



## BRITISH JOURNAL OF PHARMACEUTICAL AND MEDICAL RESEARCH

Cross Ref DOI : <https://doi.org/10.24942/bjpmr.2022.988> Volume 07, Issue 03 May- June 2022ISSN:2456-9836  
IF 5.885 & ICV: 60.37

## Research Article

## Analyses of Priming Reactions and Proofreading Functions of Primases during Initiation of Replication of Prokaryotic and Eukaryotic Genomes

Peramachi Palanivelu\*

Senior Professor & Head (Retd.), Department of Molecular Microbiology, School of Biotechnology,  
Madurai Kamaraj University, Madurai – 625 021, India

## ARTICLE INFO

Article History:Received on 09<sup>th</sup> May, 2022Peer Reviewed on 25<sup>th</sup> May, 2022Revised on 16<sup>th</sup> June, 2022Published on 30<sup>th</sup> June, 2022

## ABSTRACT

In prokaryotes and eukaryotes, the crucial step of replication initiation is performed by RNA primases along with an additional DNA primase, viz. the DNA pol  $\alpha$ , in eukaryotes. The active sites of catalytic and proofreading activities of primases involved in the initiation of replication of prokaryotic and eukaryotic genomes are analysed. The template-binding pairs, catalytic amino acid(s), metal-binding sites are identified. The template-binding pairs are found to be the typical –YG– in the primases of eubacterial and eukaryotic (animal) sources as found in most of the DNA and RNA polymerases, but an -IG- pair is identified in the yeast and higher fungal primases. The proposed catalytic amino acid is found to be invariably a basic amino acid K/R from both eubacterial and eukaryotic primases

as in other DNA and RNA polymerases. The additional primase, DNA pol  $\alpha$ , which makes DNA primers exclusively for eukaryotic replication initiation also showed the typical template-binding pairs, –YG-, and the catalytic amino acid K/R. In addition to that, a two-in-one site –YGDTDS- and a –SLYPS- motif which are common among B family polymerases are found in the DNA pol  $\alpha$ . Primases from both eubacteria and eukaryotes contain typical Mg<sup>2+</sup>-binding triads, viz. -DxD- and additional Zn<sup>2+</sup>-binding sites. Moreover, many specialized and conserved sequence motifs are also identified both in the RNA primases and DNA pol  $\alpha$ . The PR function is found to be muted in DNA pol  $\alpha$  even though most of the conserved active sites amino acids are identified. BLASTp analysis showed the eubacterial and human primases are significantly different in their identities, suggesting that inhibitors of bacterial DNA primases could act as novel antibiotics.

Br J Phar Med Res Copyright©2022 **Peramachi Palanivelu** et al. This is an Open Access article distributed under the terms of the Creative Commons Attribution 4.0 International License (<http://creativecommons.org/licenses/by/4.0/>), allowing third parties to copy and redistribute the material in any medium or format and to remix, transform, and build upon the material for any purpose, even commercially, provided the original work is properly cited and states its license.

**Corresponding Author:** Peramachi Palanivelu, Senior Professor & Head (Retd.), Department of Molecular Microbiology, School of Biotechnology, Madurai Kamaraj University, Madurai – 625 021, India

## Introduction

Genome duplication, which preserves and maintains the blueprint of life, is an indispensable process in all living organisms. Generally, two types of replicases, viz. DNA-dependent DNA polymerases (DdDps) and RNA-dependent RNA polymerases (RdRps) are involved in replication of genomes, depending upon whether the genome is of DNA or RNA, respectively. The RNA genomes of RNA viruses are replicated by RdRps, whereas the prokaryotic and eukaryotic DNA genomes are replicated by DdDps. Whereas the RdRps do not require a primer to initiate replication of RNA genomes, the DdDps do require a primer for initiation of replication of DNA genomes. The DNA genome replication process in prokaryotes and eukaryotes is a complex one, but a well-orchestrated process, involving many enzymes and regulatory proteins, like primases, helicases, single-strand binding protein (SSB), replicases, repair enzymes, ligases, etc. These enzymes and proteins essentially form a multi-structural enzyme-protein complex known as primosome to initiate the very first step in the replication process. In prokaryotes like eubacteria, the replication initiation is started by a replication initiation protein which binds at the origin of replication (Ori), which is followed by the unwinding enzyme (replicative helicase) and SSBs. The primase binds finally to form the primosome and start synthesizing short RNA primers of 7-10 nucleotides (nts) in length.<sup>[1]</sup> However, the primosome complexes are much more complex in archaeabacteria and eukaryotes, where the primosome consists of many replication initiation proteins and a heterotetrameric primosome complex consisting of an RNA primase (a heterodimer with one small catalytic subunit and one large regulatory subunit) and the DNA pol  $\alpha$  (again a heterodimer with one large catalytic subunit and one small regulatory subunit).<sup>[2]</sup> Thus, in both the cases, an RNA segment is first synthesized

by the primosomes and then extended by the replicative polymerases. In archaeabacteria and eukaryotes, a DNA primer is made on the RNA primer by the second enzyme, the DNA pol  $\alpha$ . Despite significant advances in characterization of the processes involved in DNA replication, several basic questions remain unanswered. Therefore, understanding the DNA replication process will require not only our ability to connect the data available from various techniques like biochemical, genetic, site-directed mutagenesis (SDM), X-ray crystallography, cryo-EM, etc., but also to integrate them for a complete understanding of the replication process.

Primases (EC 2.7.7.102) are DNA-dependent RNA polymerases (DdRps) that initiate DNA replication in both prokaryotes and eukaryotes through the synthesis of short RNA fragments; once for the leading-strand and at multiple points for the lagging-strand, resulting in the synthesis of a large number of DNA fragments known as Okazaki fragments, named after the Japanese scientist, Okazaki, who discovered. All the three replicative polymerases are DdDps, belong to B family polymerases (pol), and possess both replication and repair functions. The B family pols are reported from both prokaryotes and eukaryotes, e.g., pol II, pol B, pol  $\alpha$ , pol  $\epsilon$ , pol  $\delta$ , Pol  $\zeta$ . They all exhibit 3'→5' proofreading function in addition to polymerization of deoxynucleoside triphosphates (dNTPs).<sup>[3]</sup> Prokaryotic chromosomes, mitochondrial DNA and circular viral DNA genomes have a single origin of replication. In contrast, eukaryotic chromosomes are very long and linear, and possess many origins of replication, which ensure replication initiation at multiple sites simultaneously. For example, in humans there are between 30,000 and 50,000 initiation sites.

The genome replication in both prokaryotes and eukaryotes involve three major steps, viz. initiation, elongation and termination. The initiation is the rate limiting step, which is

accomplished by the primosome complex. Interestingly, RNA replicases and primases do not require a primer template, as they initiate the RNA/primer synthesis with two NTPs (usually of purines, A or G) at the initiation site. The primases are broadly classified into single-subunit (SSU) primases (e.g., eubacterial and phage primases) and two-subunit primases (from eukaryotes, archaeabacteria and eukaryotic viruses). As eubacterial primases exhibit a markedly different structure as compared to the human primase, inhibition of eubacterial primase activity is expected to selectively inhibit bacterial DNA replication and further growth. In fact, growing evidences show such selective inhibition of bacterial DNA replication initiation has a bacteriocidal effect and therefore, inhibitors of bacterial primases are considered as effective antibiotics.<sup>[4]</sup>

The precise mechanism by which these processes are coordinated at the molecular level has remained largely obscure. In this communication, the enzymes and mechanism(s) involved in the initiation process are analyzed in detail and discussed.

## Materials and Methods

The protein sequence data of primases of prokaryotic and eukaryotic primases and DNA polymerase  $\alpha$  (DNA pol  $\alpha$ ) were obtained from PUBMED and SWISS-PROT databases. The advanced version of Clustal Omega was used for multiple sequence alignment (MSA) analysis of the proteins. Along with the conserved motifs identified by the bioinformatics analysis, and the data already available from biochemical, SDM and X-ray crystallographic analyses on the primases are used to confirm the possible amino acids that make up the active sites of different primases that initiate DNA replication in prokaryotes and eukaryotes.

## Results and Discussion

### Priming of Replication Initiation in Prokaryotes

DNA replication in eubacteria and eukaryotes follows more or less a similar mechanism. To initiate replication, it requires mainly the activity of the primase, which is a DdRp that synthesizes short RNA primers on the DNA template at the Origin of replication. The RNA primer is then extended by the DdDp, DNA polymerase III (DNA pol III), during replication of the genome. Based on their structure and sequence alignments, the prokaryotic primases are classified as a family of DnaG proteins. Primases which belong to this DnaG family, contain three distinct domains: A N-terminal domain (NTD) consisting of a zinc-binding motif, a middle domain with RNA polymerase activity and a C-terminal domain (CTD) that interacts with the replicative helicase for unwinding the DNA during the replication process. All the three domains are shown to be important for the synthesis of the primers and initiation of replication. In addition to the polymerase active sites, the polymerase domain of *E. coli* and T7 primases also possesses a TOPRIM fold, similar to topoisomerases. The binding of primase with the replicative helicase increases the binding affinity of the primase to SSBs and thus, enhances the rate of primer synthesis markedly. In fact, such interactions between the helicase and primase enhanced primer activity >1000-fold as compared to the activity of the isolated primase.<sup>[5]</sup>

As mentioned elsewhere, the priming reactions in bacterial replication are accomplished by the primosome multi-protein complex. The bacterial primosome complex essentially consists of an origin-of-replication initiator protein, a replicative helicase, a helicase loader, SSBs and a primase. Replication initiation requires a replication initiation protein (DnaA) that binds to the origin of replication, followed by unzipping the DNA by the

replicative helicase (DnaB). Now, the replication process starts with unwinding of the dsDNA at the origin of replication to generate ssDNA. DnaC mediates loading of the DnaB to the DnaA at the origin of replication site. SSBs then bind to the single-stranded region to

prevent reannealing of the DNA strands so that the primase can start synthesizing the primer (Fig.1). The primase (DnaG) also interacts with the DNA pol III holoenzyme, which helps the translocation of the primer to the DNA pol III active site for further extension (Fig.1).

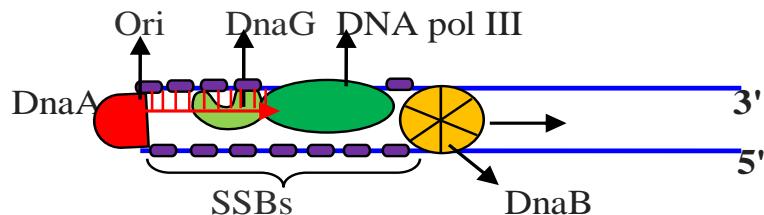
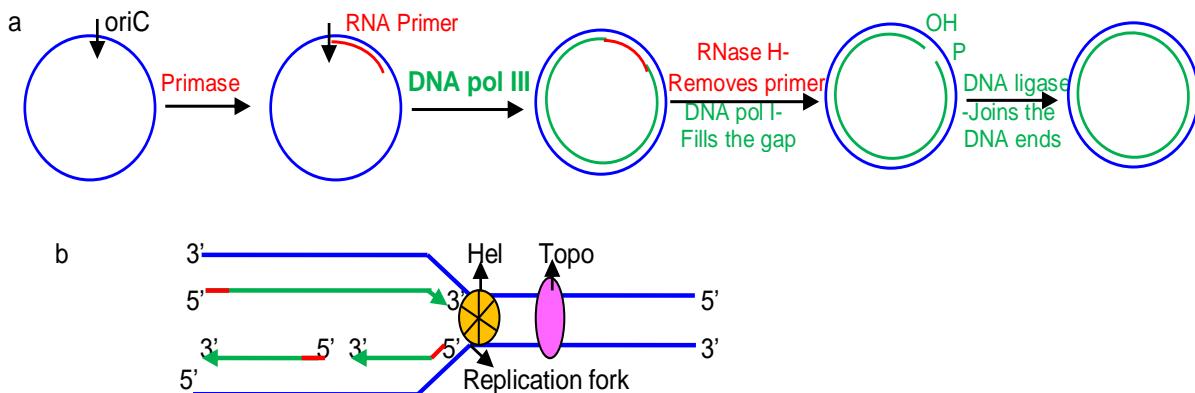


Fig. 1 A schematic diagramme showing the assembly of various proteins at the bacterial replication initiation complex

Blue lines, DNA; Red line, RNA primer; DnaA, replication initiation protein; DnaG, primase; SSBs, single-strand DNA-binding proteins; DnaB, replicative helicase, a hexamer (the helicase is also called the unwinding enzyme and it catalyzes the separation of the DNA strands; the energy required for the process is provided by the hydrolysis of ATP).

The DnaB/DnaG complex synthesizes RNA primers for both leading- and lagging-strand of the replication process. Leading-strand replication needs to be primed once, which takes place at the origin of replication whereas lagging-strand synthesis needs repeated priming and extensions, generating short DNA fragments on the lagging-stand which are also known as ‘Okazaki fragments’. The primase makes only short primers, pppAGN(7-10) with 7-10 nts in length which are extended by the DNA pol III which slides continuously along the leading-strand, but in sections on the lagging-strand. Therefore, the lagging-strand is synthesized discontinuously, (i.e.), the DNA pol III begins at the 3'-end of the RNA primers at each section and extends, thus, generating large number of Okazaki fragments on the lagging-strand (Fig. 2). Both in prokaryotes and eukaryotes, the primase initiation site is specific for ATP or GTP.<sup>[6]</sup>

In the finishing steps, RNase H (an enzyme specific DNA-RNA hybrids) removes the primers and the DNA polymerase I (DNA pol I) which recognize a “nick” or break on the phosphate backbone fills the gaps with dNTPs, and is followed by DNA ligase which covalently joins the two ends and thus, completing the replication process (Fig. 2).<sup>[7]</sup> As mentioned earlier, during replication of bacterial genomes, the primase complex initiate leading-strand synthesis once at the origin of replication, which is extended by the DNA pol III, until completion of the genome and stops at the 5'-end of the primer. However, at the lagging-strand synthesis, it primes at multiple points to complete the replication. In fact, during lagging-strand synthesis of the *E. coli* genome, the primase complex transcribes 2000 to 3000 RNA primers at a rate of about one primer per second.<sup>[6]</sup> (The *E. coli* chromosome is circular and, 4600 kb in length and has a single origin of replication. The entire replication process is completed in ~ 20 min).



*Figs. 2a and 2b. Schematic diagrammes showing the steps involved in the replication of *E. coli* chromosome. Ori, origin of replication.*

2a shows steps involved in the replication of *E. coli* chromosome, and 2b shows the continuous synthesis of DNA on the leading-strand and the generation of Okazaki fragments on the lagging-strand. Hel, helicase; Topo, Topoisomerase.

**Structural Features of Eubacterial Primases**  
On the basis of sequence analysis, the eubacterial primases are structurally distinct from the primases that initiates archaeabacterial and eukaryotic replication. Among eubacterial primases, the *E. coli* primase is the most extensively studied one. A limited proteolytic analysis of the *E. coli* primase yielded three

structural domains, viz. a 12-kD NTD consists of a Zn<sup>2+</sup>-binding motif (ZBM), followed by a 36-kDa core domain with RNA polymerase activity (RNAP domain) and a CTD, also known as DnaB-interacting domain (DnaB-ID) (Fig. 3). The CTD is responsible for interaction with the replicative helicase. [8]



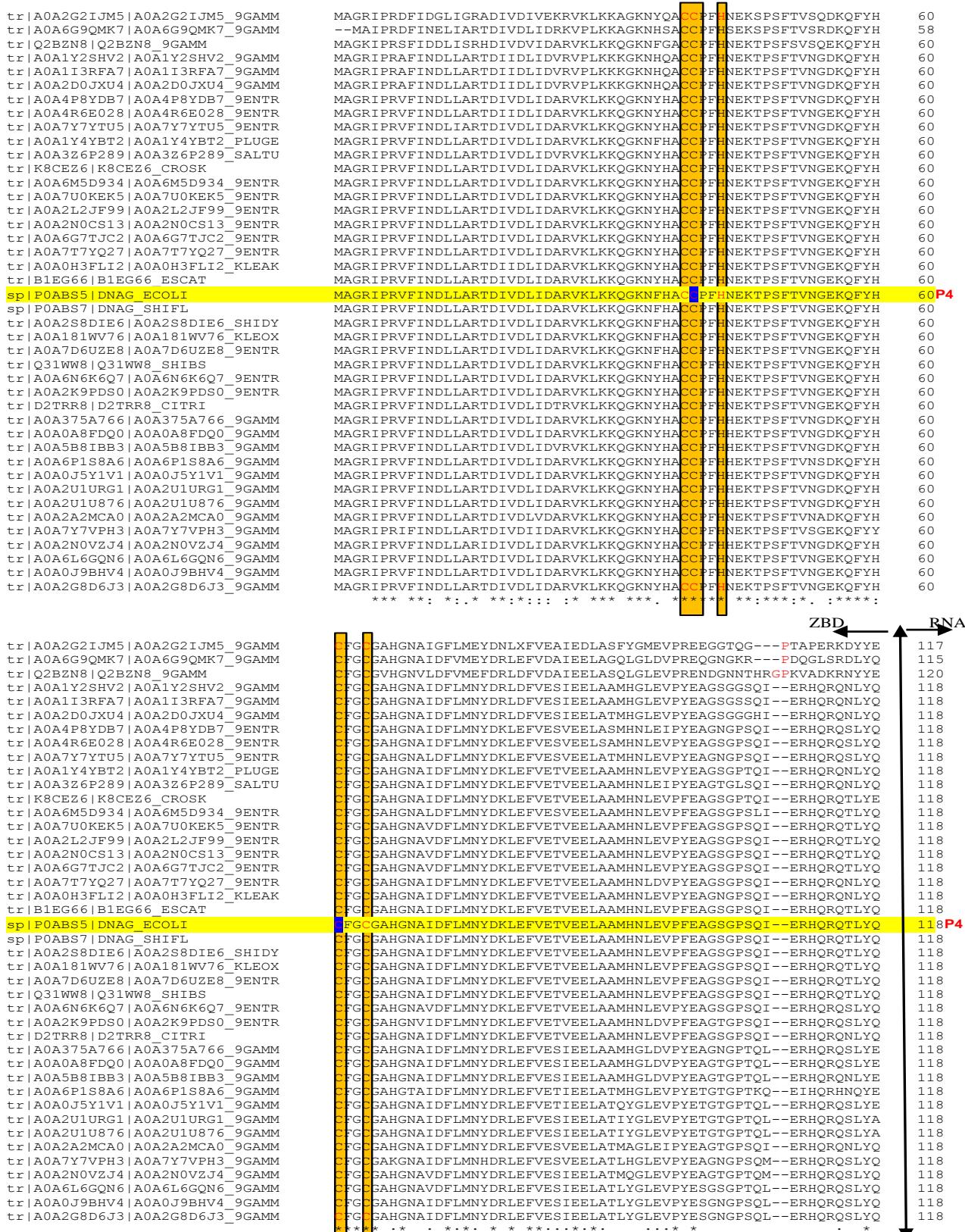
*Fig. 3 Functional domains of a eubacterial primase (numbering from the *E. coli* primase)*

Adapted from [9]

Figure 4 shows the MSA analysis of the catalytic subunit of eubacterial primases (only the required regions in the discussion are shown). It shows many highly conserved peptides regions, motifs and triads, diads, and also single amino acid conservations. The primase catalytic subunit from different bacterial sources shows an invariant -YG-, (highlighted in yellow), which is implicated in template-binding in many DdDps and DdRps. A proposed catalytic pair -KG- is also found at the upstream of the -YG- pair (highlighted in yellow).

The catalytic K and a E of the primase are based on the SDM data from P4 phage primase and is highlighted in dark blue. [10] The template binding pair is followed by two typical metal-binding -DxD- motifs, which is reported in most of the other DdDps and DdRps and the metal-binding motifs are found within two highly conserved peptides (highlighted in green). A -CxxH—CxxC- type ZBM (highlighted in orange) is located in the NTD as reported in X-ray crystallographic data. [9] SDM analysis of the ZBM from P4 primase is highlighted in dark blue. [10]

