWAIQSSGVDYLHLLIIAMDYLTRRFNLA 932
sp|Q92076|DPOG1_CHICK  PQTPVLGCHISRALEPAV---AKGEFLTSRVNWVVQSSAVDYLHLMLVSMKWLFEEYDIN 539
sp|P54098|DPOG1_HUMAN  PRTPVLGCCISRALEPSA---VQEEFMTSRVNWVVQSSAVDYLHLMLVAMKWLFEEFAID 1126
sp|Q27607|DPOG1_DROME  PRTPFLGGRLSRALEADTGPEQEQRFLPTRINWVVQSGAVDFLHLMLVSMRWLMGS---H 1030
sp|P54099|DPOG1_MOUSE  PRTPVLGCCISRALEPSV---VQGEFITSRVNWVVQSSAVDYLHLMLVAMKWLFEEFAID 1105
sp|Q91684|DPOG1_XENLA  PKTPVLGCRISRALEPTA---VKGEFITSRVNWVVQSSAVDYLHLMLVAMKWLFEAYDID 1095
                       *:**.*  :: .*          *:.:*:**.:**..**:***:  :*  :

sp|Q01941|DPOG_PICPA   ARLCITVHDEIRYLVKEEDKFRAAYALQISNLWTRAMFCQQLGINEVPQSCAFFSAVDLD 912
sp|Q12704|DPOG_SCHPO   ARLSLTVHDEVRYLSSDKDKYRVAFALQVANLWTRAFFCQRLGINELPQSVAFFSSVDID 939
sp|P15801|DPOG_YEAST   ARLCISIHDEIRFLVSEKDKYRAAMALQISNIWTRAMFCQQMGINELPQNCAFFSQVDID 943
sp|Q9Y767|DPOG_NEUCR   CRLAITVHDEIRYLAEEPDKYRVAMALQIANLWTRVMFAQQVGIQDLPQSCAFFSAVDID 992
sp|Q92076|DPOG1_CHICK  GRFCISIHDEVRYLVQEQDRYRAALALQITNLLTRCMFAYKLGLQDLPQSVAFFSAVDID 599
sp|P54098|DPOG1_HUMAN  GRFCISIHDEVRYLVREEDRYRAALALQITNLLTRCMFAYKLGLNDLPQSVAFFSAVDID 1186
sp|Q27607|DPOG1_DROME  VRFCLSFHDELRYLVKEELSPKAALAMHITNLMTRSFCVSRIGLQDLPMSVAFFSSVEVD 1090
sp|P54099|DPOG1_MOUSE  GRFCISIHDEVRYLVREEDRYRAALALQITNLLTRCMFAYKLGLNDLPQSVAFFSAVDID 1165
sp|Q91684|DPOG1_XENLA  GRFCISIHDEVRYLVHSKDRYRAALALQITNLLTRCMFASRLGIQDVPQSVAFFSAVDID 1155
                        *:.::.***:*:*  .    :.* *::::*: ** :   ::*::::* . **** *::*
```

```
sp|Q01941|DPOG_PICPA   FVLRKEVDLDCVTPSNPD------PIPCGKSLDIYQLLQQEDIKG-A--DFPRT---MHL 960
sp|Q12704|DPOG_SCHPO   HVLRKDVKMDCVTPSNKV------PIPPGEELTIESVLEKLEQSG-QS--LEPL---EQI 987
sp|P15801|DPOG_YEAST   SVIRKEVNMDCITPSNKT------AIPHGEALDINQLLDKSNSKL-GKPNLDID---SKV 993
sp|Q9Y767|DPOG_NEUCR   HVLRKEVDMDCITPSNPI------PIAHGESIDIFQILEKGDDAKLDDSIVPQSQYAPRL 1046
sp|Q92076|DPOG1_CHICK  RCLRKEVTMNCATPSNPTGMEKKYGIPRGEALDIYQIIEITKGSL-E------------- 645
sp|P54098|DPOG1_HUMAN  RCLRKEVTMDCKTPSNPTGMERRYGIPQGEALDIYQIIELTKGSL-EK--R--------- 1234
sp|Q27607|DPOG1_DROME  TVLRKECTMDCKTPSNPHGLRIGYGIQPGQSLSVAEAIEKAGGND-VS------------ 1137
sp|P54099|DPOG1_MOUSE  QCLRKEVTMDCKTPSNPTGMERRYGIPQGEALDIYQIIELTKGSL-EK--RKP------- 1215
sp|Q91684|DPOG1_XENLA  KCLRKEVTMDCSTPSNPNGMEKRYGIPQGEALDIYQILKVTKGV---------------- 1199
                         :**:  ::* ****         *  *: : : . :.
```

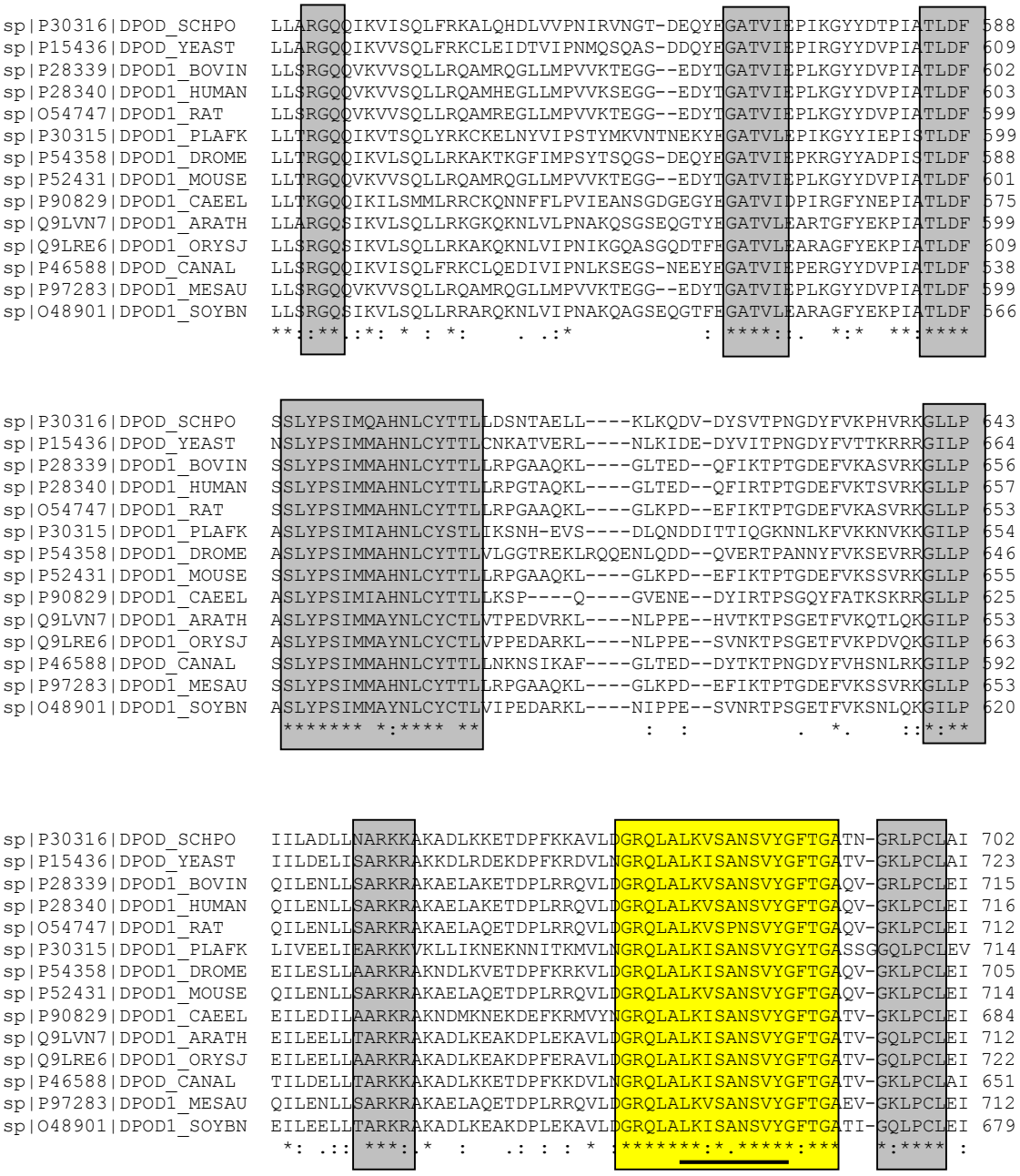
Fig. 2I. CLUSTAL format for T-COFFEE version of mitochondrial γ polymerases (eukaryotes)

The possible catalytic region is underlined and shaded yellow.
Only regions showing highly conserved blocks are shown.

Fig. 2J Delta polymerases (Eukaryotes)

```
sp|P30316|DPOD_SCHPO   PYFYVKAPVGFRPEM---LERFTQDLDATCN--------------G---GVIDHCIIEMK 176
sp|P15436|DPOD_YEAST   NYLYVPAPNSSDANDQEQINKFVHYLNETF---------------D---HAIDSIEVVSK 199
sp|P28339|DPOD1_BOVIN  PYFYTPAPPGFGPEH---LSELQRELSAAISRDQRG-G---KELTG---PAVLAVELCSR 194
sp|P28340|DPOD1_HUMAN  PYFYTPAPPGFGPEH---MGDLQRELNLAISRDSRG-G---RELTG---PAVLAVELCSR 195
sp|O54747|DPOD1_RAT    PYFYTPAPPGFGAEH---LSELQRELNAAISRDQRG-G---KELSG---PAVLAIELCSR 191
sp|P30315|DPOD1_PLAFK  PYFYVEKPDDFDNED---LIKLEMLMNENLNLNSQY------KIYE---KKILKIEIVKT 187
sp|P54358|DPOD1_DROME  PYFYIEAPSQFEEHH---CEKLQKALDQKVIADIRN-N---KDNVQ---EAVLMVELVEK 175
sp|P52431|DPOD1_MOUSE  PYFYTPAPPGFGAEH---LSELQQELNAAISRDQRG-G---KELSG---PAVLAIELCSR 193
sp|P90829|DPOD1_CAEEL  PHFYFQAPQGFGVEH---IGTAQSAICNMVAAAKRRGGSGQAQLPGKVVDNLVHVEIVHG 163
sp|Q9LVN7|DPOD1_ARATH  PYFYIACPPGMGPDD---ISNFHQSLEGRMRESNKN-A-----KVP---KFVKRIEMVQK 190
sp|Q9LRE6|DPOD1_ORYSJ  PYFYISCPMGMGPDD---ISRFHQTLEGRMKDSNRN-S-----NVP---RFVKRIELVQK 200
sp|P46588|DPOD_CANAL   HYFYCPVPKGFE-EN---LTEFTNYLKATF-------------------DGIERVEITSK 136
sp|P97283|DPOD1_MESAU  PYFYTPAPPGFGAEH---LSDLQRELSTAISRDQRG-G---KELSG---PAVLAIELCSR 191
sp|O48901|DPOD1_SOYBN  PYFYICCPPGMGPDD---ISHFHQTLEGRMREANRN-S-----NVG---KFVRRIEMVQR 157
                         ::*   *     .          :                        :    :

sp|P30316|DPOD_SCHPO   APLRIMSFDIECAGRKGV-FPDPSIDPVIQIASIVTQYGDST-PFVRNVFCVDTCSQIVG 349
sp|P15436|DPOD_YEAST   APLRIMSFDIECAGRIGV-FPEPEYDPVIQIANVVSIAGAKK-PFIRNVFTLNTCSPITG 370
sp|P28339|DPOD1_BOVIN  APLRVLSFDIECAGRKGI-FPEPERDPVIQICSLGLRWGEPE-PFLRLALTLRPCAPILG 364
sp|P28340|DPOD1_HUMAN  APLRVLSFDIECAGRKGI-FPEPERDPVIQICSLGLRWGEPE-PFLRLALTLRPCAPILG 365
sp|O54747|DPOD1_RAT    APLRVLSFDIECAGRKGI-FPEPERDPVIQICSLGLRWGEPE-PFLRLALTLRPCAPILG 361
sp|P30315|DPOD1_PLAFK  PKLRILSFDIECIKLDGKGFPEAKNDPIIQISSILYFQGEPIDNCTKFIFTLLECASIPG 359
sp|P54358|DPOD1_DROME  APFRILSFDIECAGRKGI-FPEAKIDPVIQIANMVIRQGERE-PFIRNVFTLNECAPIIG 349
sp|P52431|DPOD1_MOUSE  APLRVLSFDIECAGRKGI-FPEPERDPVIQICSLGLRWGEPE-PFLRLALTLRPCAPILG 363
sp|P90829|DPOD1_CAEEL  APIRTLSLDIECIGRRGV-FPEAIKDPIIQIANLVKIEGEAE-PFVRNCFVLGTCAPVVG 335
sp|Q9LVN7|DPOD1_ARATH  APFRVLSFDIECAGRKGH-FPEAKHDPVIQIANLVTLQGEDH-PFVRNVMTLKSCAPIVG 359
sp|Q9LRE6|DPOD1_ORYSJ  APFRILSFDIECAGRKGH-FPEPTHDPVIQIANLVTLQGEGQ-PFVRNVMTLKSCSPIVG 369
sp|P46588|DPOD_CANAL   APLRILSFDIECAGRKGV-FPEAEHDPVIQIANVVQKSGESK-PFVRNVFTVNTCSSIIG 299
sp|P97283|DPOD1_MESAU  APLRVLSFDIECAGRKGI-FPEPERDPVIQICSLGLRWGEPE-PFLRLALTLRPCAPILG 361
sp|O48901|DPOD1_SOYBN  APFRILSFDIECAGRKGH-FPEPTHDPVIQIANLVTLQGEDQ-PFIRNVMTLKSCSPIVG 326
                       . :* :*:****    *  **:.  **:***..:    *       :  : :  *: : *
```

```
sp|P30316|DPOD_SCHPO   TQVYEFQNQAEMLSSWSKFVRDVDPDVLIGYNICNFDIPYLLDRAKSLRIHNFPLLGRIH 409
sp|P15436|DPOD_YEAST   SMIFSHATEEEMLSNWRNFIIKVDPDVIIGYNTTNFDIPYLLNRAKALKVNDFPYFGRLK 430
sp|P28339|DPOD1_BOVIN  AKVQSYEREEDLLQAWSTFIRIMDPDVITGYNIQNFDLPYLISRAQTLKVPGFPLLGRVI 424
sp|P28340|DPOD1_HUMAN  AKVQSYEKEEDLLQAWSTFIRIMDPDVITGYNIQNFDLPYLISRAQTLKVQTFPFLGRVA 425
sp|O54747|DPOD1_RAT    AKVQSYEREEDLLQAWATFILAMDPDVITGYNIQNFDLPYLISRAQTLKVDRFPFLGRVT 421