## CLUSTAL O (1.2.4) MSA of eubacterial primases



tr A0A2G2IJM5 A0A2G2IJM5_9GAMM	TEDQLLTTGVLIENEGSRRYDRFRE	RVMFFIRDKR GRVIG CGRVLG DEK PKYLNSPET	237
tr A0A6G9QMK7 A0A6G9QMK7_9GAMM	AQDKLLTAGMLIEND-SGKRYDRFRDRMLMFPIRDRI	GRVIG CGRVLG DEK PKYLNSPET	234
tr Q2BZB8 Q2BZB8_9GAMM	SQKALVNTGLLIEND-NGRRYDRFRGRVMPFIHD	GRVIG CGRVLG DGT PKYLNSPET	238
tr A0A1Y2SHV2 A0A1Y2SHV2_9GAMM	DRKQLNDAGMLVTND-NGRTYDRFRGRVMPFI	GRVIG CGRVLG DGT PKYLNSPET	236
tr A0A1I3RFA7 A0A1I3RFA7_9GAMM	DRKQLNDAGMLVTND-NGRTYDRFRGRVMPFI	GRVIG CGRVLG DGT PKYLNSPET	236
tr A0A2D0JXU4 A0A2D0JXU4_9GAMM	DRNQLSDAGMLVTND-SGRTYDRFRGRIMPFI	GRVIG CGRVLG DGT PKYLNSPET	236
tr A0A4P8YDB7 A0A4P8YDB7_9ENTR	NRKSLTDAGMLVTND-RGRSYDRFRGRVMPFI	GRVIG CGRVLG DGT PKYLNSPET	236
tr A0A4R6E028 A0A4R6E028_9ENTR	NRQALTDAAGMLVTND-QGRSYDRFRGRVMPFI	GRVIG CGRVLG DGT PKYLNSPET	236
tr A0A7Y7YTU5 A0A7Y7YTU5_9ENTR	NRQALTDAAGMLVTND-QGRSYDRFRGRVMPFI	GRVIG CGRVLG DGT PKYLNSPET	236
tr A0A1Y4YBT2 A0A1Y4YBT2_PLUGE	NKQSLIDAGMLVTND-QGRSYDRFRGRVMPFI	GRVIG CGRVLG DGM PKYLNSPET	236
tr A0A3Z6P289 A0A3Z6P289_SALTU	NKALLNDAGMLVNNE-QGRTYDRFRGRVMPFI	GRVIG CGRVLG DGM PKYLNSPET	236
tr K8CEZ6 K8CEZ6_CROSK	NRKSLIDAGMLVTND-QGRSYDRFRGRVMPFI	GRVIG CGRVLG DGT PKYLNSPET	236
tr A0A6M5D934 A0A6M5D934_9ENTR	NRKSLIDAGMLVTND-QGRSYDRFRGRVMPFI	GRVIG CGRVLG DGT PKYLNSPET	236
tr A0A7U0KEK5 A0A7U0KEK5_9ENTR	DRKSLIDAGMLVTND-QGRSYDRFRGRVMPFI	GRVIG CGRVLG DGT PKYLNSPET	236
tr A0A2L2JF99 A0A2L2JF99_9ENTR	DRKSLIDAGMLVTND-QGRSYDRFRGRVMPFI	GRVIG CGRVLG DGT PKYLNSPET	236
tr A0A2N0CS13 A0A2N0CS13_9ENTR	DRKSLIDAGMLVTND-KGRSYDRFRGRVMPFI	GRVIG CGRVLG DGT PKYLNSPET	236
tr A0A6G7TJC2 A0A6G7TJC2_9ENTR	DRKSLIDAGMLVTND-KGRSYDRFRGRVMPFI	GRVIG CGRVLG DGT PKYLNSPET	236
tr A0A7T7YQ27 A0A7T7YQ27_9ENTR	NRQSLIDAGMLVTND-QGRSYDRFRGRVMPFI	GRVIG CGRVLG DGT PKYLNSPET	236
tr A0A0H3FLI2 A0A0H3FLI2_KLEAK	NRQSLIDAGMLVTND-QGRSYDRFRGRVMPFI	GRVIG CGRVLG DGT PKYLNSPET	236
tr B1EG66 B1EG66_ESCAT	NRQSLINAGMLVTND-QGRSYDRFRGRVMPFI	GRVIG CGRVLG DGT PKYLNSPET	236
sp POABS5 DNAG_ECOLI	NRQSLIDAGMLVTND-QGRSYDRFRGRVMPFI	GRVIG CGRVLG DGT PKYLNSPET	236
sp POABS7 DNAG_SHIFL	E. cc		
tr A0A2S8DIE6 A0A2S8DIE6_SHIDY	NRQSLIDAGMLVTND-QGRSYDRFRGRVMPFI	GRVIG CGRVLG DGT PKYLNSPET	236
tr A0A181WV76 A0A181WV76_KLEOX	NRQSLIDAGMLVTND-QGRSYDRFRGRVMPFI	GRVIG CGRVLG DGT PKYLNSPET	236
tr A0A7D6UZE8 A0A7D6UZE8_9ENTR	NRQSLIDAGMLVTND-QGRSYDRFRGRVMPFI	GRVIG CGRVLG DGT PKYLNSPET	236
tr Q31IW8 Q31IW8_SHIBS	NRQSLIDAGMLVTND-QGRSYDRFRGRVMPFI	GRVIG CGRVLG DGT PKYLNSPET	236
tr A0A6N6K6Q7 A0A6N6K6Q7_9ENTR	NRQSLIDAGMLVTND-QGRSYDRFRGRVMPFI	GRVIG CGRVLG DGT PKYLNSPET	236
tr A0A2K9PDSO A0A2K9PDSO_9ENTR	NRQSLIDAGMLVTND-QGRSYDRFRGRVMPFI	GRVIG CGRVLG DGT PKYLNSPET	236
tr D2TRR8 D2TRR8_CITRI	NRATLTDAAGMLVTND-NGRTYDRFRGRVMPFI	GRVIG CGRVLG DGT PKYLNSPET	236
tr A0A375A766 A0A375A766_9GAMM	DKTTLTDAAGMLVTNE-NGRTYDRFRGRVMPFI	GRVIG CGRVLG DGT PKYLNSPET	236
tr A0A0A8FDQ0 A0A0A8FDQ0_9GAMM	NRATLTDAAGMLVTNE-NGRTYDRFRGRVMPFI	GRVIG CGRVLG DGT PKYLNSPET	236
tr A0A5B8IBB3 A0A5B8IBB3_9GAMM	DRNLNDAGMLVTNE-NGRTYDRFRGRVMPFI	GRVIG CGRVLG DGT PKYLNSPET	235
tr A0A6P1S8A6 A0A6P1S8A6_9GAMM	DRSTLNDAAGMLVTNE-NGRTYDRFRGRVMPFI	GRVIG CGRVLG DGT PKYLNSPET	236
tr A0A0J5Y1V1 A0A0J5Y1V1_9GAMM	DRSALNDAGMLVTNE-NGRTYDRFRGRVMPFI	GRVIG CGRVLG DGT PKYLNSPET	236
tr A0A2U1URG1 A0A2U1URG1_9GAMM	DRSALNDAGMLVTNE-NGRTYDRFRGRVMPFI	GRVIG CGRVLG DGT PKYLNSPET	236
tr A0A2U1U876 A0A2U1U876_9GAMM	SRTALNDAGMLVNN-NGRTYDRFRGRVMPFI	GRVIG CGRVLG DGT PKYLNSPET	236
tr A0A2A2MCA0 A0A2A2MCA0_9GAMM	DRESLMEAGMLVNSD-SGRTYDRFRGRVMPFI	GRVIG CGRVLG DGT PKYLNSPET	236
tr A0A7Y7VPH3 A0A7Y7VPH3_9GAMM	DKASLIDAGMLVTNE-NGRSYDRFRGRVMPFI	GRVIG CGRVLG DGT PKYLNSPET	236
tr A0A2N0VZJ4 A0A2N0VZJ4_9GAMM	DKQSLIDAGMLVTNE-NGRSYDRFRGRVMPFI	GRVIG CGRVLG DGT PKYLNSPET	236
tr A0A6L6GQN6 A0A6L6GQN6_9GAMM	DRQSLIDAGMLVTNE-NGRSYDRFRGRVMPFI	GRVIG CGRVLG DGT PKYLNSPET	236
tr A0A0J9BHV4 A0A0J9BHV4_9GAMM	DRQSLIDAGMLVTNE-NGRSYDRFRGRVMPFI	GRVIG CGRVLG DGT PKYLNSPET	236
tr A0A2G8D6J3 A0A2G8D6J3_9GAMM			
tr A0A2G2IJM5 A0A2G2IJM5_9GAMM	P1FHKGEL YGLYEVRQAYKE EIFDQIVVV EGYMDVVALAQFGIDYAVASLGTSTSGXQMQT		297
tr A0A6G9QMK7 A0A6G9QMK7_9GAMM	P1FHKGEL YGLYEYLKQRHEDFQVHLV EGYMDVVALAQFGIDYAVASLGTSTTAEQFQL		294
tr Q2BZB8 Q2BZB8_9GAMM	DIFHKGRDI YGLYEVLQTHREPRKIMRVH EGYMDVVALAQFGVDYAVASLGTSTTGDHIQS		298
tr A0A1Y2SHV2 A0A1Y2SHV2_9GAMM	EIFHKGRDI YGLYEAQQSHSTTLTRLVV EGYMDVVALAQFGIDYAVASLGTSTAEHVQL		296
tr A0A1I3RFA7 A0A1I3RFA7_9GAMM	EIFHKGRDI YGLYEAQQNHFLVSLKLVV EGYMDVVALAQFGIDYAVASLGTSTAEHIQL		296
tr A0A2D0JXU4 A0A2D0JXU4_9GAMM	EIFHKGRDI YGLYEAQQSHNTLSRLVV EGYMDVVALAQFGIDYAVASLGTSTAEHVQL		296
tr A0A4P8YDB7 A0A4P8YDB7_9ENTR	EIFHKGRDI YGLYEARQDTSDFARLLV EGYMDVVALAQYDINYAVASLGTSTADHIQL		296
tr A0A4R6E028 A0A4R6E028_9ENTR	DIFHKGRDI YGLYEAQOSQSEPARLRLVV EGYMDVVALAQYQGSI SYAVASLGTSTADHIQL		296
tr A0A7Y7YTU5 A0A7Y7YTU5_9ENTR	DIFHKGRDI YGLYEAQOSQEPARLRLVV EGYMDVVALAQYQGSI SYAVASLGTSTADHIQL		296
tr A0A1Y4YBT2 A0A1Y4YBT2_PLUGE	DIFHKGRDI YGLYEAQQAAAEPQLRLVV EGYMDVVALAQYDINYAVASLGTSTADHIQL		296
tr A0A3Z6P289 A0A3Z6P289_SALTU	DIFHKGRDI YGLYEAQQYSEFPLRLVV EGYMDVVALAQYQDINYAVASLGTSTADHMHM		296
tr K8CEZ6 K8CEZ6_CROSK	DIFHKGRDI YGLYEVQQSDENPPRLVV EGYMDVVALAQYDINYAVASLGTSTAEHIQM		296
tr A0A6M5D934 A0A6M5D934_9ENTR	DIFHKGRDI YGLYEAQLDNEEPQLRLVV EGYMDVVALAQYDINYAVASLGTSTADHIQL		296
tr A0A7U0KEK5 A0A7U0KEK5_9ENTR	DIFHKGRDI YGLYEAQODNAEEPRLRLVV EGYMDVVALAQYDINYAVASLGTSTADHIQL		296
tr A0A2L2JF99 A0A2L2JF99_9ENTR	DIFHKGRDI YGLYEAQODNAEEPRLRLVV EGYMDVVALAQYDINYAVASLGTSTADHIQL		296
tr A0A2N0CS13 A0A2N0CS13_9ENTR	DIFHKGRDI YGLYEAQODNAEEPRLRLVV EGYMDVVALAQYDINYAVASLGTSTADHIQL		296
tr A0A6G7TJC2 A0A6G7TJC2_9ENTR	DIFHKGRDI YGLYEAQODNAEEPRLRLVV EGYMDVVALAQYDINYAVASLGTSTADHIQL		296
tr A0A7T7YQ27 A0A7T7YQ27_9ENTR	DIFHKGRDI YGLYEAQODNAEEPRLRLVV EGYMDVVALAQYDINYAVASLGTSTADHIQL		296
tr A0A0H3FLI2 A0A0H3FLI2_KLEAK	DIFHKGRDI YGLYEAQODNAEEPRLRLVV EGYMDVVALAQYDINYAVASLGTSTADHIQL		296
tr B1EG66 B1EG66_ESCAT	DIFHKGRDI YGLYEAQODNAEEPRLRLVV EGYMDVVALAQYDINYAVASLGTSTADHIQL		296
sp POABS5 DNAG_ECOLI	P4		
sp POABS7 DNAG_SHIFL			
tr A0A2S8DIE6 A0A2S8DIE6_SHIDY	DIFHKGRDI YGLYEAQODNAEPNRLVV EGYMDVVALAQYGINYAVASLGTSTADHIQL		296
tr A0A181WV76 A0A181WV76_KLEOX	DIFHKGRDI YGLYEAQODNAEPNRLVV EGYMDVVALAQYGINYAVASLGTSTADHIQL		296
tr A0A7D6UZE8 A0A7D6UZE8_9ENTR	DIFHKGRDI YGLYEAQODNAEPNRLVV EGYMDVVALAQYGINYAVASLGTSTADHIQL		296
tr Q31IW8 Q31IW8_SHIBS	DIFHKGRDI YGLYEAQODNAEPNRLVV EGYMDVVALAQYGINYAVASLGTSTADHIQL		296
tr A0A6N6K6Q7 A0A6N6K6Q7_9ENTR	DIFHKGRDI YGLYEAQODNAEPNRLVV EGYMDVVALAQYGINYAVASLGTSTADHIQL		296
tr A0A2K9PDSO A0A2K9PDSO_9ENTR	DIFHKGRDI YGLYEAQODDAEPNRLVV EGYMDVVALAQYGINYAVASLGTSTADHIQL		296
tr D2TRR8 D2TRR8_CITRI	DIFHKGRDI YGLYEAQODNAEPNRLVV EGYMDVVALAQYGINYAVASLGTSTADHIQL		296
tr A0A375A766 A0A375A766_9GAMM	DIFHKGRDI YGLYEAQODNAEPNRLVV EGYMDVVALAQYGINYAVASLGTSTADHIQL		296
tr A0A0A8FDQ0 A0A0A8FDQ0_9GAMM	EIVFHKGDI YGLYEAQQKNEELKRLVV EGYMDVVALAQFGIDYAVASLGTSTADHIQL		296
tr A0A5B8IBB3 A0A5B8IBB3_9GAMM	EIFHKGRDI YGLYEAQQKNEELKRLVV EGYMDVVALAQFGIDYAVASLGTSTADHIQL		296
tr A0A6P1S8A6 A0A6P1S8A6_9GAMM	EIFHKGRDI YGLYEAQQKNEELKRLVV EGYMDVVALAQFGIDYAVASLGTSTADHIQL		296
tr A0A0J5Y1V1 A0A0J5Y1V1_9GAMM	GIFHKGRDI YGLYEAQQNHFLDKRLVV EGYMDVVALAQFGIDYAVASLGTSTADHIQL		295
tr A0A2U1URG1 A0A2U1URG1_9GAMM	EIFHKGRDI YGLYEAQQNHFLDKRLVV EGYMDVVALAQFGIDYAVASLGTSTADHIQL		296
tr A0A2U1U876 A0A2U1U876_9GAMM	EIFHKGRDI YGLYEAQQNHFLDKRLVV EGYMDVVALAQFGIDYAVASLGTSTADHIQL		296
tr A0A2A2MCA0 A0A2A2MCA0_9GAMM	DIFHKGRDI YGLYEAQQDNEELARLRLVV EGYMDVVALAQFGIDYAVASLGTSTSEHIQL		296
tr A0A7Y7VPH3 A0A7Y7VPH3_9GAMM	PIFHKGDI YGLYEAQQKNEQPTRLVV EGYMDVVALAQFGIDYAVASLGTSTAEHVQL		296
tr A0A2N0VZJ4 A0A2N0VZJ4_9GAMM	EIFHKGRDI YGLYEAKLKHESPAKLRLVV EGYMDVVALAQFGIDYAVASLGTSTADHIQL		296
tr A0A6L6GQN6 A0A6L6GQN6_9GAMM	DIFHKGRDI YGLYEAKLKHESPAKLRLVV EGYMDVVALAQFGIDYAVASLGTSTADHIQL		296
tr A0A0J9BHV4 A0A0J9BHV4_9GAMM	DIFHKGRDI YGLYEAKQNYHEPPKLLVV EGYMDVVALAQFGIDYAVASLGTSTADHIQL		296
tr A0A2G8D6J3 A0A2G8D6J3_9GAMM	DIFHKGRDI YGLYEAKQNYHEPPKLLVV EGYMDVVALAQFGIDYAVASLGTSTADHIQL		296

tr A0A2G2IJM5 A0A2G2IJM5_9GAMM	LFRNTSQ VICCY DGD KAGRDAAWRALE NALPQLRDGKE N FVFLPDGE DPD S LVRHQGKD	357
tr A0A6G9QMK7 A0A6G9QMK7_9GAMM	LVRSAKE VICCY DGD KAGREAAWRALE TALPLLKPGDH VKFMFLPQGE DPD IMVRKIGKD	354
tr Q2BZN8_9GAMM	LFRQST VICCY DGD RAGRDAAWRAME QALPYLTDGRQ LKFMFLPDGE DPD CIRAEGKN	358
tr A0A1Y2SHV2 A0A1Y2SHV2_9GAMM	LFRTTDS VICCY DGD RAGRDAAWRLE TALPYLNDGRQ LRFMFLLPDGE DPD SLIRKEGQE	356
tr A0A1I3RFA7 A0A1I3RFA7_9GAMM	LFRTTDN VICCY DGD RAGRDAAWRLE TALPYLNDGRQ LRFMFLLPDGE DPD S LVRKEGRE	356
tr A0A2D0JXU4 A0A2D0JXU4_9GAMM	LFRTTDS VICCY DGD RAGR TAAWRLE TALPYLNDGRQ LRFMFLLPDGE DPD S LVRKEGRE	356
tr A0A4P8YDB7 A0A4P8YDB7_9ENTR	LFRVTNT VICCY DGD RAGRDAAWRLE TALPYMTDGRQ LRFMFLLPDGE DPD ILVRKEGKT	356
tr A0A4R6E028 A0A4R6E028_9ENTR	LFRVTDT VICCY DGD RAGRDAAWRLE TALPYMTDGRQ LRFMFLLPDGE DPD ILVRKEGKE	356
tr A0A7Y7YTU5 A0A7Y7YTU5_9ENTR	LFRVTDT VICCY DGD RAGRDAAWRLE TALPYMTDGRQ LRFMFLLPDGE DPD ILVRQEGKD	356
tr A0A1Y4YBT2 A0A1Y4YBT2_PLUGE	LFRVTHN VICCY DGD RAGRDAAWRLE TALPYMTDGRQ LRFMFLLPDGE DPD ILVRKEGKA	356
tr A0A3Z6P289_SALTU	LFRATNN VICCY DGD RAGRDAAWRLE TAMPYMTDGRQ LRFMFLLPDGE DPD ILVRKEGKA	356
tr K8CEZ6_K8CEZ6_CROSK	LFRVTNN VICCY DGD RAGRSAAWRLE TALPYMTDGRQ LRFMFLLPDGE DPD ILVRQEGKA	356
tr A0A6M5D934 A0A6M5D934_9ENTR	LFRVTKN VICCY DGD RAGRDAAWRLE TALPYMSDGRQ LRFMFLLPDGE DPD ILVRKEGKA	356
tr A0A7U0KEK5 A0A7U0KEK5_9ENTR	LFRVTNN VICCY DGD RAGRDAAWRLE TALPYMTDGRQ LRFMFLLPDGE DPD ILVRKEGKA	356
tr A0A2L2JF99 A0A2L2JF99_9ENTR	LFRVTNN VICCY DGD RAGRDAAWRLE TALPYMTDGRQ LRFMFLLPDGE DPD ILVRKEGKA	356
tr A0A2N0CS13 A0A2N0CS13_9ENTR	LFRATNN VICCY DGD RAGRDAAWRLE TALPYMTDGRQ LRFMFLLPDGE DPD ILVRKEGKA	356
tr A0A6G7TJC2 A0A6G7TJC2_9ENTR	LFRVTNT VICCY DGD RAGREAAWRALE TALPYMTDGRQ LRFMFLLPDGE DPD ILVRKEGKA	356
tr A0A7T7YQ27 A0A7T7YQ27_9ENTR	LFRVTNQ VICCY DGD RAGRDAAWRLE TALPYMTDGRQ LRFMFLLPDGE DPD ILVRKEGKD	356
tr A0A0H3FL12 A0A0H3FL12_KLEAK	LFRATNQ VICCY DGD RAGRDAAWRLE TALPYMTDGRQ LRFMFLLPDGE DPD ILVRQEGKE	356
tr B1EG66_B1EG66_ESCAT	LFRATNQ VICCY DGD RAGRDAAWRLE TALPYMTDGRQ LRFMFLLPDGE DPD ILVRKEGKE	356
sp P0ABS5 DNAG_ECOLI	LFRATNN VICCY DGD RAGRDAAWRLE TALPYMTDGRQ LRFMFLLPDGE DPD ILVRKEGKE	356
<b>E. coli</b>		
sp P0ABS7 DNAG_SHIFL		
tr A0A2S8DIE6 A0A2S8DIE6_SHIDY	LFRATNN VICCY DGD RAGRDAAWRLE TALPYMTDGRQ LRFMFLLPDGE DPD ILVRKEGKE	356
tr A0A181WV76 A0A181WV76_KLEOX	LFRATNN VICCY DGD RAGRDAAWRLE TALPYMTDGRQ LRFMFLLPDGE DPD ILVRKEGKE	356
tr A0A7D6U2E8 A0A7D6U2E8_9ENTR	LFRATNN VICCY DGD RAGRDAAWRLE TALPYMTDGRQ LRFMFLLPDGE DPD ILVRKEGKE	356
tr Q31WW8_Q31WW8_SHIBS	LFRATNN VICCY DGD RAGRDAAWRLE TALPYMTDGRQ LRFMFLLPDGE DPD ILVRKEGKE	356
tr A0A6N6K6Q7 A0A6N6K6Q7_9ENTR	LFRATNN VICCY DGD RAGRDAAWRLE TALPYMTDGRQ LRFMFLLPDGE DPD ILVRKEGKE	356
tr A0A2K9PDS0 A0A2K9PDS0_9ENTR	LFRATNN VICCY DGD RAGRDAAWRLE TALPYMTDGRQ LRFMFLLPDGE DPD ILVRKEGKE	356
tr D2TRR8_D2TRR8_CITRI	LFRATNN VICCY DGD RAGRDAAWRLE TALPYMTDGRQ LRFMFLLPDGE DPD ILVRKEGKE	356
tr A0A375A766 A0A375A766_9GAMM	LFRATDK VICCY DGD RAGRDAAWRLE TALPYLDDGRQ LHFMFLLPDGE DPD ILVRKEGHT	356
tr A0A0A8FDQ0 A0A0A8FDQ0_9GAMM	LFRATDQ VICCY DGD RAGRDAAWRLE TALPYLDDGRQ LRFMFLLPDGE DPD ILVRQEGKA	356
tr A0A5B8IBB3 A0A5B8IBB3_9GAMM	LFRATDQ VICCY DGD RAGRDAAWRLE TALPYLDDGRQ LRFMFLLPDGE DPD ILVRQEGKA	356
tr A0A6P1S8A6 A0A6P1S8A6_9GAMM	LFRATDL VICCY DGD RAGREAAWRALE TALPYLNDGRQ LRFMFLLPDGE DPD ILVRKEGKA	355
tr A0A0J5Y1V1 A0A0J5Y1V1_9GAMM	LFRATDL VICCY DGD RAGREAAWRALE TALPYLNDGRQ LRFMFLLPDGE DPD ILVRKEGKA	356
tr A0A2U1URG1 A0A2U1URG1_9GAMM	LFRVTDQ VICCY DGD RAGREAAWRALE TALPYLNDGRQ LRFMFLLPDGE DPD ILVRKEGKV	356
tr A0A2U1U876 A0A2U1U876_9GAMM	LFRATDQ VICCY DGD RAGREAAWRALE TALPYLNDGRQ LRFMFLLPDGE DPD ILVRKESKA	356
tr A0A2A2MCA0 A0A2A2MCA0_9GAMM	MFRSTD VICCY DGD RAGREAAWRALE TALFLSDGRQ LKFMFLLPDGE DPD ILVRKEGTA	356
tr A0A7Y7VPH3 A0A7Y7VPH3_9GAMM	LYRSTD VICCY DGD RAGREAAWRTLE TALPYMNDGRQ LRFMFLLPDGE DPD ILVRKEGKE	356
tr A0A2NOVZJ4 A0A2NOVZJ4_9GAMM	MFRSTD VICCY DGD RAGREAAWRALE TALPYLNDGRQ LRFMFLLPDGE DPD ILVRKEGKE	356
tr A0A6L6GQN6 A0A6L6GQN6_9GAMM	LFRATGT VICCY DGD RAGREAAWRALE TALPYMNDGRQ LRFMFLLPDGE DPD ILIRKEGKE	356
tr A0A0J9BHV4 A0A0J9BHV4_9GAMM	LFRATNT VICCY DGD RAGREAAWRALE TALPYMNDGRQ LRFMFLLPDGE DPD ILVRKEGKE	356
tr A0A2G8D6J3 A0A2G8D6J3_9GAMM	LFRATNT VICCY DGD RAGREAAWRALE TALPYMNDGRQ LRFMFLLPDGE DPD ILVRKEGKE	356
:	: * : . * : . * : * . * : * . * : * . * : * . * : * . * : * . * : * . * : * .	.

Fig. 4 MSA of primases from different eubacterial sources

- A0A2G2IJM5\_9GAMM*, *Moritella sp.*  
*Q2BZN8\_9GAMM*, *Photobacterium sp.*  
*A0A1I3RFA7\_9GAMM*, *Xenorhabdus mauleonii*  
*A0A4P8YDB7\_9ENTR*, *Jejubacter calystegiae*  
*A0A7Y7YTU5\_9ENTR*, *Cedecea sp.*  
**A0A3Z6P289\_SALTU**, *Salmonella typhimurium*  
*A0A6M5D934\_9ENTR*, *Kosakonia sp.*  
*A0A2L2JF99\_9ENTR*, *Lelliottia nimipressuralis*  
*A0A6G7TJC2\_9ENTR*, *Leclercia sp.*  
**A0A0H3FL12\_KLEAK**, *Klebsiella aerogenes*  
**P0ABS5/DNAG\_ECOLI**, *Escherichia coli (K12)*  
**A0A2S8DIE6\_SHIDY**, *Shigella dysenteriae*  
*A0A7D6UZE8\_9ENTR*, *Enterobacter hormaechei*  
*A0A6N6K6Q7\_9ENTR*, *Citrobacter pasteurii*  
*D2TRR8\_CITRI*, *Citrobacter rodentium*  
*A0A0A8FDQ0\_9GAMM*, *Dickeya zeae*  
*A0A6P1S8A6\_9GAMM*, *Pectobacterium polaris*  
*A0A2U1U876\_9GAMM*, *Brenneria sp.*  
*A0A2A2MCA0\_9GAMM*, *Hafnia paralvei*  
*A0A2N0CS13\_9ENTR*, *Cedecea lapagei*  
*A0A2G8D6J3\_9GAMM*, *Erwinia sp.*
- A0A6G9QMK7\_9GAMM*, *Shewanella aestuarii*  
*A0A1Y2SHV2\_9GAMM*, *Xenorhabdus vietnamensis*  
*A0A2D0JXU4\_9GAMM*, *Xenorhabdus sp.*  
*A0A4R6E028\_9ENTR*, *Buttiauxella sp.*  
*A0A1Y4YBT2\_PLUGE*, *Pluralibacter gergoviae*  
**K8CEZ6\_CROSK**, *Cronobacter sakazaki*  
*A0A7U0KEK5\_9ENTR*, *Enterobacter hormaechei*  
*A0A2N0VZJ4\_9GAMM*, *Rahnella sp.*  
*A0A7T7YQ27\_9ENTR*, *Klebsiella michiganensis*  
*B1EG66\_ESCAT*, *Escherichia albertii*  
**P0ABS7/DNAG\_SHIFL**, *Shigella flexneri*  
*A0A181WV76\_KLEOX*, *Klebsiella oxytoca*  
*Q31WW8\_SHIBS*, *Shigella boydii*  
*A0A2K9PDS0\_9ENTR*, *Citrobacter freundii*  
*A0A375A766\_9GAMM*, *Dickeya aquatic*  
*A0A5B8IBB3\_9GAMM*, *Dickeya poaceiphila*  
*A0A0J5Y1V1\_9GAMM*, *Pectobacterium peruviense*  
*A0A2U1URG1\_9GAMM*, *Brenneria nigrifluens*  
*A0A7Y7VPH3\_9GAMM*, *Pantoea sp.*  
*A0A0J9BHV4\_9GAMM*, *Frankliniella occidentalis*

### Active Site Analyses of Eubacterial Primases

A great deal of information is available on the primase from *E. coli*. For example, X-ray crystallographic data of *E. coli* primase showed that it is a modular, cashew-shaped molecule with the dimensions of 30 Å by 35 Å by 75 Å, and is composed of three distinct subdomains. The ZBM is located at the NTD and is in close proximity to the RNAP active site. The zinc ion is found to be coordinated by 3 Cys and 1 His residues and resembled a zinc-knuckle rather than the well-known zinc-finger of other primases.<sup>[9]</sup> These findings are further corroborated by the MSA analysis too. A ZBM with 3 Cys and 1 His residues (-CxxH -----CxxC-) (highlighted in orange) are found at the NTD. The involvement of the particular R<sup>221</sup> (highlighted in dark blue) in catalysis was further proved by SDM experiments.<sup>[11]</sup> It is interesting to note that two very similar direct repeats of pentapeptides (marked by arrows and highlighted) are found upstream to the proposed catalytic amino acid K, (Fig. 4), suggesting a possible region for the first two purine binding sites during primer initiation. SDM analysis of bacteriophage P4 primase provided additional insight into the *E. coli* primase functional domains. The bacteriophage P4 DNA replication depends on the product of the α gene, which is a multifunctional enzyme and harbours three different activities, viz. origin of replication recognition, DNA helicase and primase activities. By SDM experiments, Ziegelin et al. could not detect any activity in the following 3 mutant proteins of the phage P4 primase, viz. C<sup>35</sup>→G, C<sup>57</sup>→G and E<sup>214</sup>→Q (numbering from phage 4 primase).<sup>[11]</sup> (Similar amino acids are highlighted in dark blue in *E. coli* primase sequence

in Fig. 4 and marked as SDM. Furthermore, mutagenesis experiments of phage P4 primase by Ziegelin et al. coupled with the observation that mutation of the invariant K<sup>241</sup> of *E. coli* DnaG permitted transcription initiation, but inhibited the primer suggests its interaction with the DNA at the initiation site (highlighted in dark blue).<sup>[11]</sup> X-ray crystallographic data by Keck et al. showed the invariant Glu<sup>265</sup>, Asp<sup>309</sup> could serve as the metal-binding residues and bind a Mg<sup>2+</sup> in the *E. coli* primase.<sup>[9]</sup> Further insights on the *E. coli* primase are provided by SDM analysis of by Godson et al.<sup>[11]</sup> They have reported that each of the conserved amino acids on the *E. coli* primase, viz. Arg<sup>146</sup>, Arg<sup>221</sup>, Tyr<sup>230</sup>, Gly<sup>266</sup>, and Asp<sup>311</sup> were involved in the process of catalysis. The invariant residues, Glu<sup>265</sup> and Asp<sup>309</sup>, were also critical because a substitution of each amino acid irreversibly destroyed the catalytic activity (highlighted in dark blue). Furthermore, the E<sup>265</sup>→Q, D<sup>309</sup>→A, and D<sup>311</sup>→A mutant proteins of the NTD of *E. coli* primase expressed the whole P66 primase proteins but were unable to synthesize the primer RNA. Two *E. coli* primase mutant proteins, E<sup>265</sup>→Q and D<sup>309</sup>→A also showed no pRNA synthesis activity, even in the presence of Mn<sup>2+</sup> that enhanced synthesis by wild-type primase and all the mutant primases. They found that the mutant proteins R<sup>146</sup>→A, R<sup>221</sup>→A, Y<sup>230</sup>→A, G<sup>266</sup>→A, and D<sup>311</sup>→A restored to some degree (but still not as high as that of wild-type primase) by using Mn<sup>2+</sup> ions instead of Mg<sup>2+</sup> in the pRNA synthesis reaction mixture. Further studies have shown that the catalytic core of the DnaG could accommodate up to three divalent metal ions that may vary from Mg<sup>2+</sup> to Mn<sup>2+</sup> and Fe<sup>2+</sup>.<sup>[11]</sup> The active sites of viral and eubacterial primases are shown in Table 1.

Table 1 Active sites of viral and eubacterial primases

Primase	ZBM/Mg <sup>2+</sup> -binding site	Method
Phage P4	-P C <sup>5</sup> PV CGGSDRFRFDDREGRGTWY C <sup>5'</sup> NQCG-	-SDM- No primase activity. <sup>[10]</sup>
Phage P4	-WIA E <sup>214</sup> GYATAL-	-SDM- No primase activity. <sup>[10]</sup>
<i>E. coli</i>	-AC C <sup>40</sup> PFHNEKTPSFTVNGEKQFYH C <sup>61</sup> FGCG-	-By Seq. similarity with P4 phage [This work]
<i>E. coli</i>	-LVV E <sup>26</sup> GYMDV-	-SDM - No primase activity. <sup>[11]</sup>
<i>E. coli</i>	-LVVE <sup>265</sup> GYMDV -	-X-ray - No primase activity. <sup>[9]</sup>
<i>E. coli</i>	-CYD <sup>107</sup> GDR-	-SDM - No primase activity. <sup>[11]</sup>
<i>E. coli</i>	-H K <sup>241</sup> GRQLY <sup>246</sup> GL-	-SDM - No primer Synthesis. <sup>[10]</sup>
<i>E. coli</i>	-GR <sup>214</sup> V I G F G G R <sup>221</sup> V L G N D T P K Y <sup>230</sup> -	-SDM- No primase activity. <sup>[11]</sup>
<i>E. coli</i>	-CYDGD <sup>111</sup> R-	-SDM- No primase activity. <sup>[11]</sup>

The highlighted amino acids are from SDM or X-ray crystallographic data. Seq., Sequence; ZBM, dark blue; metal-binding motif, Green.

Based on the sequence similarities, SDM and X-ray crystallographic data, the proposed primase active site and the priming reactions are shown in Fig. 5. After binding to the DNA, primase should bind to

two NTPs to generate the E-DNA-NTP-NTP quaternary complex, (the two NTPs used are purine nucleotides only).

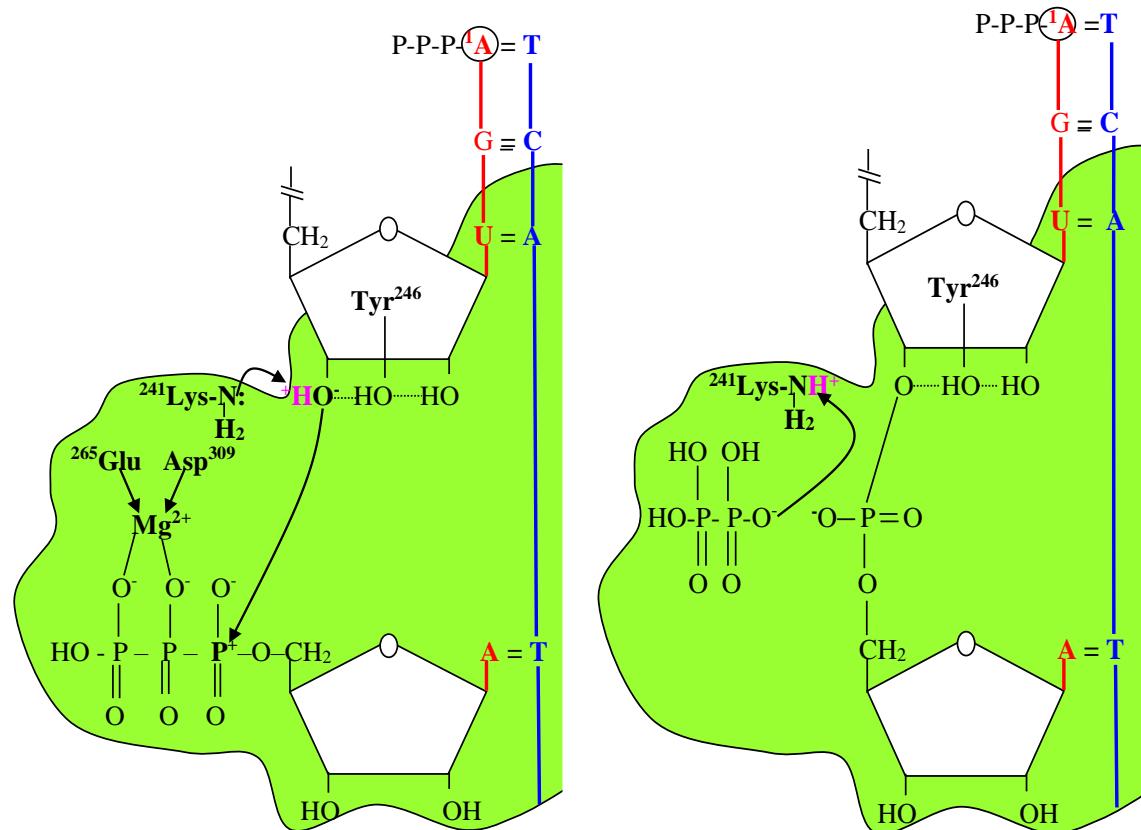


Fig. 5 The proposed active site structure and priming reactions of eubacterial primases (numbering from *E. coli* primase)

## Priming of Replication Initiation in Eukaryotes

In eukaryotes, the DNA replication involves in an incredibly more sophisticated, but highly coordinated series of molecular events. These events may be divided into six major stages: Initiation, DNA unwinding, Primer synthesis, Elongation, Termination and the final finishing-up reactions like, primer removal, gap-filling, DNA repair, ligation, mismatch repair, etc. Initiation of replication in eukaryotes is much more complex than in bacteria and at least half a dozen proteins are involved at the initiation site. During the first step of initiation, a set of initiator proteins bind to the origin of replication. This binding triggers unwinding of the DNA double helix into two single-stranded DNA strands by the replicative helicase. Because the group of proteins, the SSBs, keep the single strands

stable until primer synthesis which is followed by the elongation process. A third family of proteins, the topoisomerases, reduce some of the torsional strain caused by the unwinding of the double helix. As in eubacteria, in eukaryotes also the RNA primase activity is essential to initiate the replication. The primase is a DdRp (EC 2.7.7.102) that synthesizes short RNA primers (RNA:DNA type). Unlike prokaryotes, a second primer DNA:DNA type is synthesized by a DNA pol (DdDp), which is the one that is further recognized and extended by the replicases,  $\epsilon$  and  $\delta$  (leading- and lagging-strands, respectively) to complete the replication process. This is in stark contrast to the bacterial initiation, where the RNA primer itself is recognized and extended by the DNA pol III to complete the replication process.

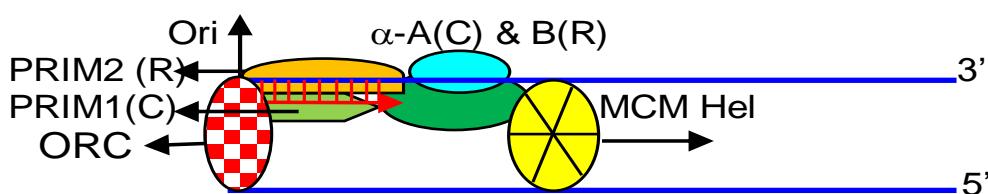


Fig. 6 A schematic diagramme showing assembly of various proteins at the initiator complex of eukaryotic replication origin (G1 phase)

The human primase subunits (PRIM1-48:PRIM2-58) and DNA pol subunits (A-180:B-70) make the heterotetramer. C, catalytic; R, Regulatory; Ori, origin of replication (an AT rich region).

The primer usually starts with an A at its 5'-end, followed by a G with a free 3'-OH, as in prokaryotes. ORC, Origin of replication recognition complex with Cdc6; MCM Hel, (minichromosome maintenance helicase is a hexameric protein and is also known as the unwinding enzyme). (The MCM helicase catalyzes the separation of the DNA strands. The energy required for this process is generated by the hydrolysis of ATP).

Thus, the eukaryotic primosome complexes are structurally different from the prokaryotic ones, even though they also follow a similar mechanism for priming DNA replication as in prokaryotes. For example, the primosome complex in eukaryotes is a heterotetramer made of two different enzyme activities and each having its own regulatory protein, e.g., the primase (a hetreodimer made of PRIM1 and PRIM2 subunits) and DNA pol  $\alpha$  (again a hetreodimer

made of A and B subunits) as shown in Fig. 6. Prim<sup>2</sup>-Pola<sup>2</sup> heterotetramer synthesizes short RNA primers (7-10 nts) on leading and lagging strands, which are further extended by the DNA pol  $\alpha$  for making short DNA primers. In other words, the eukaryotic replication is initiated and completed by a multienzyme complex involving 4 different polymerases, viz. primase, DNA pol  $\alpha$ , DNA pol  $\epsilon$  and DNA pol  $\delta$  (Fig. 7).

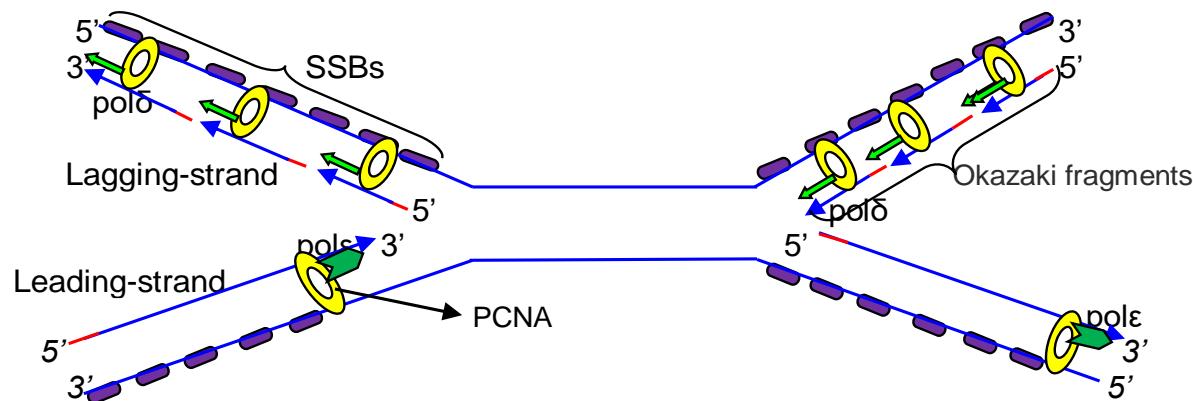


Fig. 7 A schematic diagramme showing eukaryotic replication process

*Blue lines, DNA strands; Red lines, RNA primers; Proliferating Cell Nuclear Antigen (PCNA) plays a key role in DNA replication and cell cycle regulation. It acts as a clamp to keep the polymerases on the DNA, enhancing the processivity. The DNA clamp loader, Replication Factor C (RFC), loads the PCNA to the DNA strands. [12] Both pol ε ➤ and pol δ ➤ function with the PCNA sliding clamp. [12]*

Separations of the DNA strands are initiated simultaneously in all chromosomes and at many different points along the genome at G1 phase. At the end of G1 phase, cell-cycle dependent kinase(s) triggers initiation of DNA replication at origins of replications. As a result, the double helices form “bubbles” at the separation zones. These bubbles or ‘replication units’ are called the *replicons*. In the nucleus of every human cell, ~50,000 “*replication bubbles*” are formed. This simultaneous unwinding of DNA in many different sites at the same time helps in completing the replication process much faster than if performed progressively from one end to another end of the very long double helix of each chromosome. The site, where both strands are separated is called the *replication fork*. In each separation area, two forks are formed, which progress in opposite directions. These Okazaki fragments generated during the lagging-strand synthesis are joined together by a group of processing enzymes like FEN1 (flap endonuclease-1), 5'-exonuclease, RNase H2 and DNA ligase I to create a single continuous strand. [13] The FEN-1 and the 5'-exonuclease act cooperatively with DNA pol δ to remove all the RNA/DNA primers and maintain a nick that can be closed by DNA ligase I. [14]

Figure 8 shows the MSA of the catalytic subunit of the RNA primases from various yeasts and higher fungi.

The NTD and CTD are highly conserved, but only at a few regions. Apart from the Cys residues making the ZBM, there are many invariant Cs in the eukaryotic primase structure, which may play a crucial role in maintaining the 3D structure of the enzyme. The D<sup>111</sup> and D<sup>113</sup> (numbering from *S. cerevisiae* primase) of the -DID- motif are proposed to bind a Mg<sup>2+</sup> to initiate the nucleophilic attack on the free 3'-hydroxyl end during synthesis of the RNA primer. The proposed template-binding pair, catalytic amino acid and the ZBM are found at the NTD, whereas they are also identified at the CTD of primases from animal sources (Fig. 9). The CTD is also not well conserved, except at the C-terminal ends. For example, the C-terminal ends are remarkably conserved with an invariant D/EF (with a few exceptions). Some of the common peptides in yeasts and human are highlighted in light yellow. Interestingly, the BLASTp analysis of human and *candida albicans* (a well-known fungal human pathogen) primases showed only 36.39% identity, suggesting again the highly pathogenic yeast primase could also serve as a good drug target.

## CLUSTAL O (1.2.4) MSA of primases from various yeasts and higher fungi

tr R9X8Z6 R9X8Z6_ASHAC	NHS	SPK	PQRDMVNRELAMAFRSGAYKRYNSFSSVQEKKQIQKANPDRFE	E	GAVYNKPPRD	120
tr Q75ER1 Q75ER1_ASHGO	NHS	SPK	PQRDMVNRELAMAFRSGAYKRYNSFSSVQEKKQIQKANPDRFE	E	GAVYNKPPRD	120
tr A0A1G4MEW9 A0A1G4MEW9_LACFM	NHS	SPK	PCRDMLTREFAFAMAFRSGAYKRYNSFSSVQEKKQIQQRGNPDRFE	E	GAIYNKPPRD	116
tr A0A0C7MKM8 A0A0C7MKM8_9SACH	NHS	SPK	PGRDMMHRELAMAFRSGAYKRYNSFDSVQEKKQIQRSNPDRFE	E	GAVYNKPPQ	94
tr W0TJ26 W0TJ26_KLUMD	NHS	SPK	GRDIIINREFAMAFRSGAYKRYNSFNSVQEKKQIQKGNDPRFE	E	GAVYNKPPRE	118
sp P10363 PRI1_YEAST	NHS	SPK	PSRDMINREFAMAFRSGAYKRYNSFNSVQDFKKQIEKANPDRFE	I	GAIYNKPPRE	92
tr G8ZZU9 G8ZZU9_TORDC	NHS	SPK	PKGRDIVNREFAMAFRSGAYKRYNSFNSVQEKKQIEKANPDRFE	I	GAIYNKAPRE	99
tr A0A212MCQ6 A0A212MCQ6_ZYGBA	NHS	SPK	PKLGRDIINREFAMAFRSGAYKRYNSFNSVQEKKQIEKANPDRFE	I	GAIYNKPPQD	101
tr A0A1Z8JH28 A0A1Z8JH28_PICKU	NHS	SPK	KVNNDFTHREFAFEHQNGAYQRYNSFQDAEEKKKVVKNPTFR	E	GAVYVPEPKD	96
tr A0A1B2JDX1 A0A1B2JDX1_PICKPA	NHS	SPK	VPQTDFTMREFAFFKSGAYQRYNSFVSAKEEKTIVVRTPTR	F	GAVYSINPSK	163
tr A0A1V2L196 A0A1V2L196_CYBFA	NHS	SPK	PKQPDFTMREFAFEHRSGAYQRYNSFESLENEKQSVRVRNP	F	GPVYAINPSK	140
tr A0A0H5BZM9 A0A0H5BZM9_CYBJN	NHS	SPK	PKQPDFTMREFAFEHRSGAYQRYNSFESPEDKSVVVRNP	F	GAVYAVNPSE	142
tr A0A510P6L1 A0A510P6L1_CANAR	NHS	SPK	PKTDFTMREFAFEYRSGAYQRYNSFGSPEEKKSVVTANP	F	GAVYAVNPKE	103
tr C5MJ98 C5MJ98_CANTT	NHS	SPK	PKTDFTMREFAFEYRNSFKTGTLEEKKSVVTANP	F	GAVYKINPKE	104
tr C4YS43 C4YS43_CANAW	NHS	SPK	PKTDFTMREFAFEYRNSFKTGTLEEKKSVVTANP	F	GAVYKINPKE	78
*	*	*	***:	:	***:	*
tr R9X8Z6 R9X8Z6_ASHAC	RDS	VLS	VLTTEMKPLEKELVFIDMDYD	DYRTCCSGAQVKWF	FIFLAMKVMTSLVEDF	180
tr Q75ER1 Q75ER1_ASHGO	RDS	VLS	VLTTEMKPLEKELVFIDMDYD	DYRTCCSGAQVKWF	FIFLAMKVMTSLIEDF	180
tr A0A1G4MEW9 A0A1G4MEW9_LACFM	RDS	VLS	VLTTEMKPCKEKEVVFIDMDYD	YDRTCCSGANVCSKCKWFISI	LAMKIIINIALSED	176
tr A0A0C7MKM8 A0A0C7MKM8_9SACH	MDS	LLK	ADMKPLEKEVVFIDMDYD	DAHRTCCSGANVCSKCKWFISI	LAMKIIINVALEED	154
tr W0TJ26 W0TJ26_KLUMD	RDS	ILK	TKEMKPIEKEVELFIDMDYD	SYRTCCSGAQVKWF	FITLAMKVINSLEDF	178
sp P10363 PRI1_YEAST	RDT	LLK	SELKAKELEKELVFIDMDYD	YDRAFTCCSGAQVKWF	ISLAMKITNTALRED	152
tr G8ZZU9 G8ZZU9_TORDC	RDT	LLK	SELKAKELEKELVFIDMDYD	YDRAFTCCSGAQVKWF	ITLAMLVVNTLLED	159
tr A0A212MCQ6 A0A212MCQ6_ZYGBA	RDS	LLK	SDMVPLEKELFIDMDYD	YDRTCCSGAQVKWF	ISLAMKIMETVLVDD	161
tr A0A1Z8JH28 A0A1Z8JH28_PICKU	RKT	VSN	AKNMMPKLSKEMVLDIDTDYD	EIRTCCKGTSIDKCKWFITLAI	KVVDDVAIREDF	156
tr A0A1B2JDX1 A0A1B2JDX1_PICKPA	RKT	VSN	AKNMMPKLSKEMVLDIDTDYD	EIRTCCKGTSIDKCKWFITLAI	KVVDDVAIREDF	223
tr A0A1V2L196 A0A1V2L196_CYBFA	RKS	VPS	KSAMKALEKELVFIDIDTDYD	EIRTCCKSKTIDKCKWFITTAAT	KVLDALRED	200
tr A0A0H5BZM9 A0A0H5BZM9_CYBJN	RKTL	PKS	AMKPLEKELVFIDIDTDYD	EIRTCCKSKTIDKCKWFITTAAT	KVLDALRED	202
tr A0A510P6L1 A0A510P6L1_CANAR	RRNL	LKSA	MKPLSKELVFIDIDTDYD	EIRTCCKSGTDIDKCKWF	KIASEVLSHALEDF	163
tr C5MJ98 C5MJ98_CANTT	RKLP	TKAF	PKESKELVFIDIDTDYD	EIRTCCKSGTDIDKCKWF	KIASEVLSHALEDF	164
tr C4YS43 C4YS43_CANAW	RKLP	QPK	FPKESKELVFIDIDTDYD	EIRTCCKSGTDIDKCKWF	KIQLQVGSKIIEAALRED	138
*	*	*	**:	***:	***:	***:
tr R9X8Z6 R9X8Z6_ASHAC	GLE	QFI	WI	FSGRRGAHCWISDPRARSLSDVQRKSVL	DYMNVVRDRNS--EKRISFTRPY	237
tr Q75ER1 Q75ER1_ASHGO	GLE	QFI	WI	FSGRRGAHCWISDPRARSLSDMQRKSVL	DYMNVVRDRNL--EKRISFTRPY	237
tr A0A1G4MEW9 A0A1G4MEW9_LACFM	GYE	DFI	WI	FSGRRGAHCWVSDKRARTLSDI	QRRNVLDMNVVRDRNV--GKRLNLKRPY	233
tr A0A0C7MKM8 A0A0C7MKM8_9SACH	GYT	DIW	FSGRRGAHCWVSDKRARTLSDI	QRRNVLDMNVVRDRNV--DKRQNLKRPY	211	
tr W0TJ26 W0TJ26_KLUMD	GFE	IFIW	FSGRRGAHCWISDKRARVMI	DTYRKNMVLDDVNVRDRNI--EKRLNLVRPL	235	
tr A0A7I9E8J8 A0A7I9E8J8 YEASX	GYK	DFIWI	FSGRRGAHCWVSDKDRARALTDV	QRRNVLDDVNVRDRNT--DKRLALKRPY	209	
sp P10363 PRI1_YEAST	GYKDF	IFIWI	FSGRRGAHCWVSDKDRARALTDV	QRRNVLDDVNVRDRNT--DKRLALKRPY	209	
tr C8ZAQ7 C8ZAQ7 YEAS8	GYKDF	IFIWI	FSGRRGAHCWVSDKDRARALTDV	QRRNVLDDVNVRDRNT--DKRLALKRPY	209	
tr A0A815ZF39 A0A815ZF39 YEASX	GYKDF	IFIWI	FSGRRGAHCWVSDKDRARALTDV	QRRNVLDDVNVRDRNT--DKRLALKRPY	209	
tr G8ZZU9 G8ZZU9_TORDC	AYE	DFIWI	FSGRRGAHCWVSDKDRARILHD	QRRNVLDDVNVRDRNA--DKRLALKRPY	216	
tr A0A212MCQ6 A0A212MCQ6_ZYGBA	GYEV	FFWV	FSGRRGAHCWVSDRARRT	LNDLQRRNVLDDVNVRDRSA--EKRIALKRPY	218	
tr A0A1Z8JH28 A0A1Z8JH28_PICKU	GIENR	IFIWI	FSGRRGAHCWMSDSKIRKENGRRF	IAEYLDLNVKGSK--KGIFGLKKPY	215	
tr A0A1B2JDX1 A0A1B2JDX1_PICKPA	GFEHM	IFIWI	FSGRRGAHCWVSDQRARS	MNDTLRKAVIDYLDVLNKGGHK--NRNLTFRRPL	282	
tr A0A1V2L196 A0A1V2L196_CYBFA	GFEHM	IFIWI	FSGRRGAHCWVSDQRARS	MNDTLRKAVIDYLDVLNKGGHK--NRNLTFRRPL	256	
tr A0A0H5BZM9 A0A0H5BZM9_CYBJN	GFEHM	IFIWI	FSGRRGAHCWVSDRARS	DLDEATRKSIVEYLDVLGGRSHKGMTSLSIKKP	223	
tr A0A0LONST6 A0A0LONST6_CANAR	GFDH	IFIWI	FSGRRGAHCWVSDRARS	DLDEATRKSIVEYLDVLGGRSHKGMTSLSIKKP	223	
tr A0A510P6L1 A0A510P6L1_CANAR	GFEHM	IFIWI	FSGRRGAHCWVSDRARS	DLDEATRKSIVEYLDVLGGRSHKGMTSLSIKKP	221	
tr C5MJ98 C5MJ98_CANTT	GFEHL	IFIWI	FSGRRGAHCWVSDRARELDESSR	RAIVEYLDVLGAKTQ--QGTLNLRPL	195	
tr C4YS43 C4YS43_CANAW	GFEHL	IFIWI	FSGRRGAHCWVSDKRARELDESR	KIAIDYLDVLSSKNQ--NGSLNIKKPF	195	
*	*	*	*****:	***:	***:	***:
tr R9X8Z6 R9X8Z6_ASHAC	VTK	QTI	HLLKAPCIHPATGNVC	PI	DFSPTS-VPKLIQLQHEME	390
tr Q75ER1 Q75ER1_ASHGO	VTK	QTI	HLLKAPCIHPATGNVC	PI	DFSPTG-VPKLIQLQHEME	390
tr A0A1G4MEW9 A0A1G4MEW9_LACFM	VTK	QTI	HLLKAPCIHPATGNVC	PI	GFTPSD-APRLIELQQOME	388
tr A0A0C7MKM8 A0A0C7MKM8_9SACH	VTK	QTI	HLLKAPCIHPATGNVC	PI	AFTPSD-APKLKILQNELE	367
tr W0TJ26 W0TJ26_KLUMD	VTK	QTI	HLLKAPCIHPATGNVC	PI	AFSPEM-APKLITLQNEME	391
tr A0A7I9E8J8 A0A7I9E8J8 YEASX	VTK	QTI	HLLKAPCIHPATGNVC	PI	SFAPEK-APKLIDLQTEME	362
sp P10363 PRI1_YEAST	VTK	QTI	HLLKAPCIHPATGNVC	PI	SFAPEK-APKLIDLQTEME	362
tr C8ZAQ7 C8ZAQ7 YEAS8	VTK	QTI	HLLKAPCIHPATGNVC	PI	SFAPEK-APKLIDLQTEME	362
tr A0A815ZF39 A0A815ZF39 YEASX	VTK	QTI	HLLKAPCIHPATGNVC	PI	SFAPEK-APKLIDLQTEME	362
tr G8ZZU9 G8ZZU9_TORDC	VTK	QTI	HLLKAPCIHPATGNVC	PI	SFAPEK-APKLIDLQTEME	362
tr A0A212MCQ6 A0A212MCQ6_ZYGBA	VTK	QTI	HLLKAPCIHPATGNVC	PI	KFNPDQ-APKLITLQHEME	369
tr A0A1Z8JH28 A0A1Z8JH28_PICKU	VTK	QTI	HLLKAPCIHPATGNVC	PI	KFTPAE-APKLDDLQNEME	371
tr A0A1B2JDX1 A0A1B2JDX1_PICKPA	VSRQM	HLLKSP	CVHPATGNVC	PI	REFNPFTDAPNLQQIFS--	361
tr A0A1V2L196 A0A1V2L196_CYBFA	VSKLQ	HLLKSP	CVHPATGNVC	PI	KEFWPDE-SPNLRLSQLSELE	428
tr A0A0H5BZM9 A0A0H5BZM9_CYBJN	VSKQV	HLLKSP	CVHPATGNVC	PI	WDFNPME-APNLHQIQKELI	406
tr A0A0LONST6 A0A0LONST6_CANAR	VSKQV	HLLKSP	CVHPATGNVC	PI	VSQRMHLLKSP	408
tr A0A510P6L1 A0A510P6L1_CANAR	VSKQV	HLLKSP	CVHPATGNVC	PI	KEFWPDE-SPNLRLSQLDEID	394
tr C5MJ98 C5MJ98_CANTT	VSKQV	HLLKSP	CVHPATGNVC	PI	VSKQVHLLKSP	394
tr C4YS43 C4YS43_CANAW	VSKQV	HLLKSP	CVHPATGNVC	PI	VSKQVHLLKSP	365
*	*	*	*****:	***:	***:	*