sp|P30315|DPOD1_PLAFK  SNVIWFNDEKTLLEAWNEFIIRIDPDFLTGYNIINFDLPYILNRGTALNLKKLKFLGRIK 419
sp|P54358|DPOD1_DROME  SQVLCHDKETQMLDKWSAFVREVDPDILTGYNINNFDFPYLLNRAAHLKVRNFEYLGRIK 409
sp|P52431|DPOD1_MOUSE  AKVQSYEREEDLLQAWADFILAMDPDVITGYNIQNFDLPYLISRAQALKVDRFPFLGRVT 423
sp|P90829|DPOD1_CAEEL  SNIIQCVNEKVLLEKWAEFVREVDPDIITGYNILNFDLPYILDRAKVLSLPQVSHLGRQK 395
sp|Q9LVN7|DPOD1_ARATH  VDVMSFETEREVLLAWRDLIRDVDPDIIIGYNICKFDLPYLIERAATLGIEEFPLLGRVK 419
sp|Q9LRE6|DPOD1_ORYSJ  VDVMSFDTERDVLLAWRDFIREVDPDIIIGYNICKFDLPYLIERAEVLKIVEFPILGRIR 429
sp|P46588|DPOD_CANAL   SQIFEHQREEDMLMHWKEFITKVDPDVIIGYNTANFDIPYVLNRAKALGLNDFPFFGRLK 359
sp|P97283|DPOD1_MESAU  AKVQSYEREEDLLQAWPNFILAMDPDVITGYNIQNFDLPYLISRAQTLKVDRFPFLGRVT 421
sp|O48901|DPOD1_SOYBN  VDVMPFETEREVLLAWRDFIREVDPDIIIGYNICKFDLPYLIERALNLKIAEFPILGRIR 386
                         :     :  :*  *  ::  :***.: ***  :**:**:.*.  * :  .  :**

sp|P30316|DPOD_SCHPO   NFFSVAKETTFSSKAYGTRESKTTSIPGRLQLDMLQVMQRDFKLRSYSLNAVCSQFLGEQ 469
sp|P15436|DPOD_YEAST   TVKQEIKESVFSSKAYGTRETKNVNIDGRLQLDLLQFIQREYKLRSYTLNAVSAHFLGEQ 490
sp|P28339|DPOD1_BOVIN  GLRSNIRESSFQSRQTGRRDSKVVSMVGRVQMDMLQVLLREYKLRSYTLNAVSFHFLGEQ 484
sp|P28340|DPOD1_HUMAN  GLCSNIRDSSFQSKQTGRRDTKVVSMVGRVQMDMLQVLLREYKLRSYTLNAVSFHFLGEQ 485
sp|O54747|DPOD1_RAT    GLRSNIRDSSFQSRQVGRRDSKVVSMVGRVQMDMLQVLLREYKLRSYTLNAVSFHFLGEQ 481
sp|P30315|DPOD1_PLAFK  NVASTVKDSSFSSKQFGTHETKEINIFGRIQFDVYDLIKRDYKLKSYTLNYVSFEFLKEQ 479
sp|P54358|DPOD1_DROME  NIRSVIKEQMLQSKQMGRRENQYVNFEGRVPFDLLFVLLRDYKLRSYTLNAVSYHFLQEQ 469
sp|P52431|DPOD1_MOUSE  GLRSNIRDSSFQSRQVGRRDSKVISMVGRVQMDMLQVLLREHKLRSYTLNAVSFHFLGEQ 483
sp|P90829|DPOD1_CAEEL  EKGSVVRDAAISSKQMGSRVNKSIDIHGRIIFDVLQVVLRDYKLRSYTLNSVSYQFLSEQ 455
sp|Q9LVN7|DPOD1_ARATH  NSRVRVRDSTFSSRQQGIRESKETTIEGRFQFDLIQAIHRDHKLSSYSLNSVSAHFLSEQ 479
sp|Q9LRE6|DPOD1_ORYSJ  NSRVRVRDTTFSSRQYGMRESKDVAVEGRVQFDLLQAMQRDYKLSSYSLNSVSAHFLGEQ 489
sp|P46588|DPOD_CANAL   RVKQEIKDAVFSSRAYGTRENKVVNIDGRMQLDLLQFIQREYKLRSYTLNSVSAHFLGEQ 419
sp|P97283|DPOD1_MESAU  GLRSNIRDSSFQSRQVGRRDSKVVSMVGRVQMDMLQVLLREHKLRSYTLNAVSFHFLGEQ 481
sp|O48901|DPOD1_SOYBN  NSRVRVKDTTFSSRQYGTRESKEVAVEGRVTFDLLQVMQRDYKLSSYSLNSVSSHFLSEQ 446
                             ::  :.*:  * : .:   . **. :*:   : *:.** **:** *. .** **

sp|P30316|DPOD_SCHPO   KEDVHYSIITDLQNGTADSRRRLAIYCLKDAYLPQRLMDKLMCFVNYTEMARVTGVPFNF 529
sp|P15436|DPOD_YEAST   KEDVHYSIISDLQNGDSETRRRLAVYCLKDAYLPLRLMEKLMALVNYTEMARVTGVPFSY 550
sp|P28339|DPOD1_BOVIN  KEDVQHSIITDLQNGNDQTRRRLAVYCLKDAFLPLRLLERLMVLVNAMEMARVTGVPLGY 544
sp|P28340|DPOD1_HUMAN  KEDVQHSIITDLQNGNDQTRRRLAVYCLKDAYLPLRLLERLMVLVNAVEMARVTGVPLSY 545
sp|O54747|DPOD1_RAT    KEDVQHSIITDLQNGNEQTRRRLAVYCLKDAFLPLRLLERLMVLVNNVEMARVTGVPLGY 541
sp|P30315|DPOD1_PLAFK  KEDVHYSIMNDLQNESPESRKRIATYCIKDGVLPLRLIDKLLFIYNYVEMARVTGTPFVY 539
sp|P54358|DPOD1_DROME  KEDVHHSIITDLQNGDEQTRRRLAMYCLKDAYLPLRLLEKLMAIVNYMEMARVTGVPLES 529
sp|P52431|DPOD1_MOUSE  KEDVQHSIITDLQNGNEQTRRRLAVYCLKDAFLPLRLLERLMVLVNNVEMARVTGVPLGY 543
sp|P90829|DPOD1_CAEEL  KEDVEHNIIPDLQRGDEQTRRRLAQYCLKDAYLPLRLLDKLMSIINYIEMARVTGVPMNF 515
sp|Q9LVN7|DPOD1_ARATH  KEDVHHSIITDLQNGNAETRRRLAVYCLKDAYLPQRLLDKLMFIYNYVEMARVTGVPISF 539
sp|Q9LRE6|DPOD1_ORYSJ  KEDVHHSIISDLQNGNSETRRRLAVYCLKDAYLPQRLLDKLMYIYNYVEMARVTGVPISF 549
sp|P46588|DPOD_CANAL   KEDVQHSIITDLQNGTKETRRRLAVYCLKDAFLPLRLLDKLMCLVNYTEMARVTGVPFSY 479
sp|P97283|DPOD1_MESAU  KEDVQHSIITDLQNGNEQTRRRLAVYCLRDAFLPLRLLERLMVLVNNVEMARVTGVPLGY 541
sp|O48901|DPOD1_SOYBN  KEDVHHSIISDLQNGNAETRRRLAVYCLKDAYLPQRLLDKLMFIYNYVEMARVTGVPISF 506
                       ****.:.*: ***.   ::*:*:* **::* . ** **:::*: : *  *******.*:
```


sp|P30316|DPOD_SCHPO LLARGQQIKVISQLFRKALQHDLVVPNIRVNGT-DEQYEGATVIEPIKGYYDTPIATLDF 588
sp|P15436|DPOD_YEAST LLARGQQIKVVSQLFRKCLEIDTVIPNMQSQAS-DDQYEGATVIEPIRGYYDVPIATLDF 609
sp|P28339|DPOD1_BOVIN LLSRGQQVKVVSQLLRQAMRQGLLMPVVKTEGG--EDYTGATVIEPLKGYYDVPIATLDF 602
sp|P28340|DPOD1_HUMAN LLSRGQQVKVVSQLLRQAMHEGLLMPVVKSEGG--EDYTGATVIEPLKGYYDVPIATLDF 603
sp|O54747|DPOD1_RAT LLSRGQQVKVVSQLLRQAMREGLLMPVVKTEGG--EDYTGATVIEPLKGYYDVPIATLDF 599
sp|P30315|DPOD1_PLAFK LLTRGQQIKVTSQLYRKCKELNYVIPSTYMKVNTNEKYEGATVLEPIKGYYIEPISTLDF 599
sp|P54358|DPOD1_DROME LLTRGQQIKVLSQLLRKAKTKGFIMPSYTSQGS-DEQYEGATVIEPKRGYYADPISTLDF 588
sp|P52431|DPOD1_MOUSE LLTRGQQVKVVSQLLRQAMRQGLLMPVVKTEGG--EDYTGATVIEPLKGYYDVPIATLDF 601
sp|P90829|DPOD1_CAEEL LLTKGQQIKILSMMLRRCKQNNFFLPVIEANSGDGEGYEGATVIDPIRGFYNEPIATLDF 575
sp|Q9LVN7|DPOD1_ARATH LLARGQSIKVLSQLLRKGKQKNLVLPNAKQSGSEQGTYEGATVLEARTGFYEKPIATLDF 599
sp|Q9LRE6|DPOD1_ORYSJ LLSRGQSIKVLSQLLRKAKQKNLVIPNIKGQASGQDTFEGATVLEARAGFYEKPIATLDF 609
sp|P46588|DPOD_CANAL LLSRGQQIKVISQLFRKCLQEDIVIPNLKSEGS-NEEYEGATVIEPERGYYDVPIATLDF 538
sp|P97283|DPOD1_MESAU LLSRGQQVKVVSQLLRQAMRQGLLMPVVKTEGG--EDYTGATVIEPLKGYYDVPIATLDF 599
sp|O48901|DPOD1_SOYBN LLSRGQSIKVLSQLLRRARQKNLVIPNAKQAGSEQGTFEGATVLEARAGFYEKPIATLDF 566
sp|P30316|DPOD_SCHPO SSLYPSIMQAHNLCYTTLLDSNTAELL----KLKQDV-DYSVTPNGDYFVKPHVRKGLLP 643
sp|P15436|DPOD_YEAST NSLYPSIMMAHNLCYTTLCNKATVERL----NLKIDE-DYVITPNGDYFVTTKRRRGILP 664
sp|P28339|DPOD1_BOVIN SSLYPSIMMAHNLCYTTLLRPGAAQKL----GLTED--QFIKTPTGDEFVKASVRKGLLP 656
sp|P28340|DPOD1_HUMAN SSLYPSIMMAHNLCYTTLLRPGTAQKL----GLTED--QFIRTPTGDEFVKTSVRKGLLP 657
sp|O54747|DPOD1_RAT SSLYPSIMMAHNLCYTTLLRPGAAQKL----GLKPD--EFIKTPTGDEFVKASVRKGLLP 653
sp|P30315|DPOD1_PLAFK ASLYPSIMIAHNLCYSTLIKSNH-EVS----DLQNDDITTIQGKNNLKFVKKNVKKGILP 654
sp|P54358|DPOD1_DROME ASLYPSIMMAHNLCYTTLVLGGTREKLRQQENLQDD--QVERTPANNYFVKSEVRRGLLP 646
sp|P52431|DPOD1_MOUSE SSLYPSIMMAHNLCYTTLLRPGAAQKL----GLKPD--EFIKTPTGDEFVKSSVRKGLLP 655
sp|P90829|DPOD1_CAEEL ASLYPSIMIAHNLCYTTLLKSP----Q----GVENE--DYIRTPSGQYFATKSKRRGLLP 625
sp|Q9LVN7|DPOD1_ARATH ASLYPSIMMAYNLCYCTLVTPEDVRKL----NLPPE--HVTKTPSGETFVKQTLQKGILP 653