## //C-terminal end of primases from yeasts and higher fungi

tr R9X8Z6 R9X8Z6_ASHAC	NHPEGE----DIE	EF	436
tr Q75ER1 Q75ER1_ASHGO	NHPEGE----DIE	EF	436
tr A0A1G4MEW9 A0A1G4MEW9_LACFM	EREDET---EI---	DF	433
tr A0A0C7MKM8 A0A0C7MKM8_9SACH	EREEEA---VN---	DF	412
tr W0TJ26 W0TJ26_KLUMD	EREEDS---KQ-SI	DF	438
sp P10363 PRI1_YEAST	EREDDD---EPASL	DF	409
tr G8ZZU9 G8ZZU9_TORDC	ERAEDG---ET-SI	EF	416
tr A0A212MCQ6 A0A212MCQ6_ZYGBA	SREETD---AQ-SY	EF	418
tr A0A1Z8JH28 A0A1Z8JH28_PICKU	NDE-----I	DF	402
tr A0A1B2JDX1 A0A1B2JDX1_PICPA	NRDQSVVDQDDKSLI	EF	488
tr A0A1V2L196 A0A1V2L196_CYBFA	RDDNE----EDI	EF	462
tr A0A0H5BZM9 A0A0H5BZM9_CYBJN	RDEQE----PDI	EF	463
tr A0A510P6L1 A0A510P6L1_CANAR	REREP----DPI	DF	463
tr C5MJ98 C5MJ98_CANTT	RAREE----DPI	SF	455
tr C4YS43 C4YS43_CANAW	RSREE----DPI	SF	424

Fig. 8 MSA of primases from different yeasts and higher fungi

R9X8Z6\_ASHAC, *Ashbya aceri*  
 A0A1G4MEW9\_LACFM, *Lachancea fermentati*  
 W0TJ26\_KLUMD, *Kluyveromyces marxianus*  
 G8ZZU9\_TORDC, *Torulaspora delbrueckii*  
 A0A1Z8JH28\_PICKU, *Pichia kudriavzevii*  
 A0A1V2L196\_CYBFA, *Cyberlindnera fabianii*  
 A0A510P6L1\_CANAR, *Candida auris*  
 C4YS43\_CANAW, *Candida albicans*

Figure 9 shows the MSA of the catalytic subunit of the primases from animal sources. The N-terminal and C-terminal regions are highly conserved (with a few exceptions in some regions) unlike in yeasts and higher fungi. The primase catalytic subunit from various animal sources consists of a highly conserved template-binding pair (-YG-), the basic catalytic amino acid (R), a metal-binding catalytic triads (-DxD-) (highlighted in yellow and green, respectively), and the invariant -Cs, Ds and H of the ZBM (highlighted in orange). The 6 conserved Cs, a H and a D in that region could possibly make 2 ZBMs. Apart from the C residues that form the ZBMs, there are many invariant Cs in the eukaryotic primase structure, which may involve in disulphide

Q75ER1\_ASHGO, *Ashbya gossypii*  
 A0A0C7MKM8\_9SACH, *Lachancea lanzarotensis*  
 P10363/PRI1\_YEAST, *Saccharomyces cerevisiae*  
 A0A212MCQ6\_ZYGBA, *Zygosaccharomyces bailii*  
 A0A1B2JDX1\_PICPA, *Komagataella pastoris*  
 A0A0H5BZM9\_CYBJN, *Cyberlindnera jadinii*  
 C5MJ98\_CANTT, *Candida tropicalis*

bridges and contributing to the unique 3D structure of the protein. The D<sup>109</sup>, D<sup>111</sup> (numbering from human primase) of the -DxD- triad could bind a catalytic Mg<sup>2+</sup> ion. As in prokaryotes, the eukaryotic primases also initiate the primer only with purines, viz. an ATP or a GTP. [15] Therefore, the primase must have two NTP binding sites and are referred to as the initiation and elongation sites and the binding of one may help the binding of the other, initiating the nucleophilic attack on the 3'-hydroxyl end. Interestingly, both the animal and yeast primases' ends mostly in the diad D/EF (Figs. 8, 9). The SDM experiments on the mouse and human primases are highlighted in dark blue.

## CLUSTAL O (1.2.4) MSA of eukaryotic primases from animal sources

tr | S4RXE6 | S4RXE6\_PETMA  
tr | A0A670KIZ7 | A0A670KIZ7\_PODMU  
tr | A0A670ZCJ0 | A0A670ZCJ0\_PSETE  
tr | A0A3B3DX7 | A0A3B3DX7\_9TELE  
tr | A0A4W4EQ59 | A0A4W4EQ59\_ELEEL  
tr | A0A6P6J5A1 | A0A6P6J5A1\_CARAU  
tr | F1QKT6 | F1QKT6\_DANRE  
tr | A0A674ERWO | A0A674ERWO\_SALTR  
tr | A0A665TRG3 | A0A665TRG3\_ECHNA  
tr | A0A3Q2YDY7 | A0A3Q2YDY7\_HIPCM  
tr | A0A3P9DJD8 | A0A3P9DJD8\_9CICH  
tr | I3KK59 | I3KK59\_ORENI  
tr | A0A6P8TRY8 | A0A6P8TRY8\_GYMAC  
tr | A0A6J2Q218 | A0A6J2Q218\_COTGO  
tr | A0A7N8YL88 | A0A7N8YL88\_9TELE  
tr | A0A674MIZ2 | A0A674MIZ2\_TAKRU  
tr | Q6NVT9 | Q6NVT9\_XENTR  
tr | A0A3Q0D7A7 | A0A3Q0D7A7\_MESAU  
**tr | J3QN19 | J3QN19\_MOUSE**  
tr | O89045 | O89045\_RAT  
tr | A0A671ERV8 | A0A671ERV8\_RHIFE  
tr | A0A6P3RH99 | A0A6P3RH99\_PTEVA  
tr | A0A1S3FGZ5 | A0A1S3FGZ5\_DIPOR  
tr | G1SJS0 | G1SJS0\_RABIT  
tr | F6Y809 | F6Y809\_CALJA  
tr | F6SCS2 | F6SCS2\_MACMU  
tr | A0A2K6M7L7 | A0A2K6M7L7\_RHIBE  
tr | G3RCL3 | G3RCL3\_GORGO  
**sp | P49642 | PRI1\_HUMAN**  
tr | H2RQ08 | H2RQ08\_PANTR  
tr | A0A6D2VYJ6 | A0A6D2VYJ6\_PANTR

tr | S4RXE6 | S4RXE6\_PETMA  
tr | A0A670KIZ7 | A0A670KIZ7\_PODMU  
tr | A0A670ZCJ0 | A0A670ZCJ0\_PSETE  
tr | A0A3B3XDX7 | A0A3B3XDX7\_9TELE  
tr | A0A4W4EQ59 | A0A4W4EQ59\_ELEL  
tr | A0A6P6J5A1 | A0A6P6J5A1\_CARAU  
tr | F1QKT6 | F1QKT6\_DANRE  
tr | A0A674ERW0 | A0A674ERW0\_SALTR  
tr | A0A665TRG3 | A0A665TRG3\_ECHNA  
tr | A0A3Q2YDY7 | A0A3Q2YDY7\_HIPCM  
tr | A0A3P9DJD8 | A0A3P9DJD8\_9CICH  
tr | I3KK59 | I3KK59\_ORENI  
tr | A0A6P8TRY8 | A0A6P8TRY8\_GYMAC  
tr | A0A6J2Q218 | A0A6J2Q218\_COTGO  
tr | A0A7N8YL88 | A0A7N8YL88\_9TELE  
tr | A0A674MIZ2 | A0A674MIZ2\_TAKRU  
tr | Q6NVT9 | Q6NVT9\_XENTR  
tr | A0A3Q0D7A7 | A0A3Q0D7A7\_MESAU  
tr | J3QN19 | J3QN19\_MOUSE  
tr | O89045 | O89045\_RAT  
tr | A0A671ERV8 | A0A671ERV8\_RHIFE  
tr | A0A6P3RH99 | A0A6P3RH99\_PTEVA  
tr | A0A1S3FGZ5 | A0A1S3FGZ5\_DIPOR  
tr | G1JSJS0 | G1JSJS0\_RABIT  
tr | F6Y809 | F6Y809\_CALJA  
tr | F6SCS2 | F6SCS2\_MACMUS  
tr | A0A2K6M7L7 | A0A2K6M7L7\_RHIBE  
tr | G3RCL3 | G3RCL3\_GORGO  
sp | P49642 | PR1\_HUMAN  
tr | H2R0Q8 | H2R0Q8\_PANTR  
tr | A0A6D2VYJ6 | A0A6D2VYJ6\_PANTR

tr S4RXE6 S4RXE6_PETMA	DYDDVRSQSSAADI <b>C</b> KC <b>T</b> EN_WELMTIAI <b>I</b> VIDRALSED <b>F</b> GF <b>D</b> HRLWVYSGRGV <b>C</b> W <b>CD</b> <b>I</b> A	175
tr A0A670KIZ7 A0A670KIZ7_PODMU	DYDDVRSQSSA <b>E</b> IC <b>K</b> C <b>T</b> EN_WELMTIAI <b>I</b> R <b>I</b> D <b>R</b> AL <b>V</b> E <b>D</b> FG <b>I</b> KHRLWVYSGRGV <b>C</b> W <b>CD</b> <b>E</b> S	173
tr A0A670ZCJ0 A0A670ZCJ0_PSETE	DYDDVRSQSSA <b>E</b> IC <b>K</b> C <b>T</b> EN_WELMTIAI <b>I</b> R <b>I</b> D <b>R</b> AL <b>V</b> E <b>D</b> FG <b>I</b> QHRLWVYSGRGV <b>C</b> W <b>CD</b> <b>E</b> A	173
tr A0A3B3XDX7 A0A3B3XDX7_9TELE	DYDDVRSQSSA <b>E</b> IC <b>K</b> C <b>T</b> EN_WELMTIAI <b>I</b> R <b>I</b> D <b>R</b> AL <b>V</b> E <b>D</b> FG <b>I</b> QHRLWVYSGRGV <b>C</b> W <b>CD</b> <b>E</b> A	175
tr A0A4W4EQ59 A0A4W4EQ59_ELEEL	DYDDVRSQSSA <b>E</b> IC <b>K</b> C <b>T</b> EN_WELMTIAI <b>I</b> R <b>I</b> D <b>R</b> AL <b>R</b> ED <b>F</b> GF <b>Q</b> HRLWVYSGRGV <b>C</b> W <b>CD</b> <b>E</b> A	175
tr A0A6P6J5A1 A0A6P6J5A1_CARAU	DYDDVRSQSSA <b>E</b> IC <b>K</b> C <b>T</b> EN_WELMTIAI <b>I</b> R <b>I</b> D <b>R</b> AL <b>R</b> ED <b>F</b> GF <b>H</b> LLWVYSGRGV <b>C</b> W <b>CD</b> <b>D</b> A	175
tr F1QKT6 F1QKT6_DANRE	DYDDVRSQSSA <b>E</b> IC <b>K</b> C <b>T</b> EN_WELMTIAI <b>I</b> R <b>I</b> D <b>R</b> AL <b>R</b> ED <b>F</b> GF <b>H</b> LLWVYSGRGV <b>C</b> W <b>CD</b> <b>D</b> A	175
tr A0A674ERW0 A0A674ERW0_SALTR	DYDDVRSQSSA <b>E</b> IC <b>K</b> C <b>T</b> EN_WELMTIAI <b>I</b> R <b>I</b> D <b>R</b> AL <b>R</b> DD <b>F</b> GF <b>R</b> HLWVYSGRGV <b>C</b> W <b>CD</b> <b>D</b> A	175
tr A0A665TRG3 A0A665TRG3_ECHNA	DYDDVRSQSSA <b>E</b> IC <b>K</b> C <b>T</b> EN_WELMTIAI <b>I</b> R <b>I</b> D <b>R</b> AL <b>R</b> DD <b>F</b> GF <b>Q</b> HLLWVYSGRGV <b>C</b> W <b>CD</b> <b>E</b> A	176
tr A0A3Q2YDY7 A0A3Q2YDY7_HIPCM	DYDDVRSQSSA <b>E</b> IC <b>K</b> C <b>T</b> EN_WELMTIAI <b>I</b> R <b>I</b> D <b>R</b> AL <b>R</b> DD <b>F</b> GF <b>Q</b> YLLWVYSGRGV <b>C</b> W <b>CD</b> <b>E</b> A	180
tr A0A3P9JD8 A0A3P9JD8_9CICH	DYDDVRSQSSA <b>E</b> IC <b>K</b> C <b>T</b> EN_WELMTIAI <b>I</b> R <b>I</b> D <b>R</b> AL <b>R</b> Q <b>D</b> DF <b>G</b> QHRLWVYSGRGV <b>C</b> W <b>CD</b> <b>E</b> A	175
tr I3KK59 I3KK59_ORENI	DYDDVRSQSSA <b>E</b> IC <b>K</b> C <b>T</b> EN_WELMTIAI <b>I</b> R <b>I</b> D <b>R</b> AL <b>R</b> DD <b>F</b> GF <b>H</b> QLWVYSGRGV <b>C</b> W <b>CD</b> <b>E</b> A	175
tr A0A6P8TRY8 A0A6P8TRY8_GYMAC	DYDDVRSQSSA <b>E</b> IC <b>K</b> C <b>T</b> EN_WELMTIAI <b>I</b> R <b>I</b> D <b>R</b> AL <b>R</b> DD <b>F</b> GF <b>H</b> QLLWVYSGRGV <b>C</b> W <b>CD</b> <b>E</b> A	175
tr A0A6J2Q218 A0A6J2Q218_COTGO	DYDDVRSQSSA <b>E</b> IC <b>K</b> C <b>T</b> EN_WELMTIAI <b>I</b> R <b>I</b> D <b>R</b> AL <b>R</b> DD <b>F</b> GF <b>H</b> QLLWVYSGRGV <b>C</b> W <b>CD</b> <b>E</b> A	175
tr A0A7N8YL88 A0A7N8YL88_9TELE	DYDDVRSQSSA <b>E</b> IC <b>K</b> C <b>T</b> EN_WELMTIAI <b>I</b> R <b>I</b> D <b>R</b> AL <b>R</b> DD <b>F</b> GF <b>H</b> QLLWVYSGRGV <b>C</b> W <b>CD</b> <b>E</b> A	175
tr A0A674MIZ2 A0A674MIZ2_TAKRU	DYDDVRSQSSA <b>E</b> IC <b>K</b> C <b>T</b> EN_WELMTIAI <b>I</b> R <b>I</b> D <b>R</b> AL <b>R</b> DD <b>F</b> GF <b>H</b> QLLWVYSGRGV <b>C</b> W <b>CD</b> <b>D</b> S	175
tr Q6NVT9 Q6NVT9_XENTR	DYDDVRSQSSA <b>E</b> IC <b>K</b> C <b>T</b> EN_WELMTIA <b>I</b> R <b>I</b> D <b>R</b> ALA <b>E</b> DF <b>G</b> QHRLWVYSGRGV <b>C</b> W <b>CD</b> <b>D</b> S	175
tr A0A3Q0D7A7 A0A3Q0D7A7_MESAU	DYDDVRSQSSA <b>E</b> IC <b>K</b> C <b>T</b> EN_WELMTMA <b>I</b> R <b>I</b> D <b>R</b> ALK <b>E</b> DF <b>G</b> KHRLWVYSGRGV <b>C</b> W <b>CD</b> <b>D</b> S	173
tr J3QN19 J3QN19_MOUSE	DYDDVRC <b>C</b> SSA <b>D</b> IC <b>C</b> K <b>C</b> WT <b>L</b> MT <b>M</b> AM <b>I</b> R <b>I</b> D <b>R</b> ALK <b>E</b> DF <b>G</b> KHRLWVYSGRGV <b>C</b> W <b>CD</b> <b>M</b>	173 SDM
tr O89045 O89045_RAT	DYDDVRC <b>C</b> SSA <b>D</b> IC <b>C</b> K <b>C</b> WT <b>L</b> MT <b>M</b> AM <b>I</b> R <b>I</b> D <b>R</b> ALK <b>E</b> DF <b>G</b> KHRLWVYSGRGV <b>C</b> W <b>CD</b> <b>E</b> S	175
tr A0A671ERV8 A0A671ERV8_RHIFE	DYDDVRC <b>C</b> SSA <b>D</b> IC <b>C</b> K <b>C</b> WT <b>L</b> MT <b>M</b> AI <b>I</b> R <b>I</b> D <b>R</b> ALK <b>E</b> DF <b>G</b> KHRLWVYSGRGV <b>C</b> W <b>CD</b> <b>A</b> S	173
tr A0A6P3RH99 A0A6P3RH99_PTEVA	DYDDVRC <b>C</b> SSA <b>D</b> IC <b>C</b> K <b>C</b> WT <b>L</b> MT <b>M</b> AI <b>I</b> R <b>I</b> D <b>R</b> ALK <b>E</b> DF <b>G</b> KHRLWVYSGRGV <b>C</b> W <b>CD</b> <b>E</b> A	173
tr A0A1S3FGZ5 A0A1S3FGZ5_DIPOR	DYDDVRC <b>C</b> SSA <b>D</b> IC <b>C</b> K <b>C</b> WT <b>L</b> MT <b>M</b> AI <b>I</b> R <b>I</b> D <b>R</b> ALK <b>E</b> DF <b>G</b> KHRLWVYSGRGV <b>C</b> W <b>CD</b> <b>E</b> A	173
tr GLSJS0 G1SJS0_RABIT	DYDDVRC <b>C</b> SSA <b>D</b> IC <b>C</b> K <b>C</b> WT <b>L</b> MT <b>M</b> AI <b>I</b> R <b>I</b> D <b>R</b> ALK <b>E</b> DF <b>G</b> KHRLWVYSGRGV <b>C</b> W <b>CD</b> <b>E</b> S	173
tr F6Y809 F6Y809_CALJA	DYDDVRC <b>C</b> SSA <b>D</b> IC <b>C</b> K <b>C</b> WT <b>L</b> MT <b>M</b> AI <b>I</b> R <b>I</b> D <b>R</b> ALK <b>E</b> DF <b>G</b> KHRLWVYSGRGV <b>C</b> W <b>CD</b> <b>E</b> S	173
tr F6SCS2 F6SCS2_MACMU	DYDDVRC <b>C</b> SSA <b>D</b> IC <b>C</b> K <b>C</b> WT <b>L</b> MT <b>M</b> AI <b>I</b> R <b>I</b> D <b>R</b> ALK <b>E</b> DF <b>G</b> KHRLWVYSGRGV <b>C</b> W <b>CD</b> <b>E</b> S	173
tr A0A2K6M7L7 A0A2K6M7L7_RHIBE	DYDDVRC <b>C</b> SSA <b>D</b> IC <b>C</b> K <b>C</b> WT <b>L</b> MT <b>M</b> AI <b>I</b> R <b>I</b> D <b>R</b> ALK <b>E</b> DF <b>G</b> KHRLWVYSGRGV <b>C</b> W <b>CD</b> <b>E</b> S	173
tr G3RCL3 G3RCL3_GORGO	DYDDVRC <b>C</b> SSA <b>D</b> IC <b>C</b> K <b>C</b> WT <b>L</b> MT <b>M</b> AI <b>I</b> R <b>I</b> D <b>R</b> ALK <b>E</b> DF <b>G</b> KHRLWVYSGRGV <b>C</b> W <b>CD</b> <b>E</b> S	173
sp P49642 PRI1_HUMAN	DYDDVRC <b>C</b> SSA <b>D</b> IC <b>C</b> K <b>C</b> WT <b>L</b> MT <b>M</b> AI <b>I</b> R <b>I</b> D <b>R</b> ALK <b>E</b> DF <b>G</b> KHRLWVYSGRGV <b>C</b> W <b>CD</b> <b>M</b>	173 SDM
tr H2R0Q8 H2R0Q8_PANTR	DYDDVRC <b>C</b> SSA <b>D</b> IC <b>C</b> K <b>C</b> WT <b>L</b> MT <b>M</b> AI <b>I</b> R <b>I</b> D <b>R</b> ALK <b>E</b> DF <b>G</b> KHRLWVYSGRGV <b>C</b> W <b>CD</b> <b>E</b> S	173
tr A0A6D2VYJ6 A0A6D2VYJ6_PANTR	DYDDVRC <b>C</b> SSA <b>D</b> IC <b>C</b> K <b>C</b> WT <b>L</b> MT <b>M</b> AI <b>I</b> R <b>I</b> D <b>R</b> ALK <b>E</b> DF <b>G</b> KHRLWVYSGRGV <b>C</b> W <b>CD</b> <b>E</b> S	173
tr S4RXE6 S4RXE6_PETMA	---	342
tr A0A670KIZ7 A0A670KIZ7_PODMU	KKHDSKMA <b>E</b> TE <b>I</b> MLQYC <b>P</b> RLDVNVSKGINHLLKSPFSIHPKTGRVC <b>P</b> IDVARVDEF	330
tr A0A670ZCJ0 A0A670ZCJ0_PSETE	AVFRLNC <b>E</b> KA <b>E</b> RE <b>I</b> MLQYC <b>P</b> RLDMNVSKGINHLLKSPFSVHPKTGRIS <b>P</b> IDLKKL <b>D</b> F	328
tr A0A3B3XDX7 A0A3B3XDX7_9TELE	-TTAKWQYFE <b>E</b> IMLQYC <b>P</b> RLDVNVSKGVNHLLKSPFSVHPKTGRIS <b>P</b> IDLKKL <b>D</b> DF	333
tr A0A4W4EQ59 A0A4W4EQ59_ELEEL	-MSKKGGY <b>E</b> FE <b>E</b> IMLQYC <b>P</b> RLDVNVSKGVNHLLKSPFSIHPKTGRIS <b>P</b> IDVRND <b>K</b> F	342
tr A0A6P6J5A1 A0A6P6J5A1_CARAU	-NSKKGG <b>E</b> YFD <b>E</b> IMLQYC <b>P</b> RLDVNVSKGVNHLLKSPFSVHPKTGRIS <b>P</b> IDLKE <b>N</b> TF	342
tr F1QKT6 F1QKT6_DANRE	-SSKKGG <b>E</b> YYDK <b>E</b> IMLQYC <b>P</b> RLDVNVSKGVNHLLKSPFSVHPKTGRIS <b>P</b> IDLKELDT <b>F</b>	342
tr A0A674ERW0 A0A674ERW0_SALTR	-MTAKKG <b>H</b> Y <b>E</b> FE <b>E</b> IMLQYC <b>P</b> RLDVNVSKGVNHLLKSPFSVHPKTGRIS <b>P</b> IDLRELD <b>R</b> F	343
tr A0A665TRG3 A0A665TRG3_ECHNA	-GDDKKQQY <b>E</b> FE <b>E</b> IMLQYC <b>P</b> RLDVNVSKGVNHLLKSPFSVHPKTGRIS <b>P</b> IDLKELDK <b>F</b>	341
tr A0A3Q2YDY7 A0A3Q2YDY7_HIPCM	-ATAKKQ <b>H</b> Y <b>E</b> FE <b>E</b> IMLQYC <b>P</b> RLDVNVSKGVNHLLKSPFSVHPKTGRIS <b>P</b> MDLT <b>E</b> LD <b>T</b> F	346
tr A0A3P9JD8 A0A3P9JD8_9CICH	-STAKKG <b>H</b> Y <b>E</b> FE <b>E</b> IMLQYC <b>P</b> RLDVNVSKGVNHLLKSPFSVHPKTGRIS <b>P</b> VID <b>E</b> LD <b>K</b> F	341
tr I3KK59 I3KK59_ORENI	-GAKKG <b>H</b> Q <b>E</b> FE <b>E</b> IMLQYC <b>P</b> RLDVNVSKGVNHLLKSPFSVHPKTGRIS <b>P</b> VID <b>E</b> LD <b>K</b> F	342
tr A0A6P8TRY8 A0A6P8TRY8_GYMAC	-GTSKK <b>H</b> Q <b>E</b> FE <b>E</b> IMLQYC <b>P</b> RLDVNVSKGVNHLLKSPFSVHPKTGRIS <b>P</b> VIDLKE <b>E</b> TF	341
tr A0A6J2Q218 A0A6J2Q218_COTGO	-ATGKK <b>H</b> Q <b>E</b> FE <b>E</b> IMLQYC <b>P</b> RLDVNVSKGVNHLLKSPFSVHPKTGRIS <b>P</b> MDIN <b>R</b> LET <b>F</b>	328
tr A0A7N8YL88 A0A7N8YL88_9TELE	-GAAKKG <b>H</b> Q <b>E</b> FE <b>E</b> IMLQYC <b>P</b> RLDVNVSKGVNHLLKSPFSVHPKTGRIS <b>P</b> VIDLEEL <b>D</b> F	342
tr A0A674MIZ2 A0A674MIZ2_TAKRU	-LRRHSNIP <b>E</b> IMLQYC <b>P</b> RLDVNVSKGVNHLLKSPFSVHPKTGRIS <b>P</b> VIDT <b>K</b> L <b>D</b> Q <b>F</b>	341
tr Q6NVT9 Q6NVT9_XENTR	MKNDKCG <b>P</b> W <b>E</b> IMVQYC <b>P</b> RLDVNVSKGVNHLLKSPFSVHPKTGRIS <b>P</b> VIDF <b>Q</b> K <b>V</b> D <b>Q</b> F	341
tr A0A3Q0D7A7 A0A3Q0D7A7_MESAU	MKNDKCG <b>P</b> W <b>E</b> IMVQYC <b>P</b> RLDVNVSKGVNHLLKSPFSVHPKTGRIS <b>P</b> VIDFHK <b>V</b> D <b>Q</b> F	342
tr J3QN19 J3QN19_MOUSE	TKNDKCG <b>P</b> W <b>E</b> IMLQYC <b>P</b> RLDVNVSKGVNHLLKSPFSVHPKTGRIS <b>P</b> VIDFQKV <b>D</b> Q <b>F</b>	343
tr O89045 O89045_RAT	IKNNKCG <b>P</b> W <b>E</b> IMLQYC <b>P</b> RLDVNVSKGVNHLLKSPFSVHPKTGRIS <b>P</b> VIDLQKV <b>D</b> Q <b>F</b>	341
tr A0A671ERV8 A0A671ERV8_RHIFE	IKSDKYG <b>P</b> W <b>E</b> IMLQYC <b>P</b> RLDVNVSKGVNHLLKSPFSVHPKTGRIS <b>P</b> VIDLQKV <b>D</b> Q <b>F</b>	341
tr A0A6P3RH99 A0A6P3RH99_PTEVA	---SNCLGLW <b>E</b> IMLQYC <b>P</b> RLDVNVSKGVNHLLKSPFSVHPKTGRIS <b>P</b> VIDLQNV <b>D</b> Q <b>F</b>	333
tr A0A1S3FGZ5 A0A1S3FGZ5_DIPOR	TKNDKCP <b>P</b> W <b>E</b> IMLQYC <b>P</b> RLDVNVSKGVNHLLKSPFSVHPKTGRIS <b>P</b> VIDLQKV <b>D</b> Q <b>F</b>	342
tr G1SJS0 G1SJS0_RABIT	IKNDKCG <b>P</b> W <b>E</b> IMLQYC <b>P</b> RLDVNVSKGVNHLLKSPFSVHPKTGRIS <b>P</b> VIDLQKV <b>D</b> Q <b>F</b>	342
tr F6Y809 F6Y809_CALJA	IKNDKYG <b>P</b> W <b>E</b> IMLQYC <b>P</b> RLDVNVSKGVNHLLKSPFSVHPKTGRIS <b>P</b> VIDLQKV <b>D</b> Q <b>F</b>	338
tr F6SCS2 F6SCS2_MACMU	IKNDKYG <b>P</b> W <b>E</b> IMLQYC <b>P</b> RLDVNVSKGVNHLLKSPFSVHPKTGRIS <b>P</b> VIDLQKV <b>D</b> Q <b>F</b>	342
tr A0A2K6M7L7 A0A2K6M7L7_RHIBE	IKNDKYG <b>P</b> W <b>E</b> IMLQYC <b>P</b> RLDVNVSKGVNHLLKSPFSVHPKTGRIS <b>P</b> VIDLQKV <b>D</b> Q <b>F</b>	342
tr G3RCL3 G3RCL3_GORGO	IKNDKYG <b>P</b> W <b>E</b> IMLQYC <b>P</b> RLDVNVSKGVNHLLKSPFSVHPKTGRIS <b>P</b> VIDLQKV <b>D</b> Q <b>F</b>	342
sp P49642 PRI1_HUMAN	283IKNDKYG <b>P</b> W <b>E</b> IMLQYC <b>P</b> RLDVNVSKGVNHLLKSPFSVHPKTGRIS <b>P</b> VIDLQKV <b>D</b> Q <b>F</b>	342 SDM
tr H2R0Q8 H2R0Q8_PANTR	IKNDKYG <b>P</b> W <b>E</b> IMLQYC <b>P</b> RLDVNVSKGVNHLLKSPFSVHPKTGRIS <b>P</b> VIDLQKV <b>D</b> Q <b>F</b>	342
tr A0A6D2VYJ6 A0A6D2VYJ6_PANTR	IKNDKYG <b>P</b> W <b>E</b> IMLQYC <b>P</b> RLDVNVSKGVNHLLKSPFSVHPKTGRIS <b>P</b> VIDLQKV <b>D</b> Q <b>F</b>	342

## //C-terminal end of primases from animal sources

tr S4RXE6 S4RXE6_PETMA	RVFERFVDKLEASWKGEELLKKSDLKK-----EF-----	430
tr A0A670KIZ7 A0A670KIZ7_PODMU	KVFEQFLEEMEKSHKKERLKNTLALQKI-----CTAALLHPSL-----	418
tr A0A670ZCJ0 A0A670ZCJ0_PSETE	KVFEQFLEEMEKSYKGQLRKNSGKGLLPKDANSNFPSPNCFSDIS-----VMVTWAVT--	429
tr A0A3B3DX7_9TELE	KYFDQFLAGMADSWKGKLRQSGEFTNT-GDIFTNFMLDLFLC-----FLIFFL	437
tr A0A4W4EQ59_ELEEL	KLFQFLDAMAHTRKGETLKKSDLQK-----EF-----	422
tr A0A6P6J5A1 A0A6P6J5A1_CARAU	KVFDRFLREAMARSRKGDMLKKSDLQK-----DF-----	425
tr F1QKT6 F1QKT6_DANRE	KVLDRFLDSMARSRKGEMLKKSDLQK-----DF-----	425
tr A0A674ERW0 A0A674ERW0_SALTR	KLFQFLLEGMAHRSRKGEMLKKSDLQK-----DF-----	425
tr A0A665TRG3 A0A665TRG3_ECHNA	KYFDQFLDGIACTWKGEELLKKSDLQK-----QF-----	424
tr A0A3Q2YDY7 A0A3Q2YDY7_HIPCM	KYFDAFLDGVSQSWKGGEELLRKSDLQK-----EF-----	429
tr A0A3P9DJD8 A0A3P9DJD8_9CICH	KYFDQFLDGMARSWKGERLRNSDLQK-----EF-----	426
tr I3KK59 I3KK59_ORENI	KYFDQFLDGMARSWKGERLRNSDLQK-----EF-----	426
tr A0A6P8TRY8 A0A6P8TRY8_GYMAC	KCFQFLDVMRSRWSKGGEELLKKSDLQK-----DF-----	426
tr A0A6J2Q218 A0A6J2Q218_COTGO	KYFDQFLDGMARSWKGEELLKKSDLQK-----DF-----	433
tr A0A7N8YL88 A0A7N8YL88_9TELE	KYFDQFLDGMASSWKGGEELLRKSDLQK-----EF-----	406
tr A0A674MIZ2 A0A674MIZ2_TAKRU	KHFQFLAQMAQSWSKGDDLLKKSDLQK-----QF-----	425
tr Q6NVT9 Q6NVT9_XENTR	KVFQFLDKLDQSRKGSELLSKSDLKK-----EF-----	420
tr A0A3Q0D7A7 A0A3Q0D7A7_MESAU	KVFQFLENLDKSRKGSELLKKSDLQK-----DF-----	419
tr J3QN19 J3QN19_MOUSE	KVFQFLENLDKSRKGSELLKKSDLQK-----DF-----	418
tr O89045 O89045_RAT	KVFQFLLENLDKSRKGSELLKKSDLQ-----DF-----	415
tr A0A671ERV8 A0A671ERV8_RHIFE	KVFQFLLENLDKSRKGSELLKKSDLQK-----DF-----	419
tr A0A6P3RH99 A0A6P3RH99_PTEVA	KVFQFLLENLDKSRKGSELLRKSDLQK-----DF-----	419
tr A0A1S3FGZ5 A0A1S3FGZ5_DIPOR	KVFQFLLENLDKSRKGSELLKKSDLQR-----DF-----	411
tr G1SJS0 G1SJS0_RABIT	KVFQFLLENLDKSRKGSELLKKSDLQK-----DF-----	420
tr F6Y809 F6Y809_CALJA	KVFQFLLENLDKSRKGSELLKKSDLQK-----DF-----	420
tr F6SCS2 F6SCS2_MACMU	KVFEHFLLENLDKSRKGSELLKKSDLQK-----NLDSRKGEELLKKSDLQKDF-----	435
tr AOA2K6M7L7 A0A2K6M7L7_RHIBE	KVFEHFLLENLDKSRKGSELLKKSDLQK-----DF-----	420
tr G3RCL3 G3RCL3_GORGO	KVFEHFLLENLDKSRKGSELLKKSDLQK-----DF-----	420
sp P49642 PRI1_HUMAN	KVFEHFLLENLDKSRKGSELLKKSDLQK-----DF-----	420
tr H2R0Q8 H2R0Q8_PANTR	KVFEHFLLENLDKSRKGSELLKKSDLQK-----DF-----	423
tr A0A6D2VYJ6 A0A6D2VYJ6_PANTR	KVFEHFLLENLDKSRKGSELLKKSDLQK-----DF-----	420
:	: : *: : : *	:::

Fig. 9 MSA analysis of primases from animal sources

S4RXE6\_PETMA, *Petromyzon marinus*  
 A0A670ZCJ0\_PSETE, *Pseudonaja textilis*  
 A0A4W4EQ59\_ELEEL, *Electrophorus electricus*  
 F1QKT6\_DANRE, *Danio rerio*  
 A0A665TRG3\_ECHNA, *Echeneis naucrates*  
 A0A3P9DJD8\_9CICH, *Maylandia zebra*  
 A0A6P8TRY8\_GYMAC, *Gymnodraco acuticeps*  
 A0A7N8YL88\_9TELE, *Mastacembelus armatus*  
 Q6NVT9\_XENTR, *Xenopus tropicalis*  
 J3QN19\_MOUSE, *Mus musculus*  
 A0A671ERV8\_RHIFE, *Rhinolophus ferrumequinum*  
 A0A1S3FGZ5\_DIPOR, *Dipodomys ordii*  
 F6Y809\_CALJA, *Callithrix jacchus*  
 A0A2K6M7L7\_RHIBE, *Rhinopithecus bieti*  
 P49642/PRI1\_HUMAN, *Homo sapiens*  
 A0A6D2VYJ6\_PANTR, *Pan troglodytes*

The proposed active site amino acids in bacterial and eukaryotic primases are shown in Table 2. The prokaryotic primases differ in having a

A0A670KIZ7\_PODMU, *Podarcis muralis*  
 A0A3B3DX7\_9TELE, *Poecilia Mexicana*  
 A0A6P6J5A1\_CARAU, *Carassius auratus*  
 A0A674ERW0\_SALTR, *Salmo trutta*  
 A0A3Q2YDY7\_HIPCM, *Hippocampus comes*  
 I3KK59\_ORENI, *Oreochromis niloticus*  
 A0A6J2Q218\_COTGO, *Cottoperca gobio*  
 A0A674MIZ2\_TAKRU, *Takifugu rubripes*  
 A0A3Q0D7A7\_MESAU, *Mesocricetus auratus*  
 O89045\_RAT, *Rattus norvegicus*  
 A0A6P3RH99\_PTEVA, *Pteropus vampyrus*  
 G1SJS0\_RABIT, *Oryctolagus cuniculus*  
 F6SCS2\_MACMU, *Macaca mulatta*  
 G3RCL3\_GORGO, *Gorilla gorilla*  
 H2R0Q8\_PANTR, *Pan troglodytes*

pentapeptide repeat (shown by arrows) and a metal-binding site between the repeats and the proposed catalytic pairs. The yeast primases differ in the template-binding pair.

Table 2. Proposed active and MBSSs of the primases from prokaryotic and eukaryotic organisms

Primases (DdRps)	ZBM	Catalytic region	MBSs
Prokaryotic -A <b>CC</b> PFHNEKTPSFTVN <b>E</b> KQFY <b>H</b> C <b>G</b> <b>G</b> <sup>54</sup> - <b>G</b> R <b>V</b> <b>G</b> <b>G</b> <b>R</b> <b>V</b> <b>L</b> <b>G</b> N <b>D</b> T <b>P</b> <b>K</b> <b>Y</b> <b>L</b> <b>N</b> <b>S</b> <b>P</b> <b>E</b> <b>T</b> <b>D</b> <b>I</b> <b>F</b> <b>H</b> <b>K</b> <b>G</b> <b>R</b> <b>Q</b> <b>L</b> <b>Y</b> <b>G</b> <b>LYE</b> <sup>250</sup> - <b>Y</b> <b>D</b> <b>G</b> <b>D</b> <b>R</b> <sup>312</sup> - <b>G</b> <b>E</b> <b>D</b> <b>P</b> <b>D</b> <b>T</b> <sup>348</sup> . (Bacteria, <i>E. coli</i> , DnaG)	CX2C		
Eukaryotic - <b>R</b> <b>C</b> <b>O</b> <b>S</b> <b>S</b> <b>A</b> <b>C</b> <b>P</b> <b>K</b> <b>C</b> <b>W</b> <b>T</b> <b>L</b> <b>M</b> <b>T</b> <b>M</b> <b>A</b> <b>I</b> <b>/</b> <b>L</b> <b>W</b> <b>V</b> <b>S</b> <b>G</b> <b>R</b> <b>R</b> <b>G</b> <b>V</b> <b>H</b> <b>C</b> <b>W</b> <b>C</b> <sup>171</sup> - <b>14</b> <b>K</b> <b>L</b> <b>Y</b> <b>Y</b> <b>R</b> <b>R</b> <b>I</b> <b>L</b> <b>F</b> <b>P</b> <b>Y</b> <b>S</b> <b>Q</b> <b>Y</b> <b>R</b> <b>W</b> <b>L</b> <b>N</b> <b>Y</b> <b>B</b> <b>G</b> <sup>34</sup> - <b>L</b> <b>V</b> <b>F</b> <b>D</b> <b>I</b> <b>D</b> <b>M</b> <b>T</b> <b>D</b> <b>Y</b> <b>D</b> <sup>117</sup> . (Animal, <i>H. s.</i> , PRIM1) CX2C	CX2C		
Eukaryotic (Yeast, <i>S. c.</i> , PRIM1)	- <b>R</b> <b>T</b> <b>C</b> <b>C</b> <b>S</b> <b>G</b> <b>A</b> <b>Q</b> <b>V</b> <b>C</b> <b>S</b> <b>K</b> <b>C</b> <b>W</b> <b>K</b> <b>F</b> <sup>137</sup> .		- <b>63</b> <b>N</b> <b>S</b> <b>V</b> <b>Q</b> <b>D</b> <b>F</b> <b>K</b> <b>1A</b> <b>Q</b> <b>I</b> <b>E</b> <b>K</b> <b>A</b> <b>N</b> <b>P</b> <b>D</b> <b>R</b> <b>F</b> <b>E</b> <b>I</b> <b>13</b> <b>G</b> <b>A</b> <b>I</b> <b>Y</b> <sup>86</sup> - <b>E</b> <b>L</b> <b>V</b> <b>F</b> <b>D</b> <b>I</b> <b>D</b> <b>M</b> <b>D</b> <b>D</b> <b>Y</b> <b>D</b> <sup>119</sup> .

ZBM, Zinc-binding motif; MBSs, Metal-binding sites; *H. s.*, *Homo sapiens*; *S. c.*, *Saccharomyces cerevisiae*; → direct repeats.

### Active Site Analyses of Eukaryotic Primases

Through a combination of genetic, molecular biological, biochemical and X-ray crystallographic analyses, significant advances have been made in understanding the essential roles played by the eukaryotic primase during replication. The human primase consists of a small catalytic subunit (PRIM1-p48) and a large regulatory subunit (PRIM2-p58). Unlike in prokaryotes, the primase is associated to an additional enzyme, DNA pol  $\alpha$ , in the primosome complex. The DNA pol  $\alpha$  is composed of a larger catalytic subunit A (p180) and a smaller regulatory subunit B (p70). The smaller subunit B connects the primase to the DNA pol  $\alpha$ . The primase synthesizes RNA primers of 7–10 nts in length, which are then translocated to the active site of DNA pol  $\alpha$  which is further extended by 20–30 deoxynucleotides of DNA primers. Only the DNA primers are extended by the replicative DNA pols  $\epsilon$  and  $\delta$  (Fig. 7). [16]

Crystallographic details on the structures for full-length human primase alone and within the primosome complex have been reported. [17, 18] The Crystallographic data showed the putative active sites, including three invariant D residues (D<sup>109</sup>, D<sup>111</sup>, and D<sup>306</sup>) in the catalytic subunit, p48. The proposed catalytic amino acid R and template-binding –YG- pair are based on the sequence similarity with other DNA/RNA polymerases. [3, 19] Further insights on the primase active sites were provided by SDM experiments on mouse and human primases. For example, the SDM analysis of the mutant proteins of mouse primase indicated that residues 104–111 were most critical for primer synthesis and formed part of the active site. Ala substitution in the three invariant D residues, viz.

D<sup>105</sup>→A, D<sup>109</sup>→A and D<sup>111</sup>→A (highlighted in dark blue (Fig. 9) produced proteins with no detectable activity in direct primase assays, indicating that these residues might form part of a conserved carboxylic triad which is also observed in the active sites of DNA polymerases and reverse transcriptases. Furthermore, mutation of two residues, R<sup>162</sup> and R<sup>163</sup> of mouse primase (Fig. 9) caused an increase in  $K_{M(NTP)}$  and suggested that the initiation and elongation may use the same active site. [20]

The human primase was also subjected to SDM analysis. Vaithilingam et al., found that the two residues (S<sup>160</sup> and H<sup>166</sup>, highlighted in dark blue) which are in direct contact with the nucleotide that were previously unrecognized as critical to the human primase active site. [21] Interestingly, all the 5 amino acids are found in the highly conserved block. Kilkenny et al. also reported similar findings with the human primase. [22] With SDM they found that in human primase also the Ala mutants in the catalytic triad D<sup>109</sup>, D<sup>111</sup>, and D<sup>306</sup>, lost all their ability to synthesize RNA primers. Furthermore, Ala mutation, viz. H<sup>166</sup>→A (highlighted in dark blue) abolished primer synthesis by human primase, as reported previously for the equivalent residue in archaeal and prokaryotic prim-fold polymerases. The particular H is found in the ZBM by MSA analysis. Interestingly, the H<sup>315</sup>→A mutant (highlighted in dark blue) was almost impaired and lead to the loss of primer synthesis. The important role of K<sup>318</sup> in NTP binding is highlighted by the inability of the K<sup>318</sup>→A mutant (highlighted in dark blue) to perform synthesis of RNA primers.

These results suggest the initial two NTPs could possibly bind in the highly conserved block of amino acids from 300 to 334 (Fig. 9). As BLASTp analysis showed no significant similarity between the bacterial and human primases, the bacterial primases could serve as a good candidate as drug targets. In addition to, the CTD of the regulatory subunit of the eukaryotic primases revealed an iron-sulphur (4Fe-4S) cluster coordinated by four Cys residues that are essential for primer synthesis. [23, 24] Martínez-Jiménez et al. found by deletion mutagenesis, that the ZBM was important for

NTP-binding during initiation. [25] They found that in the deletion mutant  $\Delta$ ZMB ( $\Delta$ 410-560), the ZBM was required for the binding and selection of the first nucleotide (preferentially an A) of the new primer strand. This was evident, when a preformed dimer \*pAG was provided as primer, as it was similarly elongated by the wild-type and the  $\Delta$ ZBM mutant, suggesting that the ZBM is dispensable once the dimer is formed. [25] Based on the SDM, X-ray crystallographic data and sequence similarity in MSAs, the proposed amino acids at the active site and the primase reactions of the human primase are shown in Fig. 10.

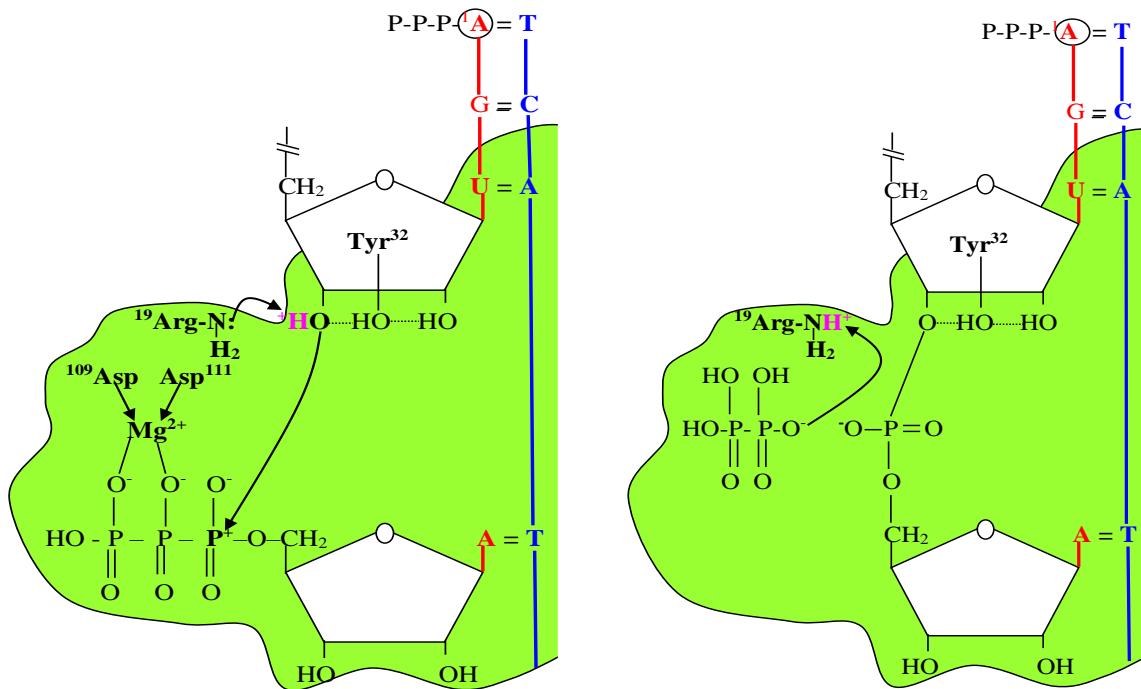


Fig. 10 The proposed active site structure of human primase and the priming reactions (numbering from *H. sapiens*)

#### DNA-Priming by DNA Polymerase $\alpha$ in Eukaryotes: Structural Features

Unlike the prokaryotic replication initiation complex, the eukaryotic replication initiation complex uses an additional priming enzyme in the initiation process. The second enzyme of the eukaryotic replication initiation process is the DNA pol  $\alpha$ , which is predominantly found in proliferating cells as it is coupled to replication in eukaryotic cells. During the S phase of the cell-

cycle, the DNA polymerase  $\alpha$  complex (composed of a catalytic subunit POLA1, a regulatory subunit POLA2 and two primase subunits PRIM1 and PRIM2) is recruited to the replication forks via direct interactions with WDHD1 (which acts as a replication initiation factor bringing together the MCM helicase and the DNA polymerase  $\alpha$ /primase complex to initiate DNA replication).