sp|Q9LRE6|DPOD1_ORYSJ ASLYPSIMMAYNLCYCTLVPPEDARKL----NLPPE--SVNKTPSGETFVKPDVQKGILP 663
sp|P46588|DPOD_CANAL SSLYPSIMMAHNLCYTTLLNKNSIKAF----GLTED--DYTKTPNGDYFVHSNLRKGILP 592
sp|P97283|DPOD1_MESAU SSLYPSIMMAHNLCYTTLLRPGAAQKL----GLKPD--EFIKTPTGDEFVKSSVRKGLLP 653
sp|O48901|DPOD1_SOYBN ASLYPSIMMAYNLCYCTLVIPEDARKL----NIPPE--SVNRTPSGETFVKSNLQKGILP 620
sp|P30316|DPOD_SCHPO IILADLLNARKKAKADLKKETDPFKKAVLDGRQLALKVSANSVYGFTGATN-GRLPCLAI 702
sp|P15436|DPOD_YEAST IILDELISARKRAKKDLRDEKDPFKRDVLNGRQLALKISANSVYGFTGATV-GKLPCLAI 723
sp|P28339|DPOD1_BOVIN QILENLLSARKRAKAELAKETDPLRRQVLDGRQLALKVSANSVYGFTGAQV-GRLPCLEI 715
sp|P28340|DPOD1_HUMAN QILENLLSARKRAKAELAKETDPLRRQVLDGRQLALKVSANSVYGFTGAQV-GKLPCLEI 716
sp|O54747|DPOD1_RAT QILENLLSARKRAKAELAQETDPLRRQVLDGRQLALKVSPNSVYGFTGAQV-GKLPCLEI 712
sp|P30315|DPOD1_PLAFK LIVEELIEARKKVKLLIKNEKNNITKMVLNGRQLALKISANSVYGYTGASSGGQLPCLEV 714
sp|P54358|DPOD1_DROME EILESLLAARKRAKNDLKVETDPFKRKVLDGRQLALKISANSVYGFTGAQV-GKLPCLEI 705
sp|P52431|DPOD1_MOUSE QILENLLSARKRAKAELAQETDPLRRQVLDGRQLALKVSANSVYGFTGAQV-GKLPCLEI 714
sp|P90829|DPOD1_CAEEL EILEDILAARKRAKNDMKNEKDEFKRMVYNGRQLALKISANSVYGFTGATV-GKLPCLEI 684
sp|Q9LVN7|DPOD1_ARATH EILEELLTARKRAKADLKEAKDPLEKAVLDGRQLALKISANSVYGFTGATV-GQLPCLEI 712
sp|Q9LRE6|DPOD1_ORYSJ EILEELLAARKRAKADLKEAKDPFERAVLDGRQLALKISANSVYGFTGATV-GQLPCLEI 722
sp|P46588|DPOD_CANAL TILDELLTARKKAKADLKKETDPFKKDVLNGRQLALKISANSVYGFTGATV-GKLPCLAI 651
sp|P97283|DPOD1_MESAU QILENLLSARKRAKAELAQETDPLRRQVLDGRQLALKVSANSVYGFTGAEV-GKLPCLEI 712
sp|O48901|DPOD1_SOYBN EILEELLTARKRAKADLKEAKDPLEKAVLDGRQLALKISANSVYGFTGATI-GQLPCLEI 679

```
sp|P30316|DPOD_SCHPO   SSSVTSYGRQMIEKTKDVVEKRYR---IENGYSHDAVVIYGDTDSVMVKFGVKTLPEAMK 759
sp|P15436|DPOD_YEAST   SSSVTAYGRTMILKTKTAVQEKYC---IKNGYKHDAVVVYGDTDSVMVKFGTTDLKEAMD 780
sp|P28339|DPOD1_BOVIN  SQSVTGFGRQMIEKTKQLVETKYT---VENGYSTSAKVVYGDTDSVMCRFGVSSVAEAMA 772
sp|P28340|DPOD1_HUMAN  SQSVTGFGRQMIEKTKQLVESKYT---VENGYSTSAKVVYGDTDSVMCRFGVSSVAEAMA 773
sp|O54747|DPOD1_RAT    SQSVTGFGRQMIEKTKQLVETKYT---LENGYDANAKVVYGDTDSVMCRFGVSSVAEAMS 769
sp|P30315|DPOD1_PLAFK  AVSITTLGRSMIEKTKERVESFYC---KSNGYEHNSTVIYGDTDSVMVKFGTNNIEEAMT 771
sp|P54358|DPOD1_DROME  SGSVTAYGRTMIEMTKNEVESHYT---QANGYENNAVVIYGDTDSVMVNFGVKTLERSME 762
sp|P52431|DPOD1_MOUSE  SQSVTGFGRQMIEKTKQLVESKYT---VENGYDANAKVVYGDTDSVMCRFGVSSVAEAMS 771
sp|P90829|DPOD1_CAEEL  SQSVTAFGRKMIDMTKLEVERIYKKGALDGKCPADAKVIYGDTDSVMVKFGVETVAQAME 744
sp|Q9LVN7|DPOD1_ARATH  SSSVTSYGRQMIEQTKKLVEDKFT---TLGGYQYNAEVIYGDTDSVMVQFGVSDVEAAMT 769
sp|Q9LRE6|DPOD1_ORYSJ  SSSVTSYGRQMIEHTKKLVEDKFT---TLGGYEHNAEVIYGDTDSVMVQFGVSTVEDAMK 779
sp|P46588|DPOD_CANAL   SSSVTAFGREMIEKTKNEVQEYYS---KKNGHPYDAKVIYGDTDSVMVKFGYQDLETCMK 708
sp|P97283|DPOD1_MESAU  SQSVTGFGRQMIEKTKQLVESKYT---LENGYNANAKVVYGDTDSVMCRFGVSSVAEAMS 769
sp|O48901|DPOD1_SOYBN  SSSVTSYGRQMIEHTKKLVEDKFT---TLNGYEHNAEVIYGDTDSVMVQFGVSAVEEAMN 736
                       : *:*  ** **  **  *:  :      .    .: *:******** .**   :  .*

sp|P30316|DPOD_SCHPO   LGEEAANYVSDQFPNPIKLEFEKVYFPYLLISKKRYAGLFW-TRTDTYDKMDSKGIETVR 818
sp|P15436|DPOD_YEAST   LGTEAAKYVSTLFKHPINLEFEKAYFPYLLINKKRYAGLFW-TNPDKFDKLDQKGLASVR 839
sp|P28339|DPOD1_BOVIN  LGREAADWVSGHFPSPIRLEFEKVYFPYLLISKKRYAGLLFSSRPDAHDRMDCKGLEAVR 832
sp|P28340|DPOD1_HUMAN  LGREAADWVSGHFPSPIRLEFEKVYFPYLLISKKRYAGLLFSSRPDAHDRMDCKGLEAVR 833
sp|O54747|DPOD1_RAT    LGREAANWVSSHFPSPIRLEFEKVYFPYLLISKKRYAGLLFSSRSDAHDRMDCKGLEAVR 829
sp|P30315|DPOD1_PLAFK  LGKDAAERISKEFLSPIKLEFEKVYCPYLLLNKKRYAGLLY-TNPNKHDKMDCKGIETVR 830
sp|P54358|DPOD1_DROME  LGREAAELVSSKFVHPIKLEFEKVYYPYLLINKKRYAGLYF-TRPDTYDKMDCKGIETVR 821
sp|P52431|DPOD1_MOUSE  LGREAANWVSSHFPSPIRLEFEKVYFPYLLISKKRYAGLLFSSRSDAHDKMDCKGLEAVR 831
sp|P90829|DPOD1_CAEEL  IGLDAAKEVSKIFTPPIKLEFEKVYSPYLLINKKRYAGLYF-TKPDVHDKMDCKGLETVR 803
sp|Q9LVN7|DPOD1_ARATH  LGREAAEHISGTFIKPIKLEFEKVYFPYLLINKKRYAGLLW-TNPQQFDKMDTKGIETVR 828
sp|Q9LRE6|DPOD1_ORYSJ  LGREAADYISGTFIKPIKLEFEKIYFPYLLISKKRYAGLYW-TNPEKFDKMDTKGIETVR 838
sp|P46588|DPOD_CANAL   LGEEAANYVSTKFKNPIKLEFEKVYFPYLLINKKRYAGLYW-TRPEKFDKMDTKGIETVR 767
sp|P97283|DPOD1_MESAU  LGREAANWVSSHFPSPIRLEFEKVYFPYLLISKKRYAGLLFSSQPDTHDRMDCKGLEAVR 829
sp|O48901|DPOD1_SOYBN  LGREAAEHISGTFTKPIKLEFEKVYYPYLLISKKRYAGLFW-TKPDNFDKMDTKGIETVR 795
                       :* :**. :*  *  **.***** * ****:*******  : :..: .*::* **: :**

sp|P30316|DPOD_SCHPO   RDNCPLVSYVIDTALRKMLIDQDVEGAQLFTKKVISDLLQNKIDMSQLVITKALSK--TD 876
sp|P15436|DPOD_YEAST   RDSCSLVSIVMNKVLKKILIERNVDGALAFVRETINDILHNRVDISKLIISKTLA---PN 896
sp|P28339|DPOD1_BOVIN  RDNCPLVANLVTASLRRLLIDRDPSGAVAHAQDVISDLLCNRIDISQLVITKELTRAAAD 892
sp|P28340|DPOD1_HUMAN  RDNCPLVANLVTASLRRLLIDRDPEGAVAHAQDVISDLLCNRIDISQLVITKELTRAASD 893
sp|O54747|DPOD1_RAT    RDNCPLVANLVTSSLRRILVDRDPDGAVAHAKDVISDLLCNRIDISQLVITKELTRAAAD 889
sp|P30315|DPOD1_PLAFK  RDFCILIQQMMETVLNKLLIEKNLNSAIEYTKSKIKELLTNNIDMSLLVVTKSLGK--TD 888
sp|P54358|DPOD1_DROME  RDNSPLVANLMNSCLQKLLIERDPDGAVAYVKQVIADLLCNRIDISHLVITKELAK--TD 879
sp|P52431|DPOD1_MOUSE  RDNCPLVANLVTSSLRRILVDRDPDGAVAHAKDVISDLLCNRIDISQLVITKELTRAAAD 891
sp|P90829|DPOD1_CAEEL  RDNCPLVAKVLGVCLEKLLIERDQQSALDFAKRTISDLLCNKIDISLLIISKELTKSGDK 863
sp|Q9LVN7|DPOD1_ARATH  RDNCLLVKNLVTESLNKILIDRDVPGAAENVKKTISDLLMNRIDLSLLVITKGLTKTGDD 888
sp|Q9LRE6|DPOD1_ORYSJ  RDNCLLVKNLVTECLHKILVDRDVPGAVQYVKNTISDLLMNRVDLSLLVITKGLTKTGED 898
sp|P46588|DPOD_CANAL   RDNCRLVQNVITKVLEFILEERDVPKAQRFVKQTIADLLQNRIDLSQLVITKAYSK--HD 825
sp|P97283|DPOD1_MESAU  RDNCPLVANLVTSSLRRILVDRDPDGAVAHAKDVISDLLCNRIDISQLVITKELTRAAAD 889
sp|O48901|DPOD1_SOYBN  RDNCLLVKNLVNDCLHKILIDRDIPGAVQYVKNAISDLLMNRMDLSLLVITKGLTKTGDD 855
                       ** . *:  ::   *. :* :::   *   .:  * ::* *.:*:* *:::*       .
```

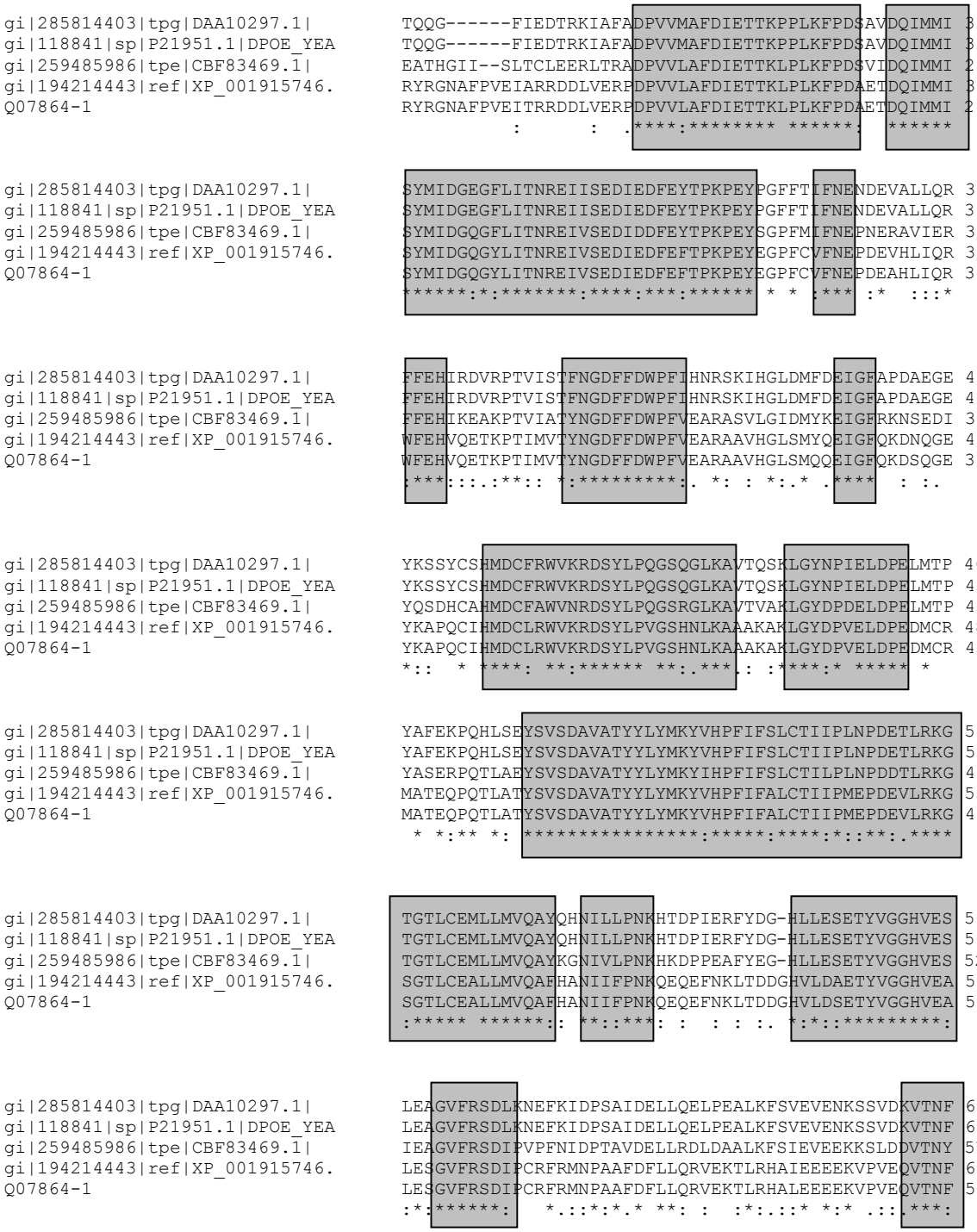
```
sp|P30316|DPOD_SCHPO   YAAKMAHVELAERMRKRDAGSAPAIGDRVAYVIIKGAQGDQFYMRSEDPIYVLENNIPID 936