The eukaryotic RNA primase forms a tight complex with DNA pol  $\alpha$  and it passes newly synthesized RNA primer to the DNA pol  $\alpha$ , via a direct handoff. Now the DNA polymerase  $\alpha$  complex initiates synthesis of DNA primers on the short RNA primers on both leading- and lagging-strands which are subsequently transferred to polymerase  $\delta$  band polymerase  $\epsilon$  for elongation of the lagging- and leading-strands, respectively. In fact, the large, p58 subunit of the eukaryotic RNA primase passes the newly generated RNA primer-template from the primase to the DNA pol  $\alpha$ . [26] Thus, DNA pol  $\alpha$  is involved both in DNA replication and cell-cycle regulation and hence, is tightly synchronized with the cell-cycle progression. [27] Therefore, this is known as the initiator polymerase of eukaryotic genome replication during cell division. DNA pol  $\alpha$  is a DdDp (EC 2.7.7.7) (also known as pol A in human and as pol I in yeasts). As mentioned elsewhere, the primase-DNA pol  $\alpha$  complex is a heterotetramer, composed of two primase subunits and two polymerase subunits. [12] The two polymerase subunits are known as A with a molecular mass of 180 kDa, (the catalytic subunit), and B with a molecular mass of 68 kDa, (the regulatory subunit). The larger subunit A possesses the polymerase activity, whereas the smaller B subunit interacts with subunit A. Its interaction with A is important for the nuclear import of the DNA pol  $\alpha$ . As such, the DNA pol  $\alpha$  is a low-fidelity enzyme with an error rate of  $10^{-4}$ - $10^{-5}$  as compared to the error rates of  $10^{-6}$ - $10^{-7}$  of the main replicative DNA polymerases, viz. pol  $\delta$  and pol  $\epsilon$ . This is because, although DNA pol  $\alpha$  contains an exonuclease domain, no proofreading activity has been observed and its errors are, therefore, corrected by the other proofreading-proficient enzyme DNA pol  $\delta$ . [28] The DNA pol  $\alpha$  exhibits a preference for an RNA primer whereas the other two replicative polymerases, viz. pol  $\delta$  and pol  $\epsilon$  use only the DNA primers for further elongation. Therefore, the

intermediate step of synthesizing DNA primers in eukaryotes is performed by the DNA pol  $\alpha$ . It has two ZBM, but reported to be devoid of any exonuclease function. [28] Like RNA primases, the DNA primase can start with dATP and dGTP.

Figure 11 shows the MSA analysis of the DNA pol  $\alpha$  from animal sources. Although their primary structures are markedly different from other DNA polymerases, the crystal structures are found to be remarkably similar in overall shape. For example, 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 D residues. The palm domain also shows the highly conserved-<sup>864</sup>DFNSLYPS<sup>871</sup>- which form a short helix and involve in dNTP binding. Many Cs are conserved but 8 of them (highlighted in orange) form the two ZBM and are found as direct and inverted repeats (-CX2C-----CX4C-) and (-CX4C-----CX2C-) and perform two different functions as discussed elsewhere (Fig. 11). The NTD is highly rich in acidic amino acids (highlighted in red); there are at least two completely conserved direct repeats of triads, 3D/Es (marked with arrows) suggesting a possible role in the polymerase function, maybe binding to highly basic histone proteins on the chromosomes. A highly conserved 8 amino acid-motif consisting of 4Ls and 4Ks, arranged consecutively (marked with arrows) is also observed suggesting an important role in the polymerase function. The two invariant -DxD-motifs, viz. -YGDTDS-and -DID- are implicated in metal-binding and are reported in other DNA polymerases too. [3] The middle and the C-terminal regions are highly conserved in all (data not shown). By sequence similarity, the proposed amino acids in the PR active site of DNA pol  $\alpha$  (Table 4) are highlighted in light blue, even though the activity is reported to be muted. [28]

CLUSTAL O (1.2.4) MSA of the DNA polymerase  $\alpha$  from animal sources

tr A0A6P8GL87 A0A6P8GL87_CLUHA	EKVGRKSALEQLKKAKAGEKVKY	ELEECEENVY	EVDVE	AQYSRMRMVRD	RQDDDI	WII	DDDGMG	94
tr A0A1SNC92 A0A1SNC92_SALSA	EKTGRKSALEQLKKAKAGEKVKY	EVEDLTSVY	EVEDE	QYSRMRMVRD	RQDDDI	WII	DDDGTG	120
tr A0A6P8T108 A0A6P8T108_GYMAC	EKTGRKSALEQLKKAKAGEKVKY	EVEKFSSVY	EVEDE	QYSRMRMVRD	RQDDDI	WII	DDDGTG	94
tr A0A3P9D223 A0A3P9D223_9CICH	EKVGRKSALEQLKKAKAGEKVKY	EVEKFSSVY	EVEDE	QYSRMRMVRD	RQDDDI	WII	DDDGTG	88
tr A0A3P9PMI28 A0A3P9MI28_ORYLA	EKVGRKSALEQLKKAKAGEKVKY	KVEEFHSVY	EVEDE	QYSRMRMVRD	RQDDDI	WII	DDDGTG	95
tr A0A6J1UIRO A0A6J1UIRO_9SAUR	SKKGRQEALERLKKAKAGEKVKY	EVEEFGTIY	DEIDE	QYSRMRMVRD	RQDDDI	WIV	DDDGVG	92
tr A0A6P9D018 A0A6P9D018_PANGU	SKKGRQEALERLKKAKAGEKVKY	EVEEFGTIY	DEIDE	QYSRMRMVRD	RQDDDI	WIV	DDDGVG	92
sp P33609 DPOLA_MOUSE	SKKGRQEALERLKKAKAGEKVKY	EVEEDLTSVY	EVEDE	QYSRMRMVRD	RQDDDI	WIV	DDDGVG	92
sp O89042 DPOLA_RAT	SKKGRQEALERLKKAKAGEKVKY	EVEDLTSVY	EVEDE	QYSRMRMVRD	RQDDDI	WIV	DDDGVG	92
tr A0A7J8AWU5 A0A7J8AWU5_RHIFE	SKKGRQEALERLKKAKAGEKVKY	EVEDLTSVY	EVEDE	QYSRMRMVRD	RQDDDI	WIV	DDDGVG	86
tr G1T8Z5 G1T8Z5_RABIT	SKKGRQEALERLKKAKAGEKVKY	EVEDLTSVY	EVEDE	QYSRMRMVRD	RQDDDI	WIV	DDDGVG	89
tr HOUXU4 HOUXU4_CAVPO	SKKGRQEALERLKKAKAGEKVKY	EVEDLTSVY	EVEDE	QYSRMRMVRD	RQDDDI	WIV	DDDGVG	92
tr G3RT59 G3RT59_GORGO	SKKGRQEALERLKKAKAGEKVKY	EVEDLTSVY	EVEDE	QYSRMRMVRD	RQDDDI	WIV	DDDGVG	93
sp P09884 DPOLA_HUMAN	SKKGRQEALERLKKAKAGEKVKY	EVEDLTSVY	EVEDE	QYSRMRMVRD	RQDDDI	WIV	DDDGVG	86
tr A0A6D2UX8 A0A6D2UX8_PANTR	SKKGRQEALERLKKAKAGEKVKY	EVEDLTSVY	EVEDE	QYSRMRMVRD	RQDDDI	WIV	DDDGVG	92
tr K7CUU4 K7CUU4_PANTR	SKKGRQEALERLKKAKAGEKVKY	EVEDLTSVY	EVEDE	QYSRMRMVRD	RQDDDI	WIV	DDDGVG	92
tr A0A6P6ITS8 A0A6P6ITS8_PUMCO	SKKGRQEALERLKKAKAGEKVKY	EVEDLTSVY	EVEDE	QYSRMRMVRD	RQDDDI	WIV	DDDGVG	92
tr A0A6J0A827 A0A6J0A827_ACIJB	SKKGRQEALERLKKAKAGEKVKY	EVEDLTSVY	EVEDE	QYSRMRMVRD	RQDDDI	WIV	DDDGVG	92
tr A0A6P4TPK1 A0A6P4TPK1_PANPR	SKKGRQEALERLKKAKAGEKVKY	EVEDLTSVY	EVEDE	QYSRMRMVRD	RQDDDI	WIV	DDDGVG	114
tr F6ZMX7 F6ZMX7_HORSE	SKKGRQEALERLKKAKAGEKVKY	EVEDLTSVY	EVEDE	QYSRMRMVRD	RQDDDI	WIV	DDDGVG	92
tr A0A2Y9M558 A0A2Y9M558_DELLE	SKKGRQEALERLKKAKAGEKVKY	EVEDLTSVY	EVEDE	QYSRMRMVRD	RQDDDI	WIV	DDDGVG	92
tr A0A3Q1LMS8 A0A3Q1LMS8_BOVIN	SKKGRQQALERLKKAKAGEKVKY	EVEDLTSVY	EVEDE	QYSRMRMVRD	RQDDDI	WIV	DDDGVG	92
tr A0A452EXN8 A0A452EXN8_CAPHI	SKKGRQEALERLKKAKAGEKVKY	EVEDLTSVY	EVEDE	QYSRMRMVRD	RQDDDI	WIV	DDDGVG	86
tr L8IUV3 L8IUV3_9CETA	SKKGRQQALERLKKAKAGEKVKY	EVEDLTSVY	EVEDE	QYSRMRMVRD	RQDDDI	WIV	DDDGVG	79
tr A0A452EXY8 A0A452EXY8_CAPHI	SKKGRQEALERLKKAKAGEKVKY	EVEDLTSVY	EVEDE	QYSRMRMVRD	RQDDDI	WIV	DDDGVG	92
tr W5Q828 W5Q828_SHEEP	SKKGRQEALERLKKAKAGEKVKY	EVEDLTSVY	EVEDE	QYSRMRMVRD	RQDDDI	WIV	DDDGVG	92
tr A0A6P3TAR3 A0A6P3TAR3_SHEEP	SKKGRQEALERLKKAKAGEKVKY	EVEDLTSVY	EVEDE	QYSRMRMVRD	RQDDDI	WIV	DDDGVG	92
tr A0A1U7R9B9 A0A1U7R9B9_ALLSI	SKKGRQEALERLKKAKAGEKVKY	EVEDLTSVY	EVEDE	QYSRMRMVRD	RQDDDI	WIV	DDDGVG	92
tr A0A151N032 A0A151N032_ALLM	SKKGRQEALERLKKAKAGEKVKY	EVEDLTSVY	EVEDE	QYSRMRMVRD	RQDDDI	WIV	DDDGVG	87
tr A0A663MCX2 A0A663MCX2_ATHCN	SRRGQQEALERLKKAKAGEKVKY	EVEEFTGTVY	DEVDE	QYSRMRMVRD	RQDDDI	WIV	DDDGVG	87
tr A0A6J0I7H9 A0A6J0I7H9_9PASS	NRRGQQEALERLKKAKAGEKVKY	EVEEFTGTVY	DEVDE	QYSRMRMVRD	RQDDDI	WIV	DDDGVG	88
tr A0A7K5Y128 A0A7K5Y128_9CHAR	SRRGQQEALERLKKAKAGEKVKY	EVEEFTGTVY	DEIDE	QYSRMRMVRD	RQDDDI	WIV	DDDGVG	77
tr A0A7L0MV49 A0A7L0MV49_9PSIT	SRRGQQEALERLKKAKAGEKVKY	EVEEFTGTVY	DEIDE	QYSRMRMVRD	RQDDDI	WIV	DDDGVG	77
.. : * : * : * : * ..	: * : * : * : * ..	: * : * : * : * ..	: * : * : * : * ..	: * : * : * : * ..	: * : * : * : * ..	: * : * : * : * ..	: * : * : * : * ..	
tr A0A6P8GL87 A0A6P8GL87_CLUHA	YVEDGREIE	DEEELDDDAALK	-GAKGKAGSKAGDAGAKNNMKKVSLAKPNNSIKSLFMNSNVK					153
tr A0A1SNC92 A0A1SNC92_SALSA	YVEDGREIE	DEDLADDV	LEKGAKGKVGAKG--ADSKNNIKAAVAKPNISKNLFMNSNVK					178
tr A0A6P8T108 A0A6P8T108_GYMAC	YVEDGREIE	DDDLDDDV	EN-NKGAGSKG--AAASKNNMKKSVAKSNTIKSLFMNSNVK					151
tr A0A3P9D223 A0A3P9D223_9CICH	YVEDGREIE	DDDLDDDV	VED---KRGAKG--ADSRKNVKSVAKPNNSIKSLFMNSNVK					142
tr A0A3P9PMI28 A0A3P9MI28_ORYLA	YVEDGREIE	DDDLDDDV	VEK--RGKGGAKG--AESKKTVKKAATKPNTIKSLFMNSNVK					151
tr A0A6J1UIRO A0A6J1UIRO_9SAUR	YVEDGREIE	DEDDELDA	ALNSDKR-KRKDA-STYDKNLKTKLTVSKPNNIKSMFMAN					150
tr A0A6P9D018 A0A6P9D018_PANGU	YVEDGREIE	DEDELEDA	ALNSDEKR-RKDTA-STYDENKLKTLVSKPNNTIKSMFVANT					150
sp P33609 DPOLA_MOUSE	YVEDGREIE	FDDLEDDALDTCGK	-GSDGKA-HRKDRKVKKPSVTKPNNIKAMFIASAGK					150
sp O89042 DPOLA_RAT	YVEDGREIE	DDDLDEDA	ALDTGCE-GSDGKA-HRKDRKVKKPSVTKPNNIKAMFIASAGK					150
tr A0A7J8AWU5 A0A7J8AWU5_RHIFE	YVEDGREIE	DDDLDEDA	ALDSREKG--KDKA-QNKEKRNVKAATKPNNNIKSMFIASAGK					143
tr G1T8Z5 G1T8Z5_RABIT	YVEDGREIE	DDDLDEDA	TLDASEK-GKDKVKT--NKGKRNVKKPSVTKPNNIKTMLAG					147
tr HOUXU4 HOUXU4_CAVPO	YVEDGREIE	DDDLDDDAFDTSEK	GKDGT-YNKKDNVKKPVTVPKPNNIKSMFIASAGK					150
tr G3RT59 G3RT59_GORGO	YVEDGREIE	DDDLDEDA	ALDADEK-GKDGA-RNKDKRNVKKLAVTKPNNIKSMFIACAG					151
sp P09884 DPOLA_HUMAN	YVEDGREIE	FDDDEDDALDADEK	-GKDGA-RNKDKRNVKKLAVTKPNNIKSMFIACAGK					144
tr A0A6D2UX8 A0A6D2UX8_PANTR	YVEDGREIE	DDDLDEDA	ALDADEK-GKDGA-RNKDKRNVKKLAVTKPNNIKSMFIACAGK					150
tr K7CUU4 K7CUU4_PANTR	YVEDGREIE	DDDLDEDA	ALDADEK-GKDGA-RNKDKRNVKKLAVTKPNNIKSMFIACAGK					150
tr A0A6P6ITS8 A0A6P6ITS8_PUMCO	YVEDGREIE	DDDLDEDA	GKDKA-CNKDKRTVRAAVTKPNNIKSMFIACAGK					150
tr A0A6J0A827 A0A6J0A827_ACIJB	YVEDGREIE	DDDLDEDA	ALDSYEK-GKDEKA-RTKDRRNVKKAVVTPKPNNIKSMFIASAGK					172
tr A0A6P4TPK1 A0A6P4TPK1_PANPR	YVEDGREIE	DDDLDEDA	ALDSHEKAGKDDKA-RNKDKRNVKKAVVTPRNSNIKSMFMASAGR					151
tr F6ZMX7 F6ZMX7_HORSE	YVEDGREIE	DDDLDEDA	ALDSHEK-GKNNKA-CSKDKRNVKKAMVTPKPNNIKSMFMASAGR					150
tr A0A2Y9M558 A0A2Y9M558_DELLE	YVEDGREIE	DDDLDEDA	ALDSHEK-GKDNKA-CNKDKRTVRAAVTPKPNNIKSMFIASAGR					150
tr A0A3Q1LMS8 A0A3Q1LMS8_BOVIN	YVEDGREIE	DDDLDEDA	ALDSHEK-GKDKA-CNKDKRTVRAAVTPKPNNIKSMFIASAGR					144
tr A0A452EXN8 A0A452EXN8_CAPHI	YVEDGREIE	DDDLDEDA	ALDSHEK-GKDKA-CNKDKRTVRAAVTPKPNNIKSMFIASAGR					137
tr L8IUV3 L8IUV3_9CETA	YVEDGREIE	DDDLDEDA	ALDSHEK-GKDKA-CNKDKRTVRAAVTPKPNNIKSMFIASAGR					150
tr A0A452EXY8 A0A452EXY8_CAPHI	YVEDGREIE	DDDLDEDA	ALDSHEK-GKDKA-CNKDKRTVRAAVTPKPNNIKSMFIASAGR					150
tr W5Q828 W5Q828_SHEEP	YVEDGREIE	DDDLDEDA	ALDSHEK-GKDKA-CNKDKRTVRAAVTPKPNNIKSMFIASAGR					150
tr A0A6P3TAR3 A0A6P3TAR3_SHEEP	YVEDGREIE	DDDLDEDA	ALDSHEK-GKDKA-CNKDKRTVRAAVTPKPNNIKSMFIASAGR					150
tr A0A1U7R9B9 A0A1U7R9B9_ALLSI	YVEDGREIE	DDDLDEDA	ALDSHEK-GKDKA-CNKDKRTVRAAVTPKPNNIKSMFIASAGR					144
tr A0A151N032 A0A151N032_ALLM	YVEDGREIE	DDDLDEDA	ALDSHEK-GKDKA-CNKDKRTVRAAVTPKPNNIKSMFIASAGR					133
tr A0A663MCX2 A0A663MCX2_ATHCN	YVEDGREIE	DDDLDEDA	ALDSHEK-GKDKA-CNKDKRTVRAAVTPKPNNIKSMFIASAGR					143
tr A0A6J0I7H9 A0A6J0I7H9_9PASS	YVEDGREIE	DDDLDEDA	ALDSHEK-GKDKA-CNKDKRTVRAAVTPKPNNIKSMFIASAGR					145
tr A0A7K5Y128 A0A7K5Y128_9CHAR	YVEDGREIE	DDDLDEDA	ALDSHEK-GKDKA-CNKDKRTVRAAVTPKPNNIKSMFIASAGR					146
tr A0A7L0MV49 A0A7L0MV49_9PSIT	YVEDGREIE	DDDLDEDA	ALGSNKK-GKGGKT-CTVDKKNVKKSVVSKPNNTIKSMFIASAGK					135
								135
*****	*****	*****	*****	*****	*****	*****	*****	