sp|P15436|DPOD_YEAST   YTNPQPHAVLAERMKRRE-GVGPNVGDRVDYVIIGG--NDKLYNRAEDPLFVLENNIQVD 953
sp|P28339|DPOD1_BOVIN  YAGKQAHVELAERMRKRDPGSAPSLGDRVPYVIISAAKGVAAYMKSEDPLFVLEHSLPID 952
sp|P28340|DPOD1_HUMAN  YAGKQAHVELAERMRKRDPGSAPSLGDRVPYVIISAAKGVAAYMKSEDPLFVLEHSLPID 953
sp|O54747|DPOD1_RAT    YAGKQAHVELAERMRKRDPGSAPNLGDRVPYVIIGAAKGVAAYMKSEDPLFVLEHSLPID 949
sp|P30315|DPOD1_PLAFK  YETRLPHVELAKKLKQRDSATAPNVGDRVSYIIVKGVKGQAQYERAEDPLYVLDNNLAID 948
sp|P54358|DPOD1_DROME  YAAKQAHVELAAKMKKRDPGTAPKLGDRVPYVICAAAKNTPAYQKAEDPLYVLENSVPID 939
sp|P52431|DPOD1_MOUSE  YAGKQAHVELAERMRKRDPGSAPSLGDRVPYVIIGAAKGVAAYMKSEDPLFVLEHSLPID 951
sp|P90829|DPOD1_CAEEL  YQAKQAHVELAARMKKRDAGSAPRLGDRVPYVFVAAAKNVPAYERAEDPTFVLQNNIPLD 923
sp|Q9LVN7|DPOD1_ARATH  YEVKSAHGELAERMRKRDAATAPNVGDRVPYVIIKAAKGAKAYERSEDPIYVLQNNIPID 948
sp|Q9LRE6|DPOD1_ORYSJ  YAVKAAHVELAERMRKRDAATAPTVGDRVPYVIIKAAKGAKAYERSEDPIYVLDNNIPID 958
sp|P46588|DPOD_CANAL   YSAKQAHVELAERMRKRDPGSAPTLGDRVAYVIIKT-GGDKNYEKSEDPLYVLENSLPID 884
sp|P97283|DPOD1_MESAU  YAGKQAHVELAERMRKRDPGSAPSLGDRVPYVIIGAAKGVAAYMKSEDPLFVLEHSLPID 949
sp|O48901|DPOD1_SOYBN  YEVKAAHVELAERMRKRDAATAPNVGDRVPYVIIKAAKGAKAYERSEDPIYVLENNIPID 915
                       *    .*  ** ::::*: . .* :**** *::     .   * ::*** :**::.: :*
```

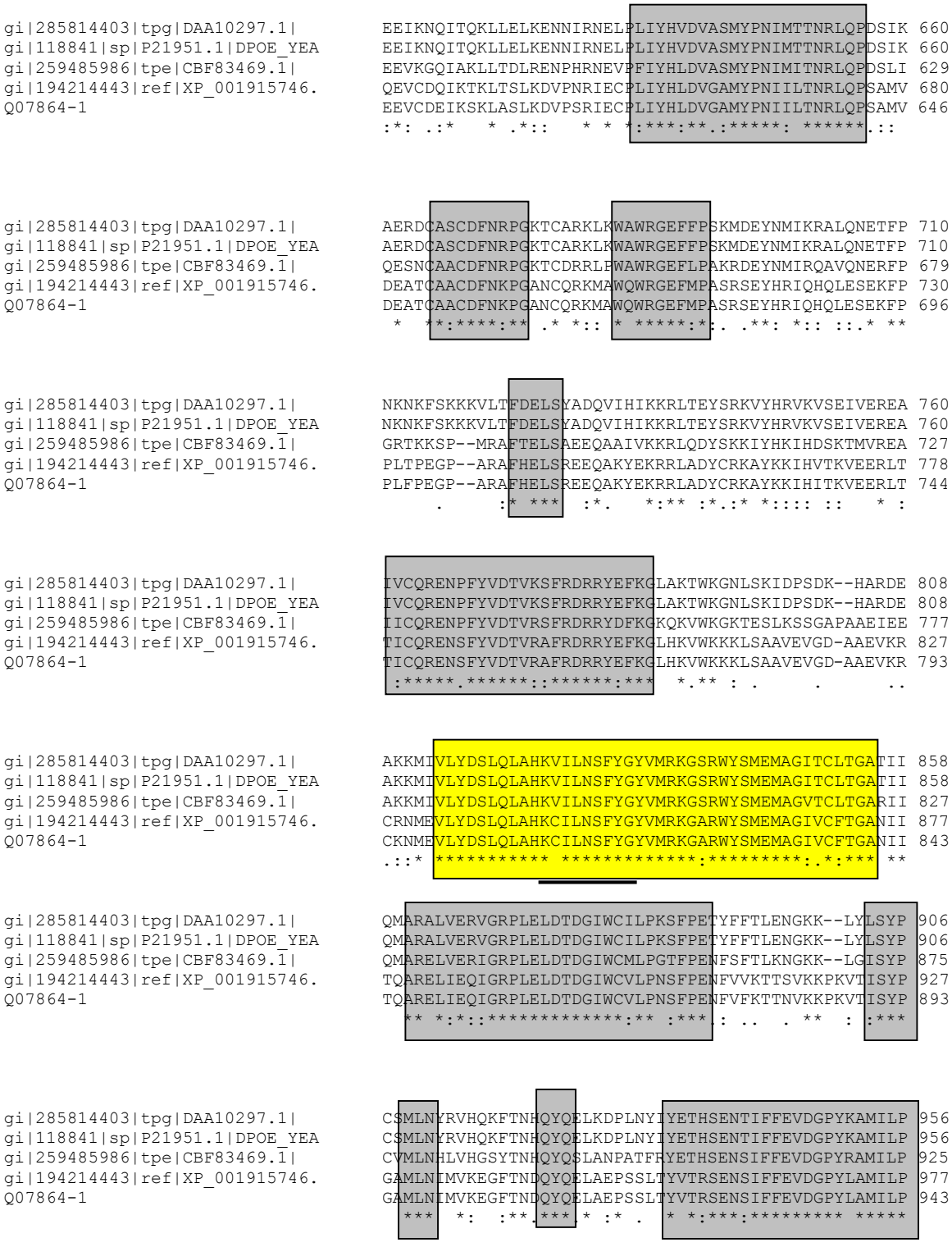
Fig. 2J. CLUSTAL format for T-COFFEE Version of delta polymerases (Eukaryotes)

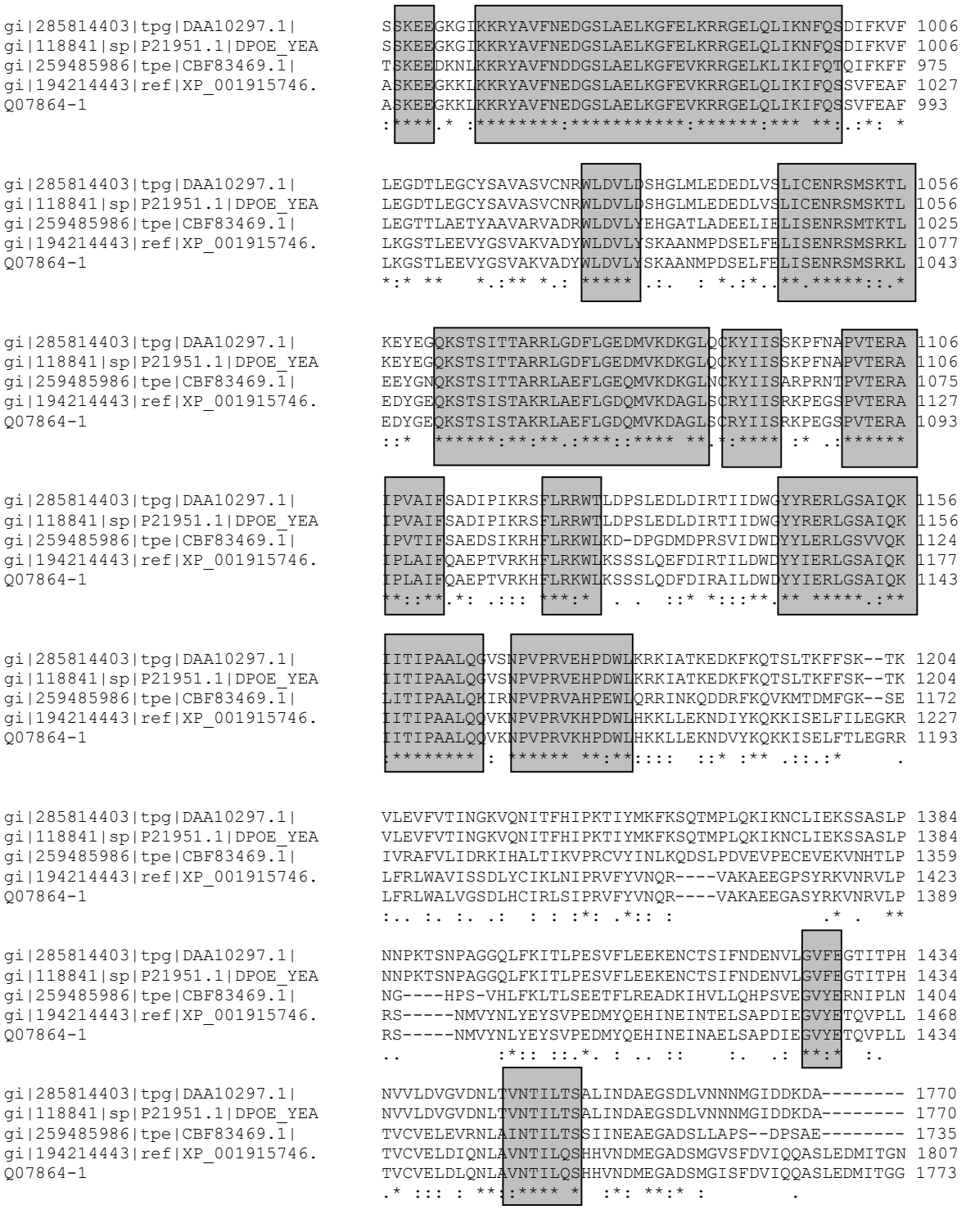
The possible catalytic region is underlined and shaded yellow.
Only regions showing highly conserved blocks are shown.

Fig. 2K Epsilon polymerases (Eukaryotes)

```
gi|285814403|tpg|DAA10297.1|       AQQLLNASKIDDIDSMMGFERYVPPQYNG-RFDAKDIDQIPGRVGWLTNM 82
gi|118841|sp|P21951.1|DPOE_YEA     AQQLLNASKIDDIDSMMGFERYVPPQYNG-RFDAKDIDQIPGRVGWLTNM 82
gi|259485986|tpe|CBF83469.1|       SLRSSEATSQDEKFEAIRLANSIDESLGFPRFEAGEK-----RVGWLINM 76
gi|194214443|ref|XP_001915746.     QLGRRDDGPSSSVSALKRLERSQWTDKMDLRFGFERLREPGEKTGWLINM 100
Q07864-1                           ----RDDGATSSVSALKRLERSQWTDKMDLRFGFERLKEPGEKTGWLINM 66
                                        :    ..      : .    .    **          :.*** **

gi|285814403|tpg|DAA10297.1|       HATLVSQETLSSGSNGGGNSNDGERVTTNQGISGVDFYFLDEEGGSFKST 132
gi|118841|sp|P21951.1|DPOE_YEA     HATLVSQETLSSGSNGGGNSNDGERVTTNQGISGVDFYFLDEEGGSFKST 132
gi|259485986|tpe|CBF83469.1|       HSTSIEDPNVPGGR------------------AGVDYYFLDDDGGSFKAT 108
gi|194214443|ref|XP_001915746.     HPTEVLDEDKRLVS-------------------AVDYYFIQDDGSRFKVA 131
Q07864-1                           HPTEILDEDKRLGS-------------------AVDYYFIQDDGSRFKVA 97
                                   *.* : :                          .**:**::::*. ** :

gi|285814403|tpg|DAA10297.1|       VVYDPYFFIACNDESRVNDVEELVKKYLESCLKSLQIIRKEDLTMDNHLL 182
gi|118841|sp|P21951.1|DPOE_YEA     VVYDPYFFIACNDESRVNDVEELVKKYLESCLKSLQIIRKEDLTMDNHLL 182
gi|259485986|tpe|CBF83469.1|       VEYDPYFLIAVKTG-HEAEVEEWCRRMFEGLIKKIKRVVKEDLKLPNHLL 157
gi|194214443|ref|XP_001915746.     LPYKPYFYIATRKG-CEREVSSFLSKKFQGKIAKVETVPKEDLDLPNHLV 180
Q07864-1                           LPYKPYFYIATRKG-CEREVSSFLSKKFQGKIAKVETVPKEDLDLPNHLV 146
                                   : *.*** ** .      :*..   : ::. : .:: : **** : ***:

gi|285814403|tpg|DAA10297.1|       ------NVAANGSEKVDAKHLIEDIREYDVPYHVRVSIDKDIRVGKWYKV 267
gi|118841|sp|P21951.1|DPOE_YEA     ------NVAANGSEKVDAKHLIEDIREYDVPYHVRVSIDKDIRVGKWYKV 267
gi|259485986|tpe|CBF83469.1|       ---------ARPNGTTNASDFIIDIREYDVPYHVRVAIDKDIRIGKWYTV 232
gi|194214443|ref|XP_001915746.     GGSVIIDEEETSKKMADQLDNIVDMREYDVPYHIRLSIDLKIHVAHWYNV 280
Q07864-1                           RGGVITDEEETSKKIADQLDNIVDMREYDVPYHIRLSIDLKIHVAHWYNV 246
                                             .  .:  . * *:********:*::** .*::.:**.*
```


gi|285814403|tpg|DAA10297.1| TQQG------FIEDTRKIAFADPVVMAFDIETTKPPLKFPDSAVDQIMMI 311
gi|118841|sp|P21951.1|DPOE_YEA TQQG------FIEDTRKIAFADPVVMAFDIETTKPPLKFPDSAVDQIMMI 311
gi|259485986|tpe|CBF83469.1| EATHGII--SLTCLEERLTRADPVVLAFDIETTKLPLKFPDSVIDQIMMI 280
gi|194214443|ref|XP_001915746. RYRGNAFPVEIARRDDLVERPDPVVLAFDIETTKLPLKFPDAETDQIMMI 330
Q07864-1 RYRGNAFPVEITRRDDLVERPDPVVLAFDIETTKLPLKFPDAETDQIMMI 296
gi|285814403|tpg|DAA10297.1| SYMIDGEGFLITNREIISEDIEDFEYTPKPEYPGFFTIFNENDEVALLQR 361
gi|118841|sp|P21951.1|DPOE_YEA SYMIDGEGFLITNREIISEDIEDFEYTPKPEYPGFFTIFNENDEVALLQR 361
gi|259485986|tpe|CBF83469.1| SYMIDGQGFLITNREIVSEDIDDFEYTPKPEYSGPFMIFNEPNERAVIER 330
gi|194214443|ref|XP_001915746. SYMIDGQGYLITNREIVSEDIEDFEFTPKPEYEGPFCVFNEPDEVHLIQR 380
Q07864-1 SYMIDGQGYLITNREIVSEDIEDFEFTPKPEYEGPFCVFNEPDEAHLIQR 346
gi|285814403|tpg|DAA10297.1| FFEHIRDVRPTVISTFNGDFFDWPFIHNRSKIHGLDMFDEIGFAPDAEGE 411
gi|118841|sp|P21951.1|DPOE_YEA FFEHIRDVRPTVISTFNGDFFDWPFIHNRSKIHGLDMFDEIGFAPDAEGE 411
gi|259485986|tpe|CBF83469.1| FFEHIKEAKPTVIATYNGDFFDWPFVEARASVLGIDMYKEIGFRKNSEDI 380
gi|194214443|ref|XP_001915746. WFEHVQETKPTIMVTYNGDFFDWPFVEARAAVHGLSMYQEIGFQKDNQGE 430
Q07864-1 WFEHVQETKPTIMVTYNGDFFDWPFVEARAAVHGLSMQQEIGFQKDSQGE 396
gi|285814403|tpg|DAA10297.1| YKSSYCSHMDCFRWVKRDSYLPQGSQGLKAVTQSKLGYNPIELDPELMTP 461
gi|118841|sp|P21951.1|DPOE_YEA YKSSYCSHMDCFRWVKRDSYLPQGSQGLKAVTQSKLGYNPIELDPELMTP 461
gi|259485986|tpe|CBF83469.1| YQSDHCAHMDCFAWVNRDSYLPQGSRGLKAVTVAKLGYDPDELDPELMTP 430
gi|194214443|ref|XP_001915746. YKAPQCIHMDCLRWVKRDSYLPVGSHNLKAAAKAKLGYDPVELDPEDMCR 480
Q07864-1 YKAPQCIHMDCLRWVKRDSYLPVGSHNLKAAAKAKLGYDPVELDPEDMCR 446
gi|285814403|tpg|DAA10297.1| YAFEKPQHLSEYSVSDAVATYYLYMKYVHPFIFSLCTIIPLNPDETLRKG 511
gi|118841|sp|P21951.1|DPOE_YEA YAFEKPQHLSEYSVSDAVATYYLYMKYVHPFIFSLCTIIPLNPDETLRKG 511
gi|259485986|tpe|CBF83469.1| YASERPQTLAEYSVSDAVATYYLYMKYIHPFIFSLCTILPLNPDDTLRKG 480
gi|194214443|ref|XP_001915746. MATEQPQTLATYSVSDAVATYYLYMKYVHPFIFALCTIIPMEPDEVLRKG 530
Q07864-1 MATEQPQTLATYSVSDAVATYYLYMKYVHPFIFALCTIIPMEPDEVLRKG 496
gi|285814403|tpg|DAA10297.1| TGTLCEMLLMVQAYQHNILLPNKHTDPIERFYDG-HLLESETYVGGHVES 560
gi|118841|sp|P21951.1|DPOE_YEA TGTLCEMLLMVQAYQHNILLPNKHTDPIERFYDG-HLLESETYVGGHVES 560
gi|259485986|tpe|CBF83469.1| TGTLCEMLLMVQAYKGNIVLPNKHKDPPEAFYEG-HLLESETYVGGHVES 529
gi|194214443|ref|XP_001915746. SGTLCEALLMVQAFHANIIFPNKQEQEFNKLTDDGHVLDAETYVGGHVEA 580
Q07864-1 SGTLCEALLMVQAFHANIIFPNKQEQEFNKLTDDGHVLDSETYVGGHVEA 546
gi|285814403|tpg|DAA10297.1| LEAGVFRSDLKNEFKIDPSAIDELLQELPEALKFSVEVENKSSVDKVTNF 610
gi|118841|sp|P21951.1|DPOE_YEA LEAGVFRSDLKNEFKIDPSAIDELLQELPEALKFSVEVENKSSVDKVTNF 610
gi|259485986|tpe|CBF83469.1| IEAGVFRSDIPVPFNIDPTAVDELLRDLDAALKFSIEVEEKKSLDDVTNY 579
gi|194214443|ref|XP_001915746. LESGVFRSDIPCRFRMNPAAFDFLLQRVEKTLRHAIEEEEKVPVEQVTNF 630
Q07864-1 LESGVFRSDIPCRFRMNPAAFDFLLQRVEKTLRHALEEEEKVPVEQVTNF 596


gi|285814403|tpg|DAA10297.1| EEIKNQITQKLLELKENNIRNELPLIYHVDVASMYPNIMTTNRLQPDSIK 660
gi|118841|sp|P21951.1|DPOE_YEA EEIKNQITQKLLELKENNIRNELPLIYHVDVASMYPNIMTTNRLQPDSIK 660
gi|259485986|tpe|CBF83469.1| EEVKGQIAKLLTDLRENPHRNEVPFIYHLDVASMYPNIMITNRLQPDSLI 629
gi|194214443|ref|XP_001915746. QEVCDQIKTKLTSLKDVPNRIECPLIYHLDVGAMYPNIILTNRLQPSAMV 680
Q07864-1 EEVCDEIKSKLASLKDVPSRIECPLIYHLDVGAMYPNIILTNRLQPSAMV 646
gi|285814403|tpg|DAA10297.1| AERDCASCDFNRPGKTCARKLKWAWRGEFFPSKMDEYNMIKRALQNETFP 710
gi|118841|sp|P21951.1|DPOE_YEA AERDCASCDFNRPGKTCARKLKWAWRGEFFPSKMDEYNMIKRALQNETFP 710
gi|259485986|tpe|CBF83469.1| QESNCAACDFNRPGKTCDRRLPWAWRGEFLPAKRDEYNMIRQAVQNERFP 679
gi|194214443|ref|XP_001915746. DEATCAACDFNKPGANCQRKMAWQWRGEFMPASRSEYHRIQHQLESEKFP 730
Q07864-1 DEATCAACDFNKPGANCQRKMAWQWRGEFMPASRSEYHRIQHQLESEKFP 696
gi|285814403|tpg|DAA10297.1| NKNKFSKKKVLTFDELSYADQVIHIKKRLTEYSRKVYHRVKVSEIVEREA 760
gi|118841|sp|P21951.1|DPOE_YEA NKNKFSKKKVLTFDELSYADQVIHIKKRLTEYSRKVYHRVKVSEIVEREA 760
gi|259485986|tpe|CBF83469.1| GRTKKSP--MRAFTELSAEEQAAIVKKRLQDYSKKIYHKIHDSKTMVREA 727
gi|194214443|ref|XP_001915746. PLTPEGP--ARAFHELSREEQAKYEKRRLADYCRKAYKKIHVTKVEERLT 778
Q07864-1 PLFPEGP--ARAFHELSREEQAKYEKRRLADYCRKAYKKIHITKVEERLT 744
gi|285814403|tpg|DAA10297.1| IVCQRENPFYVDTVKSFRDRRYEFKGLAKTWKGNLSKIDPSDK--HARDE 808
gi|118841|sp|P21951.1|DPOE_YEA IVCQRENPFYVDTVKSFRDRRYEFKGLAKTWKGNLSKIDPSDK--HARDE 808
gi|259485986|tpe|CBF83469.1| IICQRENPFYVDTVRSFRDRRYDFKGKQKVWKGKTESLKSSGAPAAEIEE 777
gi|194214443|ref|XP_001915746. TICQRENSFYVDTVRAFRDRRYEFKGLHKVWKKKLSAAVEVGD-AAEVKR 827
Q07864-1 TICQRENSFYVDTVRAFRDRRYEFKGLHKVWKKKLSAAVEVGD-AAEVKR 793
gi|285814403|tpg|DAA10297.1| AKKMIVLYDSLQLAHKVILNSFYGYVMRKGSRWYSMEMAGITCLTGATII 858
gi|118841|sp|P21951.1|DPOE_YEA AKKMIVLYDSLQLAHKVILNSFYGYVMRKGSRWYSMEMAGITCLTGATII 858
gi|259485986|tpe|CBF83469.1| AKKMIVLYDSLQLAHKVILNSFYGYVMRKGSRWYSMEMAGVTCLTGARII 827
gi|194214443|ref|XP_001915746. CRNMEVLYDSLQLAHKCILNSFYGYVMRKGARWYSMEMAGIVCFTGANII 877
Q07864-1 CKNMEVLYDSLQLAHKCILNSFYGYVMRKGARWYSMEMAGIVCFTGANII 843
gi|285814403|tpg|DAA10297.1| QMARALVERVGRPLELDTDGIWCILPKSFPETYFFTLENGKK--LYLSYP 906
gi|118841|sp|P21951.1|DPOE_YEA QMARALVERVGRPLELDTDGIWCILPKSFPETYFFTLENGKK--LYLSYP 906
gi|259485986|tpe|CBF83469.1| QMARELVERIGRPLELDTDGIWCMLPGTFPENFSFTLKNGKK--LGISYP 875
gi|194214443|ref|XP_001915746. TQARELIEQIGRPLELDTDGIWCVLPNSFPENFVVKTTSVKKPKVTISYP 927
Q07864-1 TQARELIEQIGRPLELDTDGIWCVLPNSFPENFVFKTTNVKKPKVTISYP 893
gi|285814403|tpg|DAA10297.1| CSMLNYRVHQKFTNHQYQELKDPLNYIYETHSENTIFFEVDGPYKAMILP 956
gi|118841|sp|P21951.1|DPOE_YEA CSMLNYRVHQKFTNHQYQELKDPLNYIYETHSENTIFFEVDGPYKAMILP 956
gi|259485986|tpe|CBF83469.1| CVMLNHLVHGSYTNHQYQSLANPATFRYETHSENSIFFEVDGPYRAMILP 925
gi|194214443|ref|XP_001915746. GAMLNIMVKEGFTNDQYQELAEPSSLTYVTRSENSIFFEVDGPYLAMILP 977
Q07864-1 GAMLNIMVKEGFTNDQYQELAEPSSLTYVTRSENSIFFEVDGPYLAMILP 943


gi|285814403|tpg|DAA10297.1|
gi|118841|sp|P21951.1|DPOE_YEA
gi|259485986|tpe|CBF83469.1|
gi|194214443|ref|XP_001915746.
Q07864-1

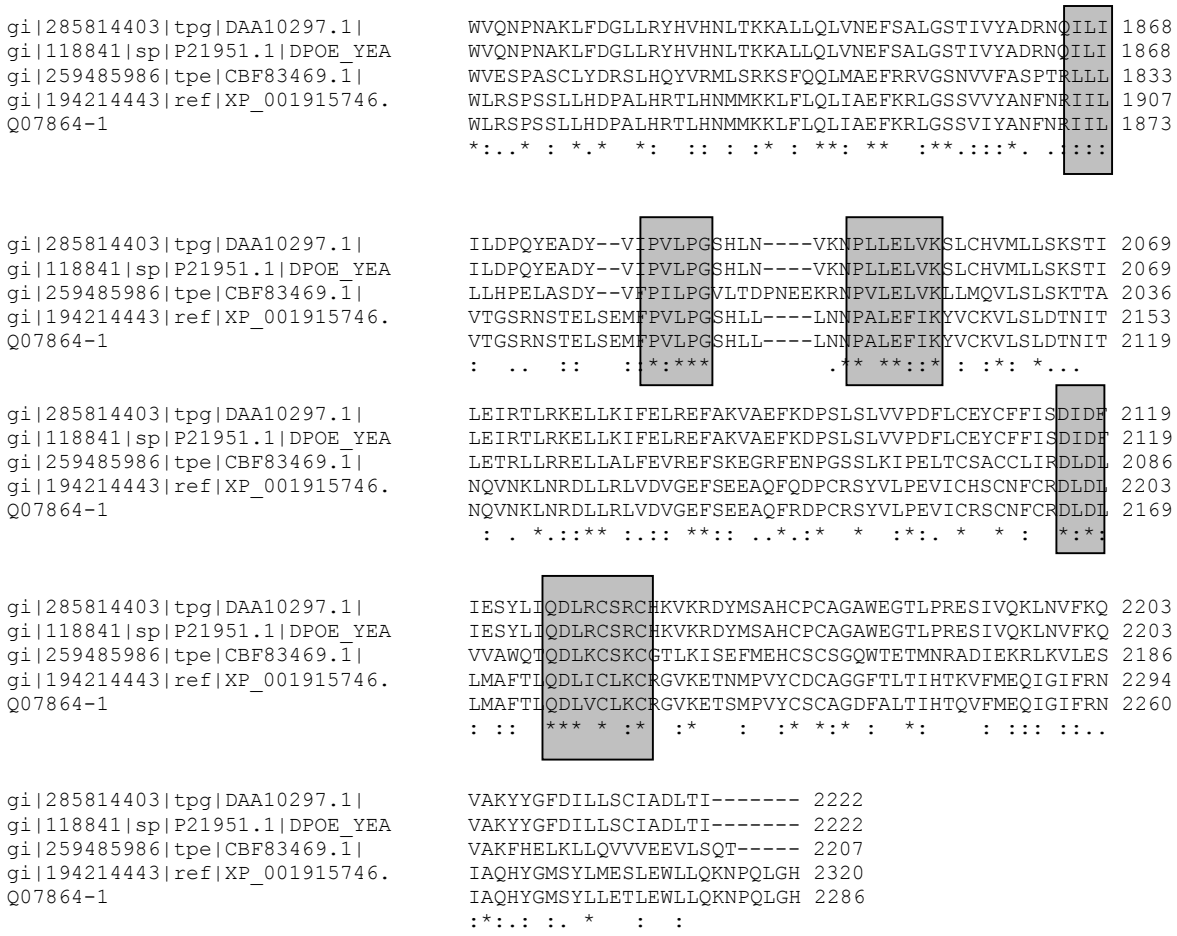


Fig. 2K. CLUSTAL 2.0.12 multiple sequence alignment of epsilon polymerases (catalytic subunit)-Eukaryotes

The possible catalytic region is underlined and shaded yellow.
Only regions showing highly conserved blocks are shown.

3.3 Conserved Amino Acid(s) in DNA Polymerases

Multiple sequence alignments of known prokaryotic and eukaryotic polymerases have clearly indicated two conserved sets of amino acids, viz., one consisting of the proposed catalytic site amino acid K and the other is an YG pair. In some cases the K is replaced by the equivalent amino acid R and in the YG pair, the G is replaced by A/I Table 1). It is interesting to note that a distance conservation is also observed in all these polymerases, i.e., the YG pair is positioned around 6-9 amino acids downstream from the catalytic amino acid K (Table 1).

Interestingly in all eukaryotic polymerases the second amino acid from the catalytic K is mostly a branched chain amino acid such as V/I/L. Similarly pol IV and pol V also use a branched chain amino acid. The steric gate amino acid Y is strictly followed by G in all cases except eukaryotic beta polymerases. In polymerase II, instead of a K, an R is conserved.

Unlike the pol I the pol II cannot use nicked duplex templates of DNA; it usually fills short gaps in DNA which act as a block to DNA polymerase and thus helps reinitiate DNA synthesis in UV induced cells. In some of the prokaryotic polymerases like pol IV and pol V, the steric gate amino acid Y is not found at the expected distance from the possible catalytic amino acid K, which probably explains the error-prone nature of the polymerases (Table 1).

Table 1. Conserved regions around the catalytic and steric-gate amino acids (Numbering from the catalytic amino acid K/R).