tr A0A6P8GL87 A0A6P8GL87_CLUHA	LGFDLEVLQLRQI SVCKVPHWSKIGRLRRANMPKLCGRGGYAEKSATCGRLVCDVEISAKE	708
tr A0A1S3NC92 A0A1S3NC92_SALSA	FGFDLEVLQLRQI TVCKVPHWSKIGRLRSNMPKLGGRSGFAEKSATCGRLVCDVEISAKE	734
tr A0A6P8T108 A0A6P8T108_GYMAC	FGFDLEVLQLRQI NVCKVPHWSKIGRLRSNMPKLGGRGSFAEKSATCGRLVCDVEISAKE	711
tr A0A3P9D2Z3 A0A3P9D2Z3_9CICH	FGFDLEVLQLRQI NVCKVPHWSKIGRLRSNMPKLGGRSGFAEKSATCGRLVCDVEISAKE	578
tr A0A3P9PMI28 A0A3P9MI28_ORYLA	FGFDLEVLQLRQI NVCKVPHWSKIGRLRSNMPKLGGRSGFAEKSATCGRLVCDVEISAKE	702
tr A0A6J1UIRO A0A6J1UIRO_9SAUR	YGFDFLELLQLRQI NFCKVPHWSKIGRLRSNMPKLGGRSGFAEKSATCGRMVGDIEISAKE	713
tr A0A6P9D018 A0A6P9D018_PANGU	YGFDFLELLQLRQI NFCKVPHWSKIGRLRSNMPKLGGRSGFAEKSATCGRMVGDIEISAKE	712
sp P33609 DPOLA_MOUSE	CSFELEVLLQLRQI NECKVPHWSKIGRLRSNMPKLGGRSGFGERNATCGRMI CDVEISAKE	703
sp O89042 DPOLA_RAT	CGFELEVLLQLRQI NECKVPHWSKIGRLRSNMPKLGGRSGFGERNATCGRMI CDVEISAKE	706
tr A0A7J8AWU5 A0A7J8AWU5_RHIFE	YGFDFELEVLLQLRQI NVCKVPHWSKIGRLRSNMPKLGGRSGFGERNATCGRMI CDVEISAKE	698
tr G1T8Z5 G1T8Z5_RABIT	YGFDFELEVLLQLRQI NVCKVPHWSKIGRLRSNMPKLGGRSGFGERNATCGRMI CDVEISAKE	704
tr HOUXU4 HOUXU4_CAVPO	YGFDFELEVLLQLRQI NVCKVPHWSKIGRLRSNMPKLGGRSGFGERNATCGRMI CDVEISAKE	705
tr G3RT59 G3RT59_GORGO	YGFDFELEVLLQLRQI NVCKVPHWSKIGRLRSNMPKLGGRSGFGERNATCGRMI CDVEISAKE	706
sp P09884 DPOLA_HUMAN	YGFDFELEVLLQLRQI INVCKAPHWSKIGRLRSNMPKLGGRSGFGERNATCGRMI CDVEISAKE	699
tr A0A6D2XUY8 A0A6D2XUY8_PANTR	YGFDFELEVLLQLRQI INVCKAPHWSKIGRLRSNMPKLGGRSGFGERNATCGRMI CDVEISAKE	705
tr K7CUU4 K7CUU4_PANTR	YGFDFELEVLLQLRQI INVCKAPHWSKIGRLRSNMPKLGGRSGFGERNATCGRMI CDVEISAKE	705
tr A0A6P6ITS8 A0A6P6ITS8_PUMCO	YGFDFELEVLLQLRQI INVCKVPHWSKIGRLRSNMPKLGGRSGLGERNATCGRMI CDVEISAKE	705
tr A0A6J0A827 A0A6J0A827_AC1JB	YGFDFELEVLLQLRQI INVCKVPHWSKIGRLRSNMPKLGGRSGLGERNATCGRMI CDVEISAKE	705
tr A0A6P4TPK1 A0A6P4TPK1_PANPR	YGFDFELEVLLQLRQI INVCKVPHWSKIGRLRSNMPKLGGRSGLGERNATCGRMI CDVEISAKE	727
tr F6ZMX7 F6ZMX7_HORSE	YGFDFELEVLLQLRQI INVCKVPHWSKIGRLRSNMPKLGGRSGLGERNATCGRMI CDVEISAKE	706
tr A0A2Y9M558 A0A2Y9M558_DELLE	YGFDFELEVLLQLRQI INVCKVPHWSKIGRLRSNMPKLGGRSGLGERNATCGRMI CDVEISAKE	705
tr A0A3Q1LMS8 A0A3Q1LMS8_BOVIN	YGFDFELEVLLQLRQI INVCKVPHWSKIGRLRSNMPKLGGRSGLGERNATCGRMI CDVEISAKE	705
tr A0A452EXN8 A0A452EXN8_CAPI	YGFDFELEVLLQLRQI INVCKVPHWSKIGRLRSNMPKLGGRSGLGERNATCGRMI CDVEISAKE	703
tr L8IUV3 L8IUV3_9CETA	YGFDFELEVLLQLRQI INVCKVPHWSKIGRLRSNMPKLGGRSGLGERNATCGRMI CDVEISAKE	692
tr A0A452EXY8 A0A452EXY8_CAPI	YGFDFELEVLLQLRQI INVCKVPHWSKIGRLRSNMPKLGGRSGLGERNATCGRMI CDVEISAKE	692
tr W5Q828 W5Q828_SHEEP	YGFDFELEVLLQLRQI INVCKVPHWSKIGRLRSNMPKLGGRSGLGERNATCGRMI CDVEISAKE	705
tr A0A6P3STAR3 A0A6P3STAR3_SHEEP	YGFDFELEVLLQLRQI INVCKVPHWSKIGRLRSNMPKLGGRSGLGERNATCGRMI CDVEISAKE	705
tr A0A1U7R9B9 A0A1U7R9B9_ALLSI	YGFDFELEVLLQLRQI INVCKVPHWSKIGRLRSNMPKLGGRGGFAERNAA CGRMI CDVEISAKE	715
tr A0A151N032 A0A151N032_ALLMI	YGFDFELEVLLQLRQI NACKVPHWSKIGRLRSNMPKLGGRGGFAERNAA CGRMI CDVEISAKE	725
tr A0A663MCX2 A0A663MCX2_ATHCN	YGFDFELEVLLQLRQI INVCKVPHWSKIGRLRSNMPKLGGRGGFAERNAA CGRMI CDVEISAKE	609
tr A0A6J0I7H9 A0A6J0I7H9_9PASS	YGFDFELEVLLQLRQI NLCKVPHWSKIGRLRSNMPKLGGRGGFAERSAA CGRMI CDVEISAKE	737
tr A0A7K5Y128 A0A7K5Y128_9CHAR	YGFDFELEVLLQLRQI NACKVPHWSKIGRLRSNMPKLGGRGGFAERNAA CGRMI CDVEISAKE	717
tr A0A7L0MV49 A0A7L0MV49_9PSIT	YGFDFELEVLLQLRQI INVCKVPHWSKIGRLRSNMPKLGGRGGFAERNAA CGRMI CDVEISAKE	716
*:***:*****: *.*:*****:*. *.*:*****:*. *.*:*****:*. *.*:*****:*. *.*:*****:*		
tr A0A6P8GL87 A0A6P8GL87_CLUHA	LIROKSYHHLTELASQVLKTERAVIPQENIRNFYSDSPHLLYVLELTWMDAKLILQIMCEL	768
tr A0A1S3NC92 A0A1S3NC92_SALSA	LIROKSYHHLTELAAQVLKTERATIPAENIKNLYSDSPHLLYLLELTWTDALKLILQIMCEL	794
tr A0A6P8T108 A0A6P8T108_GYMAC	LIROKSYHHLTELAAQVLKTERATIPOEDIKNLYSDSPHLLYLLELTWMDAKLILQIMCEL	771
tr A0A3P9D2Z3 A0A3P9D2Z3_9CICH	LIROKSYHHLTELAAQVLKTERITVPQEDIKNLYSDSPHLLYLLELTWTDALKLILQIMCEL	638
tr A0A3P9PMI28 A0A3P9MI28_ORYLA	LIROKSYHHLTELAAQVLKTERIVTVPOEEVKNLYSDSPHLLYLLELTWTDALKLILQIMCEL	762
tr A0A6J1UIRO A0A6J1UIRO_9SAUR	LIROKSYHHLSELVSQILKTERVAIPEEVVNMYSDHQLLYMLENTWIDSKYILQIMYEL	773
tr A0A6P9D018 A0A6P9D018_PANGU	LIROKSYHHLSELVSQILKTERVAIPEEVVNMYSDHQLLYMLENTWIDSKYILQIMYEL	772
sp P33609 DPOLA_MOUSE	LIROKSYHHLSELVSQILKTERVAIPEEVVNMYSDHQLLYMLENTWIDSKYILQIMCEL	772
sp O89042 DPOLA_RAT	LIROKSYHHLSELVSQILKTERVAIPEEVVNMYSDHQLLYMLENTWIDSKYILQIMCEL	763
tr A0A7J8AWU5 A0A7J8AWU5_RHIFE	LIROKSYHHLSELVSQILKTERVAIPEEVVNMYSDHQLLYMLENTWIDSKYILQIMCEL	766
tr G1T8Z5 G1T8Z5_RABIT	LIROKSYHHLSELVSQILKTERVAIPEEVVNMYSDHQLLYMLENTWIDSKYILQIMCEL	758
tr HOUXU4 HOUXU4_CAVPO	LIROKSYHHLSELVSQILKTERVAIPEEVVNMYSDHQLLYMLENTWIDSKYILQIMCEL	764
tr G3RT59 G3RT59_GORGO	LIROKSYHHLSELVSQILKTERVAIPEEVVNMYSDHQLLYMLENTWIDSKYILQIMCEL	765
sp P09884 DPOLA_HUMAN	LIROKSYHHLSELVSQILKTERVAIPEEVVNMYSDHQLLYMLENTWIDSKYILQIMCEL	766
tr A0A6D2XUY8 A0A6D2XUY8_PANTR	LIROKSYHHLSELVSQILKTERVVIPMENIQNMYSSESSQQLLYILLEHTWKDAKFILQIMCEL	759
tr K7CUU4 K7CUU4_PANTR	LIROKSYHHLSELVSQILKTERVVIPMENIQNMYSSESSQQLLYILLEHTWKDAKFILQIMCEL	765
tr A0A6P6ITS8 A0A6P6ITS8_PUMCO	LIROKSYHHLSELVSQILKTERVVIPMENIQNMYSSESSQQLLYILLEHTWKDAKFILQIMCEL	765
tr A0A6J0A827 A0A6J0A827_AC1JB	LIROKSYHHLSELVSQILKTERVVIPMENIQNMYSSESSQQLLYILLEHTWKDAKFILQIMCEL	765
tr A0A6P4TPK1 A0A6P4TPK1_PANPR	LIROKSYHHLSELVSQILKTERVVIPMENIQNMYSSESSQQLLYILLEHTWKDAKFILQIMCEL	765
tr F6ZMX7 F6ZMX7_HORSE	LIROKSYHHLSELVSQILKTERVVIPMENIQNMYSSESSQQLLYILLEHTWKDAKFILQIMCEL	787
tr A0A2Y9M558 A0A2Y9M558_DELLE	LIROKSYHHLSELVSQILKTERVVIPMENVNMYSESSHLLYLLEHTWKDAKFILQIMCEL	766
tr A0A3Q1LMS8 A0A3Q1LMS8_BOVIN	LIROKSYHHLSELVSQILKTERVVIPMENVNMYSESSHLLYLLEHTWKDAKFILQIMCEL	765
tr A0A452EXN8 A0A452EXN8_CAPI	LIROKSYHHLSELVSQILKTERVVIPMENVNMYSESSHLLYLLEHTWKDAKFILQIMCEL	765
tr L8IUV3 L8IUV3_9CETA	LIROKSYHHLSELVSQILKTERVVIPMENVNMYSESSHLLYLLEHTWKDAKFILQIMCEL	765
tr A0A452EXY8 A0A452EXY8_CAPI	LIROKSYHHLSELVSQILKTERVVIPMENVNMYSESSHLLYLLEHTWKDAKFILQIMCEL	765
tr W5Q828 W5Q828_SHEEP	LIROKSYHHLSELVSQILKTERVVIPMENVNMYSESSHLLYLLEHTWKDAKFILQIMCEL	765
tr A0A6P3STAR3 A0A6P3STAR3_SHEEP	LIROKSYHHLSELVSQILKTERVVIPIETIRNMYSDSHSHLLYLLEHTWKDAKFILQIMCEL	765
tr A0A1U7R9B9 A0A1U7R9B9_ALLSI	LIROKSYHHLSELVSQILKTERVVIPIETIRNMYSDSHSHLLYLLEHTWKDAKFILQIMCEL	765
tr A0A151N032 A0A151N032_ALLMI	LIROKSYHHLSELVSQILKTERVVIPIETIRNMYSDSHSHLLYLLEHTWKDAKFILQIMCEL	765
tr A0A663MCX2 A0A663MCX2_ATHCN	LIROKSYHHLSELVSQILKTERVVIPIETIRNMYSDSHSHLLYLLEHTWKDAKFILQIMCEL	765
tr A0A6J0I7H9 A0A6J0I7H9_9PASS	LIROKSYHHLSELVSQILKTERVVIPIETIRNMYSDSHSHLLYLLEHTWKDAKFILQIMCEL	765
tr A0A7K5Y128 A0A7K5Y128_9CHAR	LIROKSYHHLSELVSQILKTERVVIPIETIRNMYSDSHSHLLYLLEHTWKDAKFILQIMCEL	765
tr A0A7L0MV49 A0A7L0MV49_9PSIT	LIROKSYHHLSELVSQILKTERVVIPIETIRNMYSDSHSHLLYLLEHTWKDAKFILQIMCEL	776
***:*****:***: *.*:*****:*. *.*:*****:*. *.*:*****:*. *.*:*****:*		

tr A0A6P8GL87 A0A6P8GL87_CLUHA	NVLPLALQITNIAGNMSRTLMMGGRSERNEFLLLHAFHEKNYIVD <b>KQIFKPKPQHDLAED</b>	828
tr A0A1S3NC92 A0A1S3NC92_SALSA	NVLPLALQITNIAGNVLSSLRTLMGGGRSERNEYLLLHAFHDKNYIVD <b>K2SFKKAAQQELTEG</b>	854
tr A0A6P8T108 A0A6P8T108_GYMAC	NVLPLALQITNIAGNMSRTLMMGGRSERNEFLLLHAFHDKNYIVD <b>KLSFKKTQMEMGEG</b>	831
tr A0A3P9D2Z3 A0A3P9D2Z3_9CICH	NVLPLALQITSIAGNMSRTLMMGGRSERNEFLLLHAFHEKDYIVD <b>K2SFKKAAQLETAEG</b>	698
tr A0A3P9PMI28 A0A3P9MI28_ORYLA	NVLPLALQITNIAGNVLSSLRTLMGGGRSERNEYLLLHAFHDKNYIVD <b>K2SFKKAAQLEMGDG</b>	822
tr A0A6J1UIRO A0A6J1UIRO_9SAUR	NVLPLALQITNIAGNVLSSLRTLMGGGRSERNEYLLLHAFFKKQ--KHIEE	831
tr A0A6P9D018 A0A6P9D018_PANGU	NVLPLALQITNIAGNVLSSLRTLMGGGRSERNEYLLLHAFYKDYIVD <b>KQLFKKQ--KHIEE</b>	830
sp P33609 DPOLA_MOUSE	NVLPLALQITNIAGNMSRTLMMGGRSERNEFLLLHAFYENNYIVD <b>KQIFRKPPQQLGDE</b>	823
sp O89042 DPOLA_RAT	NVLPLALQITNIAGNMSRTLMMGGRSERNEFLLLHAFYENNYIVD <b>KQIFRKPPQKPGDE</b>	826
tr A0A7J8AWU5 A0A7J8AWU5_RHIFE	NVLPLALQITNIAGNMSRTLMMGGRSERNEFLLLHAFYENNYIVD <b>KQIFRKPPQQLVDE</b>	818
tr G1T8Z5 G1T8Z5_RABIT	NVLPLALQITNIAGNMSRTLMMGGRSERNEFLLLHAFYENNYIVD <b>KQIFRKPPQQLGDE</b>	824
tr HOUXU4 HOUXU4_CAVPO	NVLPLALQITNIAGNMSRTLMMGGRSERNEFLLLHAFYENNYIVD <b>KQIFRKPPQQLGDE</b>	825
tr G3RT59 G3RT59_GORGO	NVLPLALQITNIAGNMSRTLMMGGRSERNEFLLLHAFYENNYIVD <b>KQIFRKPPQQLGDE</b>	826
sp P09884 DPOLA_HUMAN	NVLPLALQITNIAGNMSRTLMMGGRSERNEY <b>IVPDKQIFRKPPQQLGDE</b>	819
tr A0A6D2XUY8 A0A6D2XUY8_PANTR	NVLPLALQITNIAGNMSRTLMMGGRSERNEFLLLHAFYENNYIVD <b>KQIFRKPPQQLGDE</b>	825
tr K7CUU4 K7CUU4_PANTR	NVLPLALQITNIAGNMSRTLMMGGRSERNEFLLLHAFYENNYIVD <b>KQIFRKPPQQLGDE</b>	825
tr A0A6P6ITS8 A0A6P6ITS8_PUMCO	NVLPLALQITNIAGNMSRTLMMGGRSERNEYLLLHAFYENNYIVD <b>KQIFRKPPQQLGDE</b>	825
tr A0A6J0A827 A0A6J0A827_ACIJB	NVLPLALQITNIAGNMSRTLMMGGRSERNEYLLLHAFYENNYIVD <b>KQIFRKPPQQLGDE</b>	825
tr A0A6P4TPK1 A0A6P4TPK1_PANPR	NVLPLALQITNIAGNMSRTLMMGGRSERNEYLLLHAFYENNYIVD <b>KQIFRKPPQQLGDE</b>	847
tr F6ZMX7 F6ZMX7_HORSE	NVLPLALQITNIAGNMSRTLMMGGRSERNEYLLLHAFYENNYIVD <b>KQIFRKPPQQLGDE</b>	826
tr A0A2Y9M558 A0A2Y9M558_DELLE	NVLPLALQITNIAGNMSRTLMMGGRSERNEYLLLHAFYENNYIVD <b>KQIFRKPPQQLGDE</b>	825
tr A0A3Q1LMS8 A0A3Q1LMS8_BOVIN	NVLPLALQITNIAGNMSRTLMMGGRSERNEYLLLHAFYENNYIVD <b>KQIFRKPPQQLGDE</b>	825
tr A0A452EXN8 A0A452EXN8_CAPI	NVLPLALQITNIAGNMSRTLMMGGRSERNEYLLLHAFYENNYIVD <b>KQIFRKPPQQLGDE</b>	823
tr L8IUV3 L8IUV3_9CETA	NVLPLALQITNIAGNMSRTLMMGGRSERNEYLLLHAFYENNYIVD <b>KQIFRKPPQQLGDE</b>	812
tr A0A452EXY8 A0A452EXY8_CAPI	NVLPLALQITNIAGNMSRTLMMGGRSERNEYLLLHAFYENNYIVD <b>KQIFRKPPQQLGDE</b>	812
tr W5Q828 W5Q828_SHEEP	NVLPLALQITNIAGNMSRTLMMGGRSERNEYLLLHAFYENNYIVD <b>KQIFRKPPQQLGDE</b>	825
tr A0A6P3STAR3 A0A6P3STAR3_SHEEP	NVLPLALQITNIAGNMSRTLMMGGRSERNEYLLLHAFYENNYIVD <b>KQIFRKPPQQLGDE</b>	825
tr A0A1U7R9B9 A0A1U7R9B9_ALLSI	NVLPLALQITNIISGNMSRTMMGGRSERNEYLLLHAFYERDYIVD <b>KQVFKKPPQQLVDE</b>	835
tr A0A151N032 A0A151N032_ALLMI	NVLPLALQITNIISGNMSRTMMGGRSERNEYLLLHAFYERDYIVD <b>KQVFKKPPQQLVDE</b>	845
tr A0A663MCX2 A0A663MCX2_ATHCN	NVLPLALQITNIISGNMSRTMMGGRSERNEYLLLHAFHEKDYIVD <b>KQVFKKAPQQLVDE</b>	729
tr A0A6J0I7H9 A0A6J0I7H9_9PASS	NVLPLALQITNIISGNMSRTMMGGRSERNEYLLLHAFHEKDYIVD <b>KQVFKKPVQQLVDE</b>	857
tr A0A7K5Y128 A0A7K5Y128_9CHAR	NVLPLALQITNIISGNMSRTMMGGRSERNEYLLLHAFHEKDYIVD <b>KQVFKKPLQQLVDE</b>	837
tr A0A7LOMV49 A0A7LOMV49_9PSIT	NVLPLALQITNIISGNMSRTMMGGRSERNEYLLLHAFHEKDYIVD <b>KQVFKKPLQQLVDE</b>	836
*****	*****	*****

tr A0A6P8GL87 A0A6P8GL87_CLUHA	EDEDGAGKVTAKKTRRKAAYAGGLVLDPKVGFYDKFILLDFN <b>SLYPSIIQEFNICFTTV</b>	888
tr A0A1S3NC92 A0A1S3NC92_SALSA	EEEGDSGKG--KSRRKKAAYAGGLVLDPKVGFYDKFVLLDFN <b>SLYPSIIQEFNICFTTV</b>	912
tr A0A6P8T108 A0A6P8T108_GYMAC	EE--DGGKG--K-RKKKAYAGGLVLDPKVGFYDKFVLLDFN <b>SLYPSIIQEFNICFTTV</b>	886
tr A0A3P9D2Z3 A0A3P9D2Z3_9CICH	EDDVDAKGK--K-RKKKAYAGGLVLDPKVGFYDKFVLLDFN <b>SLYPSIIQEFNICFTTV</b>	755
tr A0A3P9PMI28 A0A3P9MI28_ORYLA	EEDVDAGKG--K-RKKKAYAGGLVLDPKVGFYDKFVLLDFN <b>SLYPSIIQEFNICFTTV</b>	879
tr A0A6J1UIRO A0A6J1UIRO_9SAUR	DEDLENDQNKSKLGRKKAAYSGGLVLPEPKVGFYDKFILLDFN <b>SLYPSIIQEFNICFTTV</b>	891
tr A0A6P9D018 A0A6P9D018_PANGU	DEDLENDQNKSKLGRKKAAYSGGLVLPEPKVGFYDKFILLDFN <b>SLYPSIIQEFNICFTTV</b>	890
sp P33609 DPOLA_MOUSE	DEEIDGDTNKYKKGRKKATYAGGLVLDPKVGFYDKFILLDFN <b>SLYPSIIQEFNICFTTV</b>	883
sp O89042 DPOLA_RAT	DEEIDGDTNKYKKGRKKATYAGGLVLDPKVGFYDKFILLDFN <b>SLYPSIIQEFNICFTTV</b>	886
tr A0A7J8AWU5 A0A7J8AWU5_RHIFE	DEDIDGDTSKYKKGRKKAAAYAGGLVLDPKVGFYDKFILLDFN <b>SLYPSIIQEFNICFTTV</b>	878
tr G1T8Z5 G1T8Z5_RABIT	D-EMDGDTDKYKKGRKKAAAYAGGLVLDPKVGFYDKFVLLDFN <b>SLYPSIIQEFNICFTTV</b>	883
tr HOUXU4 HOUXU4_CAVPO	EEEIDGDTSKYKKGRKKAAAYAGGLVLDPKVGFYDKFILLDFN <b>SLYPSIIQEFNICFTTV</b>	885
tr G3RT59 G3RT59_GORGO	DEEIDGDTNKYKKGRKKAAAYAGGLVLDPKVGFYDKFILLDFN <b>SLYPSIIQEFNICFTTV</b>	886
sp P09884 DPOLA_HUMAN	DEEIDGDTNKYKKGRKKAAAYAGGLVLDPKVGFYDKFILLDFN <b>SLYPSIIQEFNICFTTV</b>	879
tr A0A6D2XUY8 A0A6D2XUY8_PANTR	DEEIDGDTNKYKKGRKKAAAYAGGLVLDPKVGFYDKFILLDFN <b>SLYPSIIQEFNICFTTV</b>	885
tr K7CUU4 K7CUU4_PANTR	DEEIDGDTNKYKKGRKKAAAYAGGLVLDPKVGFYDKFILLDFN <b>SLYPSIIQEFNICFTTV</b>	885
tr A0A6P6ITS8 A0A6P6ITS8_PUMCO	DEDLDGDKANKYKKGRKKAAAYAGGLVLDPKAGLYDKFILLDFN <b>SLYPSIIQEFNICFTTV</b>	885
tr A0A6J0A827 A0A6J0A827_ACIJB	DEDLDGDKANKYKKGRKKAAAYAGGLVLDPKAGLYDKFILLDFN <b>SLYPSIIQEFNICFTTV</b>	885
tr A0A6P4TPK1 A0A6P4TPK1_PANPR	DEDIDGDTNKYKKGRKKAAAYAGGLVLDPKAGLYDKFILLDFN <b>SLYPSIIQEFNICFTTV</b>	907
tr F6ZMX7 F6ZMX7_HORSE	DEDIDGDTNKYKKGRKKAAAYAGGLVLDPKAGLYDKFILLDFN <b>SLYPSIIQEFNICFTTV</b>	886
tr A0A2Y9M558 A0A2Y9M558_DELLE	DEDIDGDTNKYKKGRKKAAAYAGGLVLDPKAGLYDKFILLDFN <b>SLYPSIIQEFNICFTTV</b>	885
tr A0A3Q1LMS8 A0A3Q1LMS8_BOVIN	DEDIDGDTNKYKKGRKKAAAYAGGLVLDPKAGLYDKFILLDFN <b>SLYPSIIQEFNICFTTV</b>	885
tr A0A452EXN8 A0A452EXN8_CAPI	DEDIDGDTNKYKKGRKKAAAYAGGLVLDPKAGLYDKFILLDFN <b>SLYPSIIQEFNICFTTV</b>	883
tr L8IUV3 L8IUV3_9CETA	DEDIDGDTNKYKKGRKKAAAYAGGLVLDPKAGLYDKFILLDFN <b>SLYPSIIQEFNICFTTV</b>	872
tr A0A452EXY8 A0A452EXY8_CAPI	DEDIDGDTNKYKKGRKKAAAYAGGLVLDPKAGLYDKFILLDFN <b>SLYPSIIQEFNICFTTV</b>	872
tr W5Q828 W5Q828_SHEEP	DEDIDGDTNKYKKGRKKAAAYAGGLVLDPKAGLYDKFILLDFN <b>SLYPSIIQEFNICFTTV</b>	885
tr A0A6P3STAR3 A0A6P3STAR3_SHEEP	DEDIDGDTNKYKKGRKKAAAYAGGLVLDPKAGLYDKFILLDFN <b>SLYPSIIQEFNICFTTV</b>	885
tr A0A1U7R9B9 A0A1U7R9B9_ALLSI	DEDIF-DQTKSKTGRKKAAAYAGGLVLDPKAGLYDKFILLDFN <b>SLYPSIIQEFNICFTTV</b>	894
tr A0A151N032 A0A151N032_ALLMI	DEDIF-DQTKSKTGRKKAAAYAGGLVLDPKAGLYDKFILLDFN <b>SLYPSIIQEFNICFTTV</b>	904
tr A0A663MCX2 A0A663MCX2_ATHCN	DEDIF-DQNKSKIGKKKAAAYAGGLVLDPKAGLYDKFILLDFN <b>SLYPSIIQEFNICFTTV</b>	788
tr A0A6J0I7H9 A0A6J0I7H9_9PASS	DEDIF-DQNKSKIGKKKAAAYAGGLVLDPKAGLYDKFILLDFN <b>SLYPSIIQEFNICFTTV</b>	916
tr A0A7K5Y128 A0A7K5Y128_9CHAR	DEDIF-DQNKSKIGKKKAAAYAGGLVLDPKAGLYDKFILLDFN <b>SLYPSIIQEFNICFTTV</b>	896
tr A0A7LOMV49 A0A7LOMV49_9PSIT	DEDIF-DQNKSKIGKKKAAAYAGGLVLDPKAGLYDKFILLDFN <b>SLYPSIIQEFNICFTTV</b>	895
:	*****	*****