Prokaryotic polymerases		Eukaryotic polymerases	
pol I	A/V /K[1]A/T/SV/I/MN/SY/FGLI/AY[9]G	Alpha	L/I/K[1]LTANSMY[8]GCLG
pol II	K/R/M/R[1]H/LL/VXPXF/Y[7]G	Beta	K/S/K[1]V/I/LD/FS/PE/KY[6]I/LA/GT
pol III	NK/R[1]S/AHS/A/A/Y[7]G/A/S	Gamma	AK[1]I/VFNYG/ARIY[9]GAG
pol IV	L/M/AK[1]L/MA/GSXXXKP[9]XG[11]	Delta	LK[1]I/VSANSVY[8]GFTGA
pol V	A/S/K[1]L/VAN/M/QXXXXX	Epsilon	HK[1]V/CILNSFY[8]GYVMRKG
Viral	A/N/K[1]/RXF/IXXGXLY[9]GA/S		

Less common amino acids around the catalytic amino acid K, are shown in lower font sizes.
X represents any amino acid that is not conserved.

3.4 Absence of YG Pair in Error Prone Polymerases

Another interesting observation is that the error prone *E coli* polymerases pol IV and V did not show any YG pair at the expected positions from the catalytic K (Table 1). These polymerases were reported in 1999 and are grouped under error prone DNA polymerases [3]. (DNA polymerase IV is induced in stationary phase cells and is involved in adaptive mutagenesis whereas the DNA polymerase V replicates past gaps in DNA. It is made of two different subunits (*umuC* and *umuD')* and is involved in SOS DNA repair pathway, when *E. coli* is exposed to high levels of radiation or mutagens and a major damage is done to the DNA).

Thus, polymerase IV (*dinB*) and Pol V (*umuD'C*) belong to the newly discovered class of Y DNA polymerases [3]. Polymerases that belong to Y family are often referred to as specialized or error-prone DNA polymerases to distinguish them from the previously described DNA polymerases (pol I, pol II, and pol III) that are essentially involved in DNA replication or error-free DNA repair. These Y-family DNA polymerases are characterized by their ability to replicate DNA, through chemically damaged template bases, or to elongate mismatched primer termini. These properties stem from their capacity to accommodate and use distorted primer templates within their active site and from the lack of an associated exonuclease activity. It is Interesting to note that the explicit absence of the YG steric gate pair in these polymerases explains their error prone nature.

In the eukaryotic polymerase β, a nuclear DNA repair polymerase that plays a key role in base-excision repair, the G in YG pair is replaced with a branched chain amino acid, like I or L. The implication of such a replacement is not known.

3.5 Distance Conservation between Catalytic K and YG Pair

It is interesting to note that the catalytic amino acid K and the gate keeper pair YG are completely conserved in different polymerases from diverse group of organisms. The mismatched regions in some of the polymerases were aligned as suggested by Palanivelu [4]. Table 2 further corroborates the above findings. In this analysis, only the amino acids around the active site regions of different DNA polymerases from different sources ranging from virus to plant and animals were selected and analyzed by T-COFFEE advanced version. It is interesting to note that irrespective of the type of polymerases and their origin, all showed a completely conserved K at the catalytic site and YG pair at the steric gate position. A distance conservation is also observed in all these polymerases, (i.e.), the YG pair is 8/9 amino acids downstream of the catalytic K.

Table 2. The catalytic amino acid (K) and gate keeper pair (YG) in different polymerases from diverse sources

T4 DNA pol	546 ATLANTNQLNR**K**[1]**ILINSL**[8]**YG**ALGNIH
Human HSV 1	800 AVLLDKQQAAI**K**[1]**VVCNSV**[8]**YG**FTGVQH
E. coli DNA pol I	748 TVTSEQRRSA**K**[1]**AINFGLI**[9]**YG**MSAFGLAR
E. coli DNA pol II	482 RQGNKPLSQAL**K**[1]**IIMNAF**[8]**YG**VLGTTA
E. coli DNA pol III (alpha subunit)	663 YPDVQWQHESL**K**[1]**PVLEPT**[8]**YG**IILYQE
P. furiosus DNA pol	477 KILLDYRQKAI**K**[1]**LLANSF**[8]**YG**YYGYAK
Yeast Alpha DNA pol	933 RVQCDIRQQAL**K**[1]**LTANSM**[8]**YG**CLGYVN
Human Alpha DNA pol	939 ILQYDIRQKAL**K**[1]**LTANSM**[8]**YG**CLGFSY
Human Gamma DNA pol	917 TTVGISREHA**K**[1]**IFNYGRI**[9]**YG**AGQPFAER
Human Delta DNA pol (Catalytic subunit)	683 RQVLDGRQLAL**K**[1]**VSANSV**[8]**YG**FTGAQV
Human Epsilon DNA pol	798 EVLYDSLQLAH**K**[1]**CILNSF**[8]**YG**YVMRKGAR
A. thaliana Delta DNA pol (Catalytic subunit)	679 KAVLDGRQLAL**K**[1]**ISANSV**[8]**YG**FTGATV

N.B: Some of the above polymerases did not align in T COFFEE advanced version. So the conserved regions were selected and aligned as suggested by Palanivelu [4].

3.6 Other Conserved Motif(s) in Prokaryotic and Eukaryotic Polymerases

Apart from these highly conserved regions, not much conservation is observed among different polymerase families. However, there are conserved motifs among the members of the same family. Prokaryotic polymerase pol V and viral polymerases did not show much conservation among themselves. Eukaryotic polymerases δ and ε showed the maximum number of conserved motifs. ε polymerases showed long stretches of conserved motifs among them (Table 3).

A notable one is the DXD motif which is found in all eukaryotic polymerases but only in one of the prokaryotic polymerases viz., in the replicative polymerase, pol III, This triad is also implicated in catalysis. Delta and Epsilon polymerases have two such triads.

Table 3. Other highly conserved regions found among different prokaryotic and eukaryotic polymerases

Table 3A. Prokaryotic polymerases

pol I	EADDV/I IXS/T SP/TKQL TGRLSS PNLQNIP DYS QIEL/MRL/I/VL/MAH IQGS/T QVHDEL VPL
pol II	V/IV/IG V/I/LED PGNHD PQP GHI/VH V/INS
pol III	GPGRGS LL/IFERFLNPER PDX**DV/ID**F LIKXFDFLGL RPGP IV/L/IYQEQV/I
pol IV	I/VIHV/ID FFA RGV/IV/I LSL/IDEA/G GV/IG VRLL/VGV/L/I
pol V	D/EV/IWGV/IGR/K
Viral	YNV/I V/ILN ESI

Table 3B. Eukaryotic polymerases

Alpha	LFGK LLHXF PDK GGLVLD/EP ILLLDFNSLYPSIIQEY/FNICFTTV/I SRFY VIYG**DTD**SIMI KKKYAA KGLD RRDW/Y/F
Beta	CGSY/F **DXD**VL KIL RRXDIR GVLY
Gamma	KLPH GTITRRAVEXTWLTASN RLGSELK F/LVGADVDSQ/EELWI GT/SD TKG SRV/IN VDY LHL HDEV/IRYL AFFSX**DID** LRKE TPSN
Delta	LSFDIEC FPE DPVIQI **DPDI/V** GYN FDLPYL SYS/TLN KEDVDLQN YCLKD EMARVTGXP RGQ GATVL/I TLDF SLYPSIMMAXNLCYXTL GL/ILP ARKR GXLPCL VI/VYG**DTD**SVM PIK/RLEFEKVYXPYLLIXKKRYAGL KGI/L LVITK V/LGDRVXY EDP
Epsilon	GWLTNM VDYYFIQD/ED/EG PYFXIA KEDL NHLL/V DXREYDVPYHIRLSID DPVVL/MAFDIETTKXPLKFPDXXDQIMMI SYMIDGXGY/FLITNREIVSEDIEDFEF/YTPKPEY I/VFNEF/WFEH TYNGDFFDWPFV/I EIGF HMDCXXWVXRDSYLPXGSXXXLKA KLGYXPXELDPE YSVSDAVATYYLYMKYVHPFIFALCTIIPMEPDEXLRKGS/TGTLCEXLLMVQAY/F NII/L/VXPNK HVLDSETYVGGHVE GVFRSDI/L VTNF/Y IYHLDVG/AXMYPNIXXNRLQP CAXCDFNKPG WXWRGEFXP FXELS V/ICQRENXFYVDTVR/KXFRDRRYD/EFKG ARXLIEXIGRPLEL**DTD**GIWCXLPXSFPE L/ISYP MLN QYQYXTR/HSENS/TIFFEVDGPYXAMILPXSKEE KKRYAVFNEDGSLAELKGFEVKRRGELXLIKXFQS/T WLDVL LIXENRSMSR/KXL QKSTSIS/TTAK/RRLXEFLGDXMVKDXGLXCR/KYIIS PVTERA IPLXIF FLRKW YYXERLGSXIQKIITIPAALQXXXPVPRVXKHPDWL VNTILXS PVLPG NPXLEXI/VK **DL/ID** QDLXCK/RC

NB: X represents any amino acid smaller font sizes represent less frequently used amino acids.

3.7 Three Critical Pairs in Different Types of Polymerases and their Possible Functions

These analyses have revealed three critical pairs of amino acids in all these polymerases belonging to different types and origins. Table 4 shows these three critical pairs from different representative polymerases.

Table 4. Critical pairs of amino acids found in different types of polymerases

Enzyme	Lys//Arg	Tyr/Phe-Gly	Asp//Asp
T_7 pol	Arg518/Lys522	Tyr530/Gly531	Asp475/Asp654
E. coli pol I	Arg754/Lys758	Tyr766/Gly 767	Asp705/Asp882
E. coli pol II	Leu*523/Arg527	Phe533/Gly534	Asp452/ Asp545
E. coli pol III	^----/Lys758	Tyr764/^Ala765	Asp405/Asp733
Human α	Lys947/Lys950	Tyr957/Gly958	Asp860/Asp1004
Human β	Leu*163/Lys168	Tyr 173/Leu*174	Asp192/Asp256
Human γ	Arg943/Lys947	Tyr955/Gly956	Asp890/Asp935
Human δ	Arg689/Lys694	Tyr701/Gly702	Asp602/Asp757
Yeast ε	Leu*819/Lys824	Tyr831/Gly832	Asp669/ Asp2118

Based on multiple sequence analysis.

** Instead of Arg, a Leu is found at the corresponding position in the repair polymerases viz., pol II and pol β.*

^ No Arg or leu is found near vicinity of the probable catalytic K. A good number of prokaryotic replicative polymerases (pol III) had an Ala adjacent to the Tyr (Fig.2C)

In almost all the pol IV polymerases, only a G (PXG) is seen at the 11th position from the catalytic K; no regular gate keeper Y is found which possibly explains the error prone nature of these polymerases. The completely conserved P known to act as a helix breaker may introduce the necessary sharp kinks into polypeptide backbone (Fig. 2D).

The ε polymerases also maintain a Leu near the catalytic K, as it is also involved in DNA repair. Pol ε's main function is to extend the leading strand during replication while Pol δ is involved in the lagging strand synthesis. The most striking difference between the two DNA polymerases is that processive DNA synthesis by DNA polymerase delta is dependent on proliferating cell nuclear antigen (PCNA), a replication factor, while DNA polymerase epsilon is inherently processive.

- The pair 1 involves in polymerization, (the K functions as proton abstractor);
- The pair 2 acts as "*steric gate"* and allows only dNTPs at polymerization site and
- The pair 3 involves as "*charge shielder"* of dNTPs through a Mg ion (Fig. 3) and orients the α-phosphates of dNTPs for polymerization. One Mg^{2+} ion binds specifically to the beta and gamma phosphates of the dNTP as a bidentate.

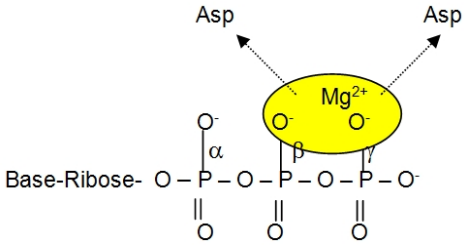


Fig. 3. dNTP-Mg ion complex where β- and γ-phosphates are shielded by the Mg ion

3.8 Functional domains of *E. coli* DNA polymerase I

As DNA polymerase I of *E. coli* is the most well studied enzyme, it is used as a model enzyme to elucidate the mechanism of action in this communication. DNA polymerase I of *E. coli* is made up of 928 amino acid residues and is a multifunctional enzyme with three different activities, viz., the polymerizing activity, the 3'→5' exonuclease activity (proof-reading function) and the 5'→3' exonuclease activity (DNA repair function).Thus, the enzyme molecule is made up of three distinct domains to perform the three different activities. These three different domains have been dissected and studied further [5].

Hans Klenow subjected the enzyme to a limited proteolysis with subtilisin, which yielded two fragments. Upon purification of the fragments by gel filtration chromatography, he found that the larger fragment, 68 kDa peptide, contained the polymerizing and proof-reading activities and the smaller fragment, 35 kDa peptide, contained the 5'→3' exonuclease activity. Further analysis of the enzyme has shown that the amino acids from 1-324 perform the 5'→3' exonuclease activity (DNA repair function); 324-517 perform the 3'→5' exonuclease activity (proof-reading function) and 517-928 perform the polymerizing function. The distance between the last two active domains is ~ 30 Å (Fig. 4).

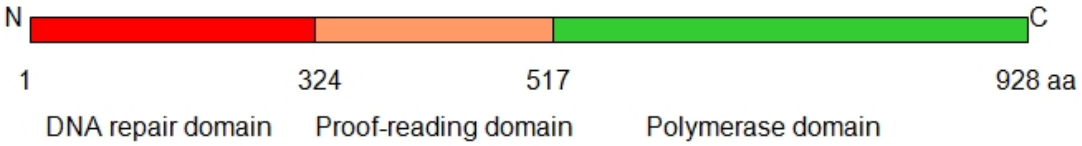


Fig. 4. Dissection of the 3 functional domains of *E. coli* DNA polymerase I

The proof-reading site is shown to bind two metal ions and one mole of dNMP [6]. The polymerizing site consists of a deep cleft that accommodates the double stranded DNA and follows a spiral path along the DNA.

3.9 Analysis of DNA polymerase I active site

The polymerase active site was probed by a variety of techniques:

DNase foot-printing assay using DNase I and methidium-propyl EDTA-Fe^{2+} indicated that the enzyme binds to the primer terminus and covers 8 base pairs.

Photo affinity labeling of the enzyme with dNTP analogue, 8-azido-dATP, and sequencing of the labeled peptide, identified Tyr^{766} at the active site of the enzyme. Thus, the foot printing and photo cross-linking experiment has suggested the Tyr^{766} in the active site [7].

However, Basu and Modak [8] who have probed the polymerase active site with pyridoxal phosphate, found a Lys^{758} at the active site. (Pyridoxal phosphate binds competitively to the dNTP site through Schiff's base formation and covalently links the amino acid responsible in polymerization reaction). These results suggest that the polymerase active site is in the bigger domain remote from the 3'→5' exonuclease activity and totally not connected to the dNMP site. Similar observations were made with an adenovirus DNA polymerase, e.g., the pyridoxal phosphate modification of an adenovirus DNA polymerase resulted in loss of DNA polymerase activity, whereas the 3'→5' exonuclease activity was unaffected. Inhibition of adenovirus DNA polymerase by pyridoxal phosphate was time-dependent and displayed saturation kinetics [9]. Zaldivar et al. [10] have shown that rat liver RNA polymerases I and II and yeast RNA polymerase I were also inactivated by pyridoxal phosphate and hence suggested a possible involvement of a Lysine residue in the catalytic site of RNA polymerases too.

From the above experiments, two different amino acids were implicated at the active site of the DNA polymerase I, viz., Tyr^{766} and Lys^{758}. Interestingly, both the Lys and Tyr are completely conserved in some of the prokaryotic DNA polymerases analyzed by Palanivelu [11]. The phi 29 viral DNA polymerase shares several regions of amino acid similarity with other alpha-like DNA polymerases. Among them, the conserved region characterized by the amino acid motif "Kx3NSxYG" has been proposed to form part of the polymerization active site of alpha-like DNA polymerases [12].

The following observations also support Lys^{758} as a probable amino acid involved in the polymerization reaction rather than Tyr^{766}:

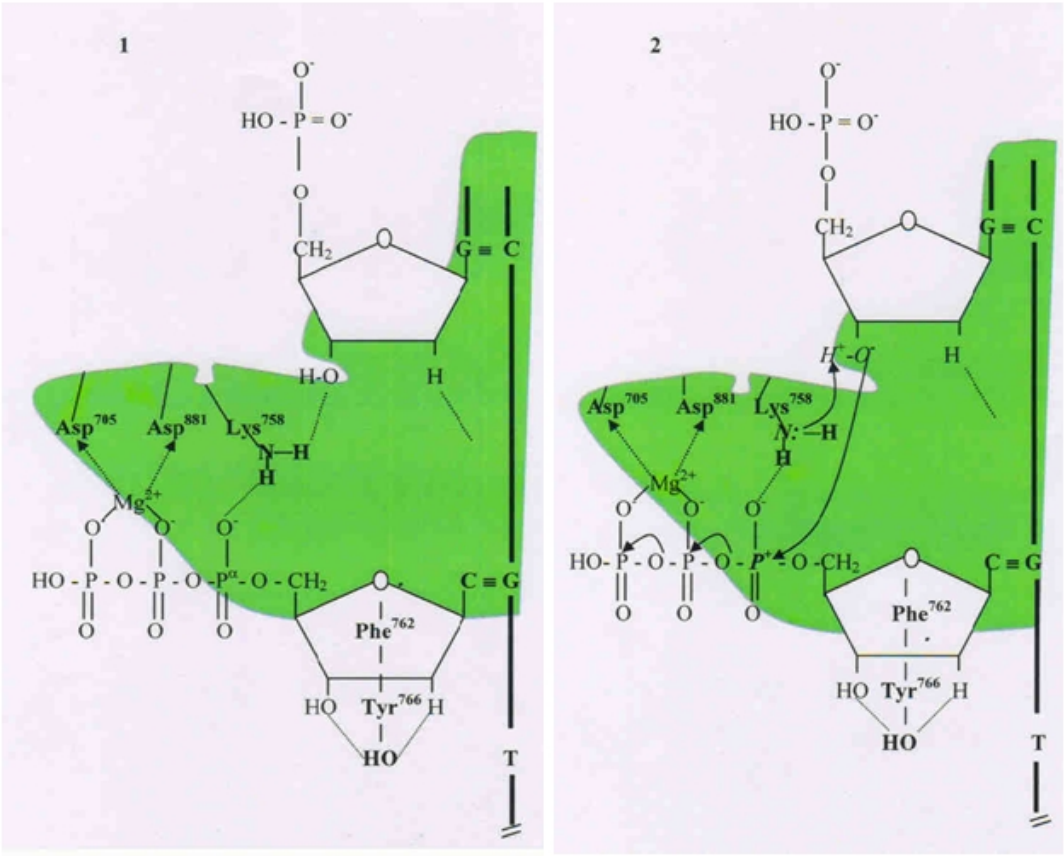
Lys is the active site amino acid in NAD and ATP dependent ligases, and also GTP dependent mRNA capping enzymes, which are all involved in making a phosphodiester bond as in polymerases [13]. Like DNA polymerases the DNA ligases are also inhibited by pyridoxal 5'-phosphate indicating the presence of a Lys at the catalytic domain of the enzyme [14]. Both the types of ligases (ATP dependent and NAD dependent DNA ligases) from various organisms showed a highly conserved motif KY/I/VDGXR with the reactive K residue, followed by a Tyr or a hydrophobic amino acid [13]. Interestingly, not only in DNA ligases, but also in RNA ligases the catalytic Lys is conserved [14].

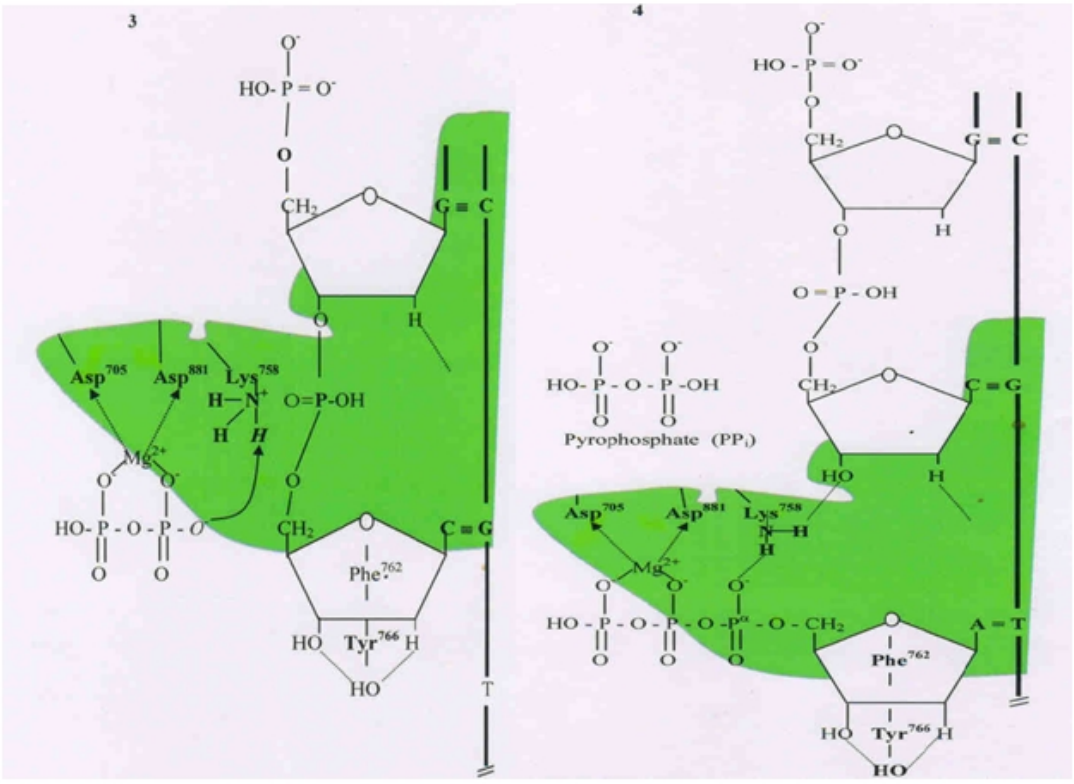
In contrast, Tyr is found in the active site of topoisomerases, which actually involve in breaking of phosphodiester bonds. Interestingly, Tyr is conserved in all topoisomerases. In fact, the active site Tyr sometimes makes a covalent bond with the 5' end of the DNA. It is interesting to note that these topoisomerases also have highly conserved Arg/Lys, which participates in the subsequent phosphodiester bond formation [15]. Kausik *et al.* [16] and Singh and Modak [17] have shown by site directed mutagenesis (SDM) experiments that replacing Y^{766} and F^{771} by Ala, significantly affected formation of Enzyme-DNA binary complex; but most importantly the catalytic activity could not be restored in a K758A mutant. Tyr^{766} is suggested to be in close proximity to the 3'-OH of the primer and highly conserved in all polymerases; it could still play an important role in the polymerization function. Further

analysis by SDM by Doublie and Ellenberger [18] and Astatke et al. [19] have shown that the critical Tyr may possibly be involved in template recognition and dNTP selection in DNA polymerases. It is also known that a highly conserved Tyr residue in reverse transcriptase controls substrate selection. In Klenow fragment, the homologous residue, Tyr^{766}, occupies a structurally equivalent position at the C-terminus of a long helix in the fingers subdomain. It is interesting to note that the highly conserved Tyr^{955} residue is critical for nucleotide recognition among Family A DNA polymerases, i.e., γ polymerases from eukaryotes. Furthermore, Tyr^{955} is a highly conserved residue among a wide variety of DNA polymerases (Table 2). As a Family-A DNA polymerase, the γ polymerases are related to *E. coli* DNA polymerase I and bacteriophage T7 DNA polymerase, and amino acid sequence alignments reveal that Tyr^{955} in γ polymerases is equivalent to Tyr^{766} in *E. coli* pol I and Tyr^{530} in T7 DNA polymerase. Further proof of Tyr^{766} involvement in nucleotide selection was obtained from site directed mutagenesis; an $Y^{766} \rightarrow F$ substitution in the Klenow polymerase did not show an appreciable increase in nucleotide misinsertion; however, substitution with Ala or Ser generated an error-prone DNA polymerase attributable to decreased stringency for selection of dNTPs [20]. Interestingly the YG doublet is highly conserved and found to be a common pair in different types of DNA polymerases (Table 2).

Further proof is provided by crystallographic analysis of T7 DNA polymerase. The T7 DNA replication complex at 2.2 Å resolution have shown that the highly conserved Lys^{522} ($\equiv Lys^{758}$ in *E. coli* DNA pol I) actually makes contact with the α-phosphate of dNTP [21].
Earlier structural, mutagenesis, and labeling studies have suggested that the incoming dNTP molecule contacts a region on one side of the polymerase cleft, primarily involving residues within the so-called “fingers” subdomain [22].

Since the mechanism of action for polymerization reactions, proposed in this article is based on a proton abstraction at the catalytic site, Lys is placed as the catalytic amino acid in this communication. Thus, active site Tyr holds the complementary base possibly inserted by the finger domain onto the catalytic site and the catalytic Lys adds the dNTP to the 3’-OH of the primer terminus. The reaction essentially occurs through proton abstraction followed by a nucleophilic attack at the growing primer terminus (Figs. 5.1-5.4).


1
2
Asp705
Asp881
Lys758
Mg2+
Phe762
Tyr766
G ≡ C
C ≡ G
T



Figs. 5.1 - 5.4. Proposed mechanism for polymerization reaction of *E. coli* DNA polymerase I

5.1. Watson-Crick base pairing of the incoming nucleotide with the template and nucleotide discrimination by steric gate amino acid tyrosine. The tyrosine-OH possibly discriminates between the 2'-H of the incoming dNTPs and 2'-OH of rNTPs. Gly/Ala residue, adjacent to the Tyr provides the flexibility necessary for the active site to change conformation to accommodate different dNTPs (not shown in the figure).

5.2. Electronic transition at the active site for proton abstraction by K followed by an electrophilic and nucleophilic attack.

5.3. Proton abstraction by the active site amino acid Lys and simultaneous formation of 3'→5' phosphodiester bond.

5.4. Transfer of the proton from Lys and formation of inorganic pyrophosphate and the next complementary nucleotide in position to get polymerized.

3.10 Analysis of the 3'→5' Exonuclease Active Site (Proof-Reading function) of the DNA polymerase I

The exonuclease's proof reading site was analyzed by Derbyshire et al. [23, 24] by genetic, crystallographic and SDM methods. The exonuclease active site (EAS) essentially consists of a dNMP site and a metal-binding site. Therefore, dNMPs can inhibit the exonuclease reaction by product inhibition. The metal binding site consists of two subsites, viz., Subsite-A and Subsite-B and thus, EAS can bind two divalent metal ions. The subsite-A is coordinated by three amino acids, viz., Asp^{355}- Glu^{357}- Asp^{501} and dNMP-phosphate provides the fourth ligand. Usually a Zn atom is associated to the subsite-A.

The second metal binding site, sub-site B is mainly coordinated by Asp^{424} and possibly to a Mg atom. The subsite-B is located between dNMP-phosphate and the carboxylate of Asp^{424}. The site A is very close to the 3' O- of the susceptible bond to be cleaved and the site B is very close to A. X-ray crystallographic data show that the distance between the two metal atoms is ~ 4.0 to 4.5 Å in *E. coli* (pol I) and T7 polymerases (Fig. 6).

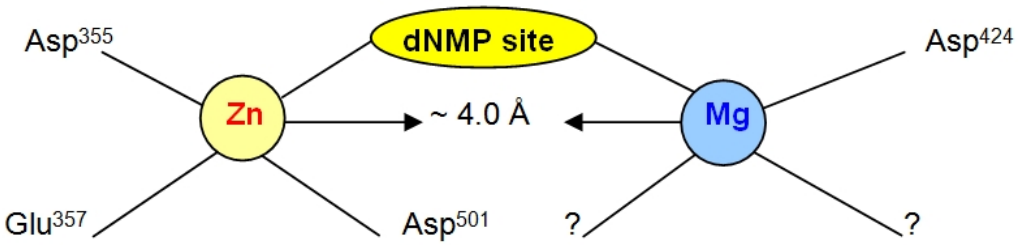


Fig. 6. Schematic diagram showing the subsites A and B of proof-reading activity of *E. coli* DNA polymerase I

The active site amino acids, which constitute the EAS, were further analyzed by site directed mutagenesis [5].

a) In a double mutant with $Asp^{355}\rightarrow$Ala and $Glu^{357}\rightarrow$Ala, both the dNMP binding site and the metal binding site A were completely abolished. This mutant protein had lost the exonuclease activity, *but exhibited the polymerase activity*. This suggested that the dNMP site is coordinated to both the metal binding sites.

b) In the second SDM experiment, the Asp^{424} was replaced by Ala ($Asp^{424}\rightarrow$Ala). In this mutant enzyme, the metal binding site B was abolished and exhibited no exonuclease activity. However, in this mutant protein also the polymerase activity was found to be preserved.

These two experiments suggest that the metal ion plays a direct role in proof-reading function and does not participate in polymerizing function. The SDM studies have further shown that both the metal binding sites are *functionally connected* and in the absence of one, the other cannot function. The Zn-site possibly involves in catalysis and the Mg-site linked to dNMP-phosphate could bind the dNMP site and also may provide the necessary "strain and distortion" for cleaving the susceptible bond of the wrongly added nucleotide.

3.11 Mechanism of Action of Proof-Reading Function

Polymerases occasionally make any mistake, as incorporation of even one wrong nucleotide may be detrimental to the growth and survival of organisms. In the event of a wrong nucleotide is incorporated, they cannot probably proceed further until the wrong nucleotide is removed with the help of the proof-reading function associated with them. It is just like if one tooth is bent in the zip fastener, the zip fastener cannot move further, until it is straightened. It is likely a similar situation operates in polymerases too. In such situations, the polymerase excise the wrong nucleotide by the proof-reading action, incorporates the correct nucleotide and then proceeds further.

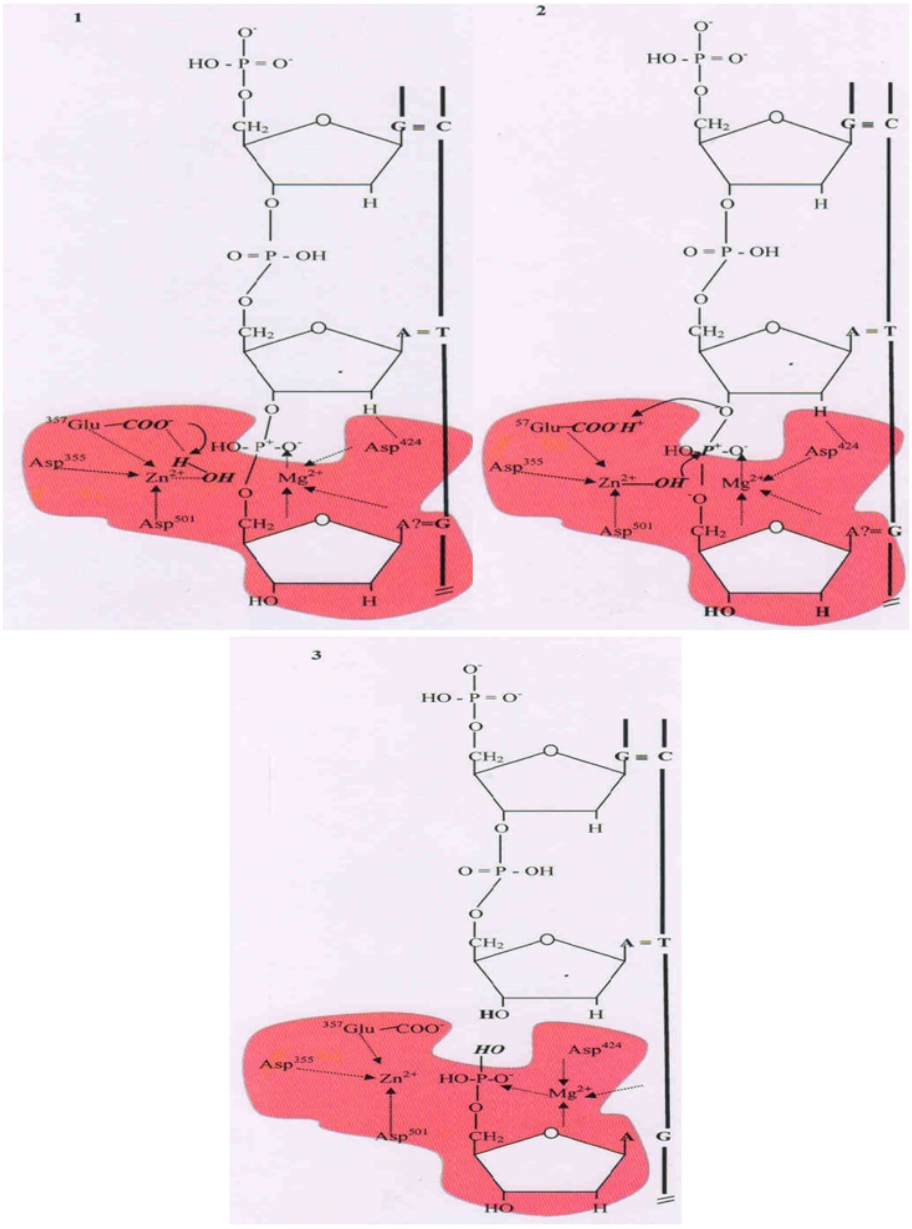
The proof reading function is essentially explained based on Zn based catalysis. As discussed elsewhere, the exonuclease active site of *E. coli* polymerase I consists of two metal binding sites, sub-site A and B. The sub-site A is coordinated by Asp^{355}, Glu^{357}, Asp^{501} and likely to a Zn atom. There is a large number of Zn containing enzymes like alkaline phosphatases, carbonic anhydrases, carboxy peptidases, thermolysin, etc. In all these enzymes, Zn plays a critical role in hydrolysis. For all catalytic Zn sites except coenzyme dependent alcohol dehydrogenase, the first two ligands are separated by a short spacer of 1 to 3 amino acids which are separated by a third ligand by a long spacer of about 20 to 120 amino acids [25,26]. Therefore, analysis of the mechanism of action of the Zn containing enzymes will throw more light on the proof-reading function of DNA polymerases too.

Generally, in all Zn containing enzymes, a water molecule is invariably coordinated to the Zn atom as the fourth ligand, the other three ligands could be contributed by carboxylic acids (Glu, Asp), His or Cys.

During catalysis, the Zn acts as a Lewis acid and displaces the proton from the water molecule and it is held transiently by the Glu^{357}, an active site amino acid at the metal binding site (the active site amino acid in other Zn containing enzymes are Glu^{270} in carboxypeptidase A; Glu^{143} in thermolysin; and Glu^{117} in carbonic anhydrase) and thus the Zn becomes Zn-hydroxide (Fig. 7).

Fig. 7. Formation Zn-hydroxide during Zn mediated enzyme catalysis

It is well known that metal hydroxides are very reactive species in chemical reactions and the Zn-hydroxide readily attacks the susceptible electrophilic center on the susceptible bond on the substrate molecule resulting in cleavage of the bond and simultaneous transfer of the proton from the active site carboxylic acid. The mechanism of 3'→5' exonuclease action, i.e., the proof-reading function, is also explained based on Zn-hydroxide formation and proton transfer. Derbyshire *et al.* [23] have also indicated that the nucleophilic attack of the terminal phosphodiester bond is initiated by a hydroxide ion coordinated to one of the enzyme-bound metal ions. The proposed mechanism of proof reading action is essentially based on the observations made by Beese and Steitz [6], and is illustrated in Figs. 8.1-8.3. The mismatched 3' end is a very poor substrate for further polymerization. Therefore, the 3'→5' exonuclease → polymerase activities switch between excision and incorporation modes without dissociation of the enzyme-substrate complex [27].



Figs. 8.1- 8.3. Proposed mechanism for poof-reading function of *E. coli* DNA polymerase I

8.1 A wrong nucleotide is placed during polymerization (Error rate is usually one in a million).
8.2 Transient proton transfer to the active site amino acid Glu^{357} and Zn-hydroxide free radical formation at the proof reading active site.
8.3 Nucleophilic attack by Zn-hydroxide free radical at the susceptible phosphodiester bond resulting in removal of the wrongly added nucleotide with concomitant transfer of the proton from Glu^{357} to the 3' growing end. The second metal ion possibly involves in stabilizing the transient pentavalent phosphorous group during the reaction.

4. CONCLUSION

Multiple sequence analyses have shown that a basic amino acid K/R and an YG pair are highly conserved in almost all DNA polymerases except in error prone polymerases where the YG pair is not found at the expected distance from the catalytic K/R. Site directed mutagenesis, biochemical and X-ray crystallographic analyses of DNA polymerase I from *E coli* have also suggested their involvement in catalysis and substrate binding. Based on these results, a mechanism of action is proposed for the polymerization reactions as well as for the proof reading function of DNA polymerase I from *E coli* as a model enzyme. Similar mechanism may be followed by other polymerase as the highly conserved K/R is present in all of them.

ACKNOWLEDGEMENTS

The author wishes to thank Prof. S. Krishnaswamy, Coordinator, Centre of Excellence in Bioinformatics and Prof. R. Usha, Department of Plant Biotechnology, School of Biotechnology, Madurai Kamaraj University, Madurai, for their useful suggestions on the manuscript.

REFERENCES

1. Kornberg A. DNA Replication. J Biol Chem. 1988;263:1-4.
2. Steitz TA. DNA Polymerases: Structural Diversity and Common Mechanisms. J. biol chem. 274:1999;17395-17398**.**
3. Ohmori H, Friedberg EC, Fuchs RP, Goodman MF, Hanaoka F, *et al*. The Y-family of DNA polymerases. 2001; Mol Cell, 8:7–8.
4. Palanivelu P. Multiple sequence analysis of polygalacturonases and invertases and phase shift in conserved motifs. Indian J Biotechnol. 2007;6:24-30.
5. Joyce CM, Steitz TA. DNA polymerase I: from crystal structure to function via genetics. Trends Biochem Sci. 1987;12:288-292.
6. Beese LS, Steitz TA. Structural basis for the 3'– 5' exonuclease activity of *Escherichia coli* DNA polymerase I: a two metal ion mechanism. EMBO J. 1991;10:25-33.
7. Rush J, Konisberg WH. Photoaffinity labeling of the Klenow fragment with 8-azido-dATP. J Biol Chem. 1990;265:4821-4827.
8. Basu A, Modak MJ. Affinity labeling of *E. coli* DNA polymerase I by pyridoxal 5'-phosphate. Biochemistry. 1987;26:1704-1709.
9. Monaghan N, Hay RT. Pyridoxal 5'-phosphate inhibition of adenovirus DNA polymerase. J Biol Chem. 1996;271:24242-24248.
10. Zaldivar MJ, Bull P, Venegas A, Valenzuela P. Inactivation of rat liver RNA polymerases I and II and yeast RNA polymerase I by pyrodixal 5'-phosphate. Evidence for the participation of lysyl residues at the active site. Biochemistry. 1975;14:4907-4911.
11. Palanivelu P. Enzymes, Ribozymes and DNAzymes. Twentyfirst Century Publications, Palkalai Nagar, Madurai, India; 2007.
12. Blasco MA, Lázaro JM, Blanco L , Salas M. Phi 29 DNA polymerase active site. The conserved amino acid motif "Kx3NSxYG" is involved in template-primer binding and dNTP selection. Proc Natl Acad Sci, (USA). 1997;94:1619–1622.

13. Sriskanda V, Schwer B, Ho CK, Shuman S. Mutational analysis of *Escherichia coli* DNA ligase identities amino acids required for nick-ligation *in vitro* and for *in vivo* complementation of their growth of yeast cells for CDC9 and LIG4. Nucleic acids Res. 1999;27:3953-3963.
14. Tomkinson AE, Totty NF, Ginsburg M, Lindahl T. Location of the active site for enzyme-adenylate formation in DNA ligases. Pro Natl Acad Sci (USA). 1991;88:400-404.
15. Redinbo MR, Stewart L, Kuhn P, Champoux JJ, Hol WGJ. Crystal Structures of Human Topoisomerase I in Covalent and Noncovalent Complexes with DNA. Science. 1998;279:1504-1513.
16. Kaushik N, Pandey VN, Modak MJ. Significance of the O-helix residues of *Escherichia coli* DNA polymerase I in DNA synthesis; dynamics of the dNTP binding pocket. Biochemistry. 1996;35:7256-66.
17. Singh K, Modak MJ. A unified DNA- and dNTP binding mode for DNA polymerases. Trends Biochem Sci. 1998;23:277-280.
18. Doublie S, Ellenberger T. The mechanism of action of T7 DNA polymerase. Curr Opin Struct Biol. 1998;8:704-712.
19. Astatke M, Grindley ND, Joyce CM. How *E. coli* DNA polymerase I (Klenow fragment) distinguishes between deoxy- and dideoxynucleotides. J Mol Biol. 1998;278:147-165.
20. Ponamarev MV, Longley MJ, Nguyen D, Kunkel TA, Copeland WC. Active site mutation in DNA polymerase γ associated with progressive external ophthalmoplegia causes error-prone DNA synthesis. J Biol Chem. 2002;277:15225-15228.
21. Doublie S, Tabor S, Long AM, Richardson, CC, Ellengerger T. Crystal structure of a bacteriophage T7 DNA replication complex at 2.2 Å resolution. Nature. 1998;391:251-258.
22. Joyce CM. Choosing the right sugar: How polymerases select a nucleotide substrate. Proc Natl Acad Sci (USA). 1997;94:1619-1622.
23. Derbyshire V, Freemont PS, Sanderson MR, Beese L, Friedman JM, Joyce CM , Steitz TA. Genetic and crystallographic studies of the 3,5'-exonucleolyitc site of DNA polymerase I. Science. 1988;240:199-201.
24. Derbyshire V, Grindley ND, Joyce CM. 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. 1991;10:17-24.
25. Vallee BL, Auld DS. Active-site zinc ligands and activated H_2O of zinc enzymes. Proc Natl Acad Sci (USA). 1990;87:220-224.
26. Vallee BL, Auld DS. Active zinc binding sites of zinc metalloenzymes 1: Matrix Suppl. 1992;5-19.
27. Mizrahi V, Benkovic P, Benkovic SJ. Mechanism of DNA polymerase I: exonuclease/ polymerase activity switch and DNA sequence dependence of pyrophosphorolysis and misincorporation reactions. Proc Natl Acad Sci, (USA). 1986;83:5769-5773.

Peer-review history:
The peer review history for this paper can be accessed here:
http://www.sciencedomain.org/review-history.php?iid=231&id=3&aid=1432