tr A0A6P8GL87 A0A6P8GL87_CLUHA	QRVASTTHKKT	EDDDQDEIPEIPDPDLEMILPKEIRKLVERRRQVKQLMKQPDLPDL	947
tr A0A1S3NC92 A0A1S3NC92_SALSA	QREAHSRKKT	EEHDPDEIPEVPDADLEMILPKEIRKLVERRKQVKQLMKQAQDINPDL	971
tr A0A6P8T108 A0A6P8T108_GYMAC	QREAYNAKKNN	EEEDGSDIPEVPDQSLEMGILPKEIRKLVERRKQVKQLMKQQDLSNL	946
tr A0A3P9D223 A0A3P9D223_9CICH	QREAQNQSKKK	EVDDPEEIPEIPDPNLEMILPKEIRKLVERRKQVKQLMKQQDINPDL	814
tr A0A3P9MI28 A0A3P9MI28_ORYLA	QREAQNRSRKR	EEDESEEIPEIPDPGLEMGILPKEIRKLVERRKQVKQLMKQQDINPDL	938
tr A0A6J1UIRO A0A6J1UIRO_9SAUR	HRMSFSVHKKT	EDGD-----	906
tr A0A6P9D018 A0A6P9D018_PANGU	HRMSFSVHKKT	EDGEQEPIPELPDSLEMILPKEIKKLVERRHQVKQLMKQPDLPDL	949
sp P33609 DPOLA_MOUSE	QRVTSEVQKAT	EDEEQECIPELPDPNLEMILPKEIRKLVERRKQVKQLMKQQDLPDL	942
sp 089042 DPOLA_RAT	QRVASETLKAT	EDEEQECIPELPDPNLDGMILPREIRKLVERRKQVKQLMKQQDLPDL	945
tr A0A7J8AWU5 A0A7J8AWU5_RHIFE	QRVASEAQKV7	EGEEQECIPELPDPSLEMILPKEIRKLVERRKQVKQLMKQQDLPDL	937
tr G1T8Z5 G1T8Z5_RABIT	QRVVAEAKV7	EDGEQEPIPELPDPGLEMGILPREIRKLVERRKHVKQLMKQQDLPDL	942
tr HOUXU4 HOUXU4_CAVPO	QRVASELQKAA	EDGEEECIPELPDASLEMILPKEIRKLVERRKQVKQLMKQQDLPDL	944
tr G3RT59 G3RT59_GORGO	QRVASEAQKV7	EDGEQEPIPELPDPGLEMGILPREIRKLVERRKQVKQLMKQQDLPDL	945
sp P09884 DPOLA_HUMAN	QRVASEAQKV7	EDGEQEPIPELPDPGLEMGILPREIRKLVERRKQVKQLMKQQDLPDL	938
tr A0A6D2XUY8 A0A6D2XUY8_PANTR	QRVASEAQKV7	EDGEQEPIPELPDPGLEMGILPREIRKLVERRKQVKQLMKQQDLPDL	944
tr K7CUU4 K7CUU4_PANTR	QRVASEAQKV7	EDGEQEPIPELPDPGLEMGILPREIRKLVERRKQVKQLMKQQDLPDL	944
tr A0A6P6ITS8 A0A6P6ITS8_PUMCO	QRVASEAHKAT	EDGEQEPIPELPDPNLEMILPKEIRKLVERRKQVKQLMKQQDLPDL	944
tr A0A6J0A827 A0A6J0A827_ACIJB	QRVASEAHKAT	EDGEQEPIPELPDPNLEMILPKEIRKLVERRKQVKQLMKQQDLPDL	944
tr A0A6P4TPK1 A0A6P4TPK1_PANPR	QRVASEAHKAT	EDGEQEPIPELPDPNLEMILPKEIRKLVERRKQVKQLMKQQDLPDL	966
tr F6ZMX7 F6ZMX7_HORSE	QRVASEAQRT	EDGEQEPIPELPDPGLEMGILPREIRKLVERRKQVKQLMKQQDLPDL	945
tr A0A2Y9M558 A0A2Y9M558_DELLE	QRVASEAQKV7	EDGEQEPIPELPDPGLEMGILPREIRKLVERRKQVKQLMKQQDLPDL	944
tr A0A3Q1LMS8 A0A3Q1LMS8_BOVIN	QRVASEAQKV7	EDGEQEPIPELPDPGLEMGILPREIRKLVERRKQVKQLMKQQDLPDL	944
tr A0A452EXN8 A0A452EXN8_CAPI	QRVASEAQKV7	EDGEQEPIPELPDPGLEMGILPREIRKLVERRKQVKQLMKQQDLPDL	942
tr L8IUV3 L8IUV3_9CETA	QRVASEAQKV7	EDGEQEPIPELPDPGLEMGILPREIRKLVERRKQVKQLMKQQDLPDL	931
tr A0A452EXY8 A0A452EXY8_CAPI	QRVASEAQKV7	EDGEQEPIPELPDPGLEMGILPREIRKLVERRKQVKQLMKQQDLPDL	931
tr W5Q828 W5Q828_SHEEP	QRVASEAQKV7	EDGEQEPIPELPDPGLEMGILPREIRKLVERRKQVKQLMKQQDLPDL	944
tr A0A6P3STAR3 A0A6P3STAR3_SHEEP	QRVASEAQKV7	EDGEQEPIPELPDPGLEMGILPREIRKLVERRKQVKQLMKQQDLPDL	944
tr A0A1U7R9B9 A0A1U7R9B9_ALLSI	QRVSSDVQKRT	EEGEEEEIPELPDSSLEMILPKEIRKLVERRKQVKQLMKQPHLNPD	953
tr A0A151N032 A0A151N032_ALLMI	QRVSSDVQKRT	EEGEEEEIPELPDSSLEMILPKEIRKLVERRKQVKQLMKQPHLNPD	963
tr A0A663MCX2 A0A663MCX2_ATHCN	QLSSEAQKRA	EVEEEEEEIPELPDPGLEMGVLPEIRKLVERRKQVKQLMKQPDLPDL	847
tr A0A6J0I7H9 A0A6J0I7H9_9PASS	QLSSEAQKRA	EVEEEEEEIPELPDPGLEMGVLPEIRKLVERRKQVKQLMKQPDLPDL	975
tr A0A7K5Y128 A0A7K5Y128_9CHAR	QLSSEAQKRA	EVKEEEEEEIPELPDPGLEMGVLPEIRKLVERRKQVKQLMKQPDLPDL	955
tr A0A7L0MV49 A0A7L0MV49_9PSIT	QLSSEAQKRA	EVEEEEEEEIPELPDPGLEMGILPKEIKKLVERRKQVKQLMKQPDLPDL	954
:	:	*	

tr A0A6P8GL87 A0A6P8GL87_CLUHA	HMDYDIRQKALKI	TANSMYCCLGFSFSRFYAKPLAAMVTHKGRE-----ILL	994
tr A0A1S3NC92 A0A1S3NC92_SALSA	YLDYDIRQKALKI	TANSMYCCLGFSFSRFYAKPLAALVTHKGRE-----ILM	1018
tr A0A6P8T108 A0A6P8T108_GYMAC	YLDYDIRQKALKI	TANSMYCCLGFSFSRFYAKPLAALVTHKGRE-----ILM	993
tr A0A3P9D223 A0A3P9D223_9CICH	YLDYDIRQKALKI	TANSMYCCLGFSFSRFYAKPLAALVTHKGRE-----ILM	861
tr A0A3P9MI28 A0A3P9MI28_ORYLA	YMDYDIRQKALKI	TANSMYCCLGFSFSRFYAKPLAALVTHKGRE-----ILM	985
tr A0A6J1UIRO A0A6J1UIRO_9SAUR	--QDYDIRQKALKI	TANSMYCCLGFSFSRFYAKPLAALVTFKGRE-----ILL	951
tr A0A6P9D018 A0A6P9D018_PANGU	YLDYDIRQKALKI	TANSMYCCLGFSFSRFYAKPLAALVTFKGRE-----ILM	996
sp P33609 DPOLA_MOUSE	VLDYDIRQKALKI	TANSMYCCLGFSFSRFYAKPLAALVTFKGRE-----ILM	989
sp 089042 DPOLA_RAT	VLDYDIRQKALKI	TANSMYCCLGFSFSRFYAKPLAALVTFKGRE-----ILM	992
tr A0A7J8AWU5 A0A7J8AWU5_RHIFE	VLDYDIRQKALKI	TANSMYCCLGFSFSRFYAKPLAALVTFKGRE-----ILM	984
tr G1T8Z5 G1T8Z5_RABIT	VLDYDIRQKALKI	TANSMYCCLGFSFSRFYAKPLAALVTFKGREQTPTVIFVCLIM	1002
tr HOUXU4 HOUXU4_CAVPO	ALDYDIRQKALKI	TANSMYCCLGFSFSRFYAKPLAALVTFKGRE-----ILM	991
tr G3RT59 G3RT59_GORGO	ILDYDIRQKALKI	TANSMYCCLGFSFSRFYAKPLAALVTFKGRE-----	989
sp P09884 DPOLA_HUMAN	ILQYDIRQKALKI	TANSMYCCLGFSFSRFYAKPLAALVTFKGRE-----ILM	985
tr A0A6D2XUY8 A0A6D2XUY8_PANTR	ILQYDIRQKALKI	TANSMYCCLGFSFSRFYAKPLAALVTFKGRE-----ILM	991
tr K7CUU4 K7CUU4_PANTR	ILQYDIRQKALKI	TANSMYCCLGFSFSRFYAKPLAALVTFKGRE-----ILM	991
tr A0A6P6ITS8 A0A6P6ITS8_PUMCO	VLDYDIRQKALKI	TANSMYCCLGFSFSRFYAKPLAALVTFKGRE-----ILM	991
tr A0A6J0A827 A0A6J0A827_ACIJB	VLDYDIRQKALKI	TANSMYCCLGFSFSRFYAKPLAALVTFKGRE-----ILM	991
tr A0A6P4TPK1 A0A6P4TPK1_PANPR	VLDYDIRQKALKI	TANSMYCCLGFSFSRFYAKPLAALVTFKGRE-----ILM	1013
tr F6ZMX7 F6ZMX7_HORSE	VLDYDIRQKALKI	TANSMYCCLGFSFSRFYAKPLAALVTFKGRE-----ILM	992
tr A0A2Y9M558 A0A2Y9M558_DELLE	VLDYDIRQKALKI	TANSMYCCLGFSFSRFYAKPLAALVTFKGRE-----ILM	991
tr A0A3Q1LMS8 A0A3Q1LMS8_BOVIN	HLQYDIRQKALKI	TANSMYCCLGFSFSRFYAKPLAALVTFKGRE-----ILM	991
tr A0A452EXN8 A0A452EXN8_CAPI	HLQYDIRQKALKI	TANSMYCCLGFSFSRFYAKPLAALVTFKGRE-----ILM	989
tr L8IUV3 L8IUV3_9CETA	HLQYDIRQKALKI	TANSMYCCLGFSFSRFYAKPLAALVTFKGRE-----ILM	978
tr A0A452EXY8 A0A452EXY8_CAPI	HLQYDIRQKALKI	TANSMYCCLGFSFSRFYAKPLAALVTFKGRE-----ILM	978
tr W5Q828 W5Q828_SHEEP	HLQYDIRQKALKI	TANSMYCCLGFSFSRFYAKPLAALVTFKGRE-----ILM	991
tr A0A6P3STAR3 A0A6P3STAR3_SHEEP	HLQYDIRQKALKI	TANSMYCCLGFSFSRFYAKPLAALVTFKGRE-----ILM	991
tr A0A1U7R9B9 A0A1U7R9B9_ALLSI	HLQYDIRQKALKI	TANSMYCCLGFSFSRFYAKPLAALVTFKGRE-----ILM	991
tr A0A151N032 A0A151N032_ALLMI	HLQYDIRQKALKI	TANSMYCCLGFSFSRFYAKPLAALVTFKGRE-----ILM	991
tr A0A663MCX2 A0A663MCX2_ATHCN	HLQYDIRQKALKI	TANSMYCCLGFSFSRFYAKPLAALVTFKGRE-----ILM	1000
tr A0A6J0I7H9 A0A6J0I7H9_9PASS	YLQYDIRQKALKI	TANSMYCCLGFSFSRFYAKPLAALVTFKGRE-----ILM	1010
tr A0A7K5Y128 A0A7K5Y128_9CHAR	YLQYDIRQKALKI	TANSMYCCLGFSFSRFYAKPLAALVTFKGRE-----ILM	894
tr A0A7L0MV49 A0A7L0MV49_9PSIT	YLQYDIRQKALKI	TANSMYCCLGFSFSRFYAKPLAALVTFKGRE-----ILM	1022
tr A0A7L0MV49 A0A7L0MV49_9PSIT	YLQYDIRQKALKI	TANSMYCCLGFSFSRFYAKPLAALVTFKGRE-----ILM	1002
*****	*****	*****:*****:*****:*****:*****:*****	1001

tr A0A6P8GL87 A0A6P8GL87_CLUHA	HTKEMVQRMNLDV1  <b>YGDTDS</b>  IMINTNSTL LDEVFKLGNKVV-----	1034
tr A0A1S3NC92 A0A1S3NC92_SALSA	HTKEMVQKMNLLEV1  <b>YGDTDS</b>  IMINTNSS LEEVEVKLGNKV-----	1058
tr A0A6P8T108 A0A6P8T108_GYMAC	HTKEMVQKMNLLEV1  <b>YGDTDS</b>  IMINTNSK LEEVEVKLGNKV-----	1033
tr A0A3P9D223 A0A3P9D223_9CICH	HTKEMVQKMNLLEV1  <b>YGDTDS</b>  IMINTNSK LDEVFKLGNKV-----	901
tr A0A3P9M128 A0A3P9M128_ORYLA	HTKEMVQKMNLLEV1  <b>YGDTDS</b>  IMINTNSK LEEVEVKLGNKCVDTLQQEKLWFLNVEQTT	1045
tr A0A6J1UIR0 A0A6J1UIR0_9SAUR	HTKEMVQKMNLLEV1  <b>YGDTDS</b>  IMINTNLTH LEEVEVKLGNK1-----	991
tr A0A6P9D018 A0A6P9D018_PANGU	HTKEMVQKMNLLEV1  <b>YGDTDS</b>  IMINTNSTL LEEVEVKLGNK1-----	1036
sp P33609 DPOLA_MOUSE	HTKEMVQKMNLLEV1  <b>YGDTDS</b>  IMINTNSTL LEEVEVKLGNK1-----	1029
sp O89042 DPOLA_RAT	HTKEMVQKMNLLEV1  <b>YGDTDS</b>  IMINTNSTL LEEVEVKLGNK1-----	1032
tr A0A7J8AWU5 A0A7J8AWU5_RHIFE	HTKEMVQKMNLLEV1  <b>YGDTDS</b>  IMINTNSTL LEEVEVKLGNK1-----	1024
tr G1T8Z5 G1T8Z5_RABIT	HTKEMVQKMNLLEV1  <b>YGDTDS</b>  IMINTNSTL LEEVEVKLGNK1-----	1042
tr HOUXU4 HOUXU4_CAVPO	HTKEMVQKMNLLEV1  <b>YGDTDS</b>  IMINTNSTL LEEVEVKLGNK1-----	1031
tr G3RT59 G3RT59_GORGO	HTKEMVQKMNLLEV1  <b>YGDTDS</b>  IMINTNSTL LEEVEVKLGNK1-----	1021
sp P09884 DPOLA_HUMAN	HTKEMVQKMNLLEV1  <b>YGDTDS</b>  IMINTNSTL LEEVEVKLGNK1-----	1025
tr A0A6D2XUY8 A0A6D2XUY8_PANTR	HTKEMVQKMNLLEV1  <b>YGDTDS</b>  IMINTNSTL LEEVEVKLGNK1-----	1031
tr K7CUU4 K7CUU4_PANTR	HTKEMVQKMNLLEV1  <b>YGDTDS</b>  IMINTNSTL LEEVEVKLGNK1-----	1031
tr A0A6P6ITS8 A0A6P6ITS8_PUMCO	HTKEMVQKMNLLEV1  <b>YGDTDS</b>  IMINTNSTL LEEVEVKLGNK1-----	1031
tr A0A6J0A827 A0A6J0A827_ACIJB	HTKEMVQKMNLLEV1  <b>YGDTDS</b>  IMINTNSTL LEEVEVKLGNK1-----	1031
tr A0A6P4TPK1 A0A6P4TPK1_PANPR	HTKEMVQKMNLLEV1  <b>YGDTDS</b>  IMINTNSTL LEEVEVKLGNK1-----	1053
tr F6ZMX7 F6ZMX7_HORSE	HTKEMVQKMNLLEV1  <b>YGDTDS</b>  IMINTNSTL LEEVEVKLGNK1-----	1032
tr A0A2Y9M558 A0A2Y9M558_DELLE	HTKEMVQKMNLLEV1  <b>YGDTDS</b>  IMINTNSTL LEEVEVKLGNK1-----	1031
tr A0A3Q1LMS8 A0A3Q1LMS8_BOVIN	HTKEMVQKMNLLEV1  <b>YGDTDS</b>  IMINTNSTL LEEVEVKLGNK1-----	1031
tr A0A452EXN8 A0A452EXN8_CAPI	HTKEMVQKMNLLEV1  <b>YGDTDS</b>  IMINTNSTL LEEVEVKLGNK1-----	1029
tr L8IUV3 L8IUV3_9CETA	HTKEMVQKMNLLEV1  <b>YGDTDS</b>  IMINTNSTL LEEVEVKLGNK1-----	1018
tr A0A452EXY8 A0A452EXY8_CAPI	HTKEMVQKMNLLEV1  <b>YGDTDS</b>  IMINTNSTL LEEVEVKLGNK1-----	1018
tr W5Q828 W5Q828_SHEEP	HTKEMVQKMNLLEV1  <b>YGDTDS</b>  IMINTNSTL LEEVEVKLGNK1-----	1031
tr A0A6P3STAR3 A0A6P3STAR3_SHEEP	HTKEMVQKMNLLEV1  <b>YGDTDS</b>  IMINTNSTL LEEVEVKLGNK1-----	1031
tr A0AIU7R9B9 A0AIU7R9B9_ALLSI	HTKEMVQKMNLLEV1  <b>YGDTDS</b>  IMINTNSTL LEEVEVKLGNK1-----	1040
tr A0A151N032 A0A151N032_ALLMI	HTKEMVQKMNLLEV1  <b>YGDTDS</b>  IMINTNSTL LEEVEVKLGNK1-----	1050
tr A0A663MCX2 A0A663MCX2_ATHCN	HTKEMVQKMNLLEV1  <b>YGDTDS</b>  IMINTNSTL LDEVFKLGNK1-----	934
tr A0A6J0I7H9 A0A6J0I7H9_9PASS	HTKEMVQKMNLLEV1  <b>YGDTDS</b>  IMINTNSTL LDEVFKLGNK1-----	1062
tr A0A7K5Y128 A0A7K5Y128_9CHAR	HTKEMVQKMNLLEV1  <b>YGDTDS</b>  IMINTNSTL LDEVFKLGNK1-----	1042
tr A0A7L0MV49 A0A7L0MV49_9PSIT	HTKEMVQKMNLLEV1  <b>YGDTDS</b>  IMINTNSTL LDEVFKLGNK1-----	1041
	***:***:***** . :***:*****	
tr A0A6P8GL87 A0A6P8GL87_CLUHA	----- KSEVNKLKYKLL  <b>EID</b>  GVFKS  <b>LLLLKKKK</b>  YAAL MVVEHHGEGRFTTKQELKGKD-----	1087
tr A0A1S3NC92 A0A1S3NC92_SALSA	----- KAEVNKLKYKLL  <b>EID</b>  GVFKS  <b>LLLLKKKK</b>  YAAL VVEPLGGGRYSTKQELKGKD-----	1111
tr A0A6P8T108 A0A6P8T108_GYMAC	----- KAEVNKLKYKLL  <b>EID</b>  GVFKS  <b>LLLLKKKK</b>  YAAL VVENHGEGRSVKQELKGKD-----	1086
tr A0A3P9D223 A0A3P9D223_9CICH	----- KAEVNKLKYKLL  <b>EID</b>  GVFKS  <b>LLLLKKKK</b>  YAAL VVEQHGEGRFSVKQELKGKD-----	954
tr A0A3P9M128 A0A3P9M128_ORYLA	----- LLFCFQVKA EVNKLKYKLL  <b>EID</b>  GVFKS  <b>LLLLKKKK</b>  YAAL VVEHHGDGRYSVKQELKGKD-----	1105
tr A0A6J1UIR0 A0A6J1UIR0_9SAUR	----- KSEINKLKYKLL  <b>EID</b>  GVFKS  <b>LLLLKKKK</b>  YAAL VVEPAGDGKYITKQELKGKD-----	1044
tr A0A6P9D018 A0A6P9D018_PANGU	----- KSEVNKLKYKLL  <b>EID</b>  GVFKS  <b>LLLLKKKK</b>  YAAL VVEPAGDGKYITKQELKGKD-----	1089
sp P33609 DPOLA_MOUSE	----- KSEVNKLKYKLL  <b>EID</b>  GVFKS  <b>LLLLKKKK</b>  YAAL VVEPTSDGNYITKQELKGKD-----	1082
sp O89042 DPOLA_RAT	----- KSEVNKLKYKLL  <b>EID</b>  GVFKS  <b>LLLLKKKK</b>  YAAL VVEPTSDGNYITKQELKGKD-----	1085
tr A0A7J8AWU5 A0A7J8AWU5_RHIFE	----- KSEVNKLKYKLL  <b>EID</b>  GIFKS  <b>LLLLKKKK</b>  YAAL VVEPTSDGNYATKQELKGKD-----	1077
tr G1T8Z5 G1T8Z5_RABIT	----- KSEVNKLKYKLL  <b>EID</b>  GIFKS  <b>LLLLKKKK</b>  YAAL VVEPTSDGNYITKQELKGKD-----	1095
tr HOUXU4 HOUXU4_CAVPO	----- KSEVNKLKYKLL  <b>EID</b>  GIFKS  <b>LLLLKKKK</b>  YAAL VVEPASGDNVYTQKELKGKD-----	1084
tr G3RT59 G3RT59_GORGO	----- KSEVNKLKYKLL  <b>EID</b>  GIFKS  <b>LLLLKKKK</b>  YAAL VVEPTSDGNYVTQKELKGKD-----	1074
sp P09884 DPOLA_HUMAN	----- KSEVNKLKYKLL  <b>EID</b>  GIFKS  <b>LLLLKKKK</b>  YAALVVEPTSDGNYVTQKELKGKD-----	1078
tr A0A6D2XUY8 A0A6D2XUY8_PANTR	----- KSEVNKLKYKLL  <b>EID</b>  GIFKS  <b>LLLLKKKK</b>  YAAL VVEPTSDGNYVTQKELKGKD-----	1084
tr K7CUU4 K7CUU4_PANTR	----- KSEVNKLKYKLL  <b>EID</b>  GIFKS  <b>LLLLKKKK</b>  YAAL VVEPTSDGNYVTQKELKGKD-----	1084
tr A0A6P6ITS8 A0A6P6ITS8_PUMCO	----- KSEVNKLKYRLLE  <b>EID</b>  GIFKS  <b>LLLLKKKK</b>  YAAL VVEPTSDGNYITQKELKGKD-----	1084
tr A0A6J0A827 A0A6J0A827_ACIJB	----- KSEVNKLKYRLLE  <b>EID</b>  GIFKS  <b>LLLLKKKK</b>  YAAL VVEPTSDGNYITQKELKGKD-----	1084
tr A0A6P4TPK1 A0A6P4TPK1_PANPR	----- KSEVNKLKYRLLE  <b>EID</b>  GIFKS  <b>LLLLKKKK</b>  YAAL VVEPTLTDGNYITQKELKGKD-----	1106
tr F6ZMX7 F6ZMX7_HORSE	----- KSEVNRLYKLL  <b>EID</b>  GIFKS  <b>LLLLKKKK</b>  YAAL VVEPTSDGNYFTQKELKGKD-----	1085
tr A0A2Y9M558 A0A2Y9M558_DELLE	----- KSEVNRLYKLL  <b>EID</b>  GIFKS  <b>LLLLKKKK</b>  YAAL VVEPTVDGNYVTQKEVKGLD-----	1084
tr A0A3Q1LMS8 A0A3Q1LMS8_BOVIN	----- KSEVNRLYKLL  <b>EID</b>  GIFKS  <b>LLLLKKKK</b>  YAAL VVEPTSDGNYVTQKEVKGLD-----	1084
tr A0A452EXN8 A0A452EXN8_CAPI	----- KSEVNRLYKLL  <b>EID</b>  GIFKS  <b>LLLLKKKK</b>  YAAL VVEPTSDGNYVTQKEVKGLD-----	1082
tr L8IUV3 L8IUV3_9CETA	----- KSEVNRLYKLL  <b>EID</b>  GIFKS  <b>LLLLKKKK</b>  YAAL VVEPTSDGNYVTQKEVKGLD-----	1071
tr A0A452EXY8 A0A452EXY8_CAPI	----- KSEVNRLYKLL  <b>EID</b>  GIFKS  <b>LLLLKKKK</b>  YAAL VVEPTSDGNYVTQKEVKGLD-----	1071
tr W5Q828 W5Q828_SHEEP	----- KSEVNRLYKLL  <b>EID</b>  GIFKS  <b>LLLLKKKK</b>  YAAL VVEPTSDGNYVTQKEVKGLD-----	1084
tr A0A6P3STAR3 A0A6P3STAR3_SHEEP	----- KSEVNRLYKLL  <b>EID</b>  GIFKS  <b>LLLLKKKK</b>  YAAL VVEPTSDGNYVTQKEVKGLD-----	1084
tr A0AIU7R9B9 A0AIU7R9B9_ALLSI	----- KSEVNRLYKLL  <b>EID</b>  GIFKS  <b>LLLLKKKK</b>  YAAL VVEPTGDKGYITQKELKGKD-----	1093
tr A0A151N032 A0A151N032_ALLMI	----- KSEVNRLYKLL  <b>EID</b>  GIFKS  <b>LLLLKKKK</b>  YAAL VVEPTGDKGYITQKELKGKD-----	1103
tr A0A663MCX2 A0A663MCX2_ATHCN	----- KSEVNRLYKLL  <b>EID</b>  GIFKS  <b>LLLLKKKK</b>  YAAL VVEPTGDKGYITQKELKGKD-----	987
tr A0A6J0I7H9 A0A6J0I7H9_9PASS	----- KSEVNRLYKLL  <b>EID</b>  GIFKS  <b>LLLLKKKK</b>  YAAL VVEPTGDKGYITQKELKGKD-----	1115
tr A0A7K5Y128 A0A7K5Y128_9CHAR	----- KSEVNRLYKLL  <b>EID</b>  GIFKS  <b>LLLLKKKK</b>  YAAL VVEPTGDKGYITQKELKGKD-----	1095
tr A0A7L0MV49 A0A7L0MV49_9PSIT	----- KSEVNRLYKLL  <b>EID</b>  GIFKS  <b>LLLLKKKK</b>  YAAL VVEPTGDKGYITQKELKGKD-----	1094
	* *:***:***** . :***:*****	



//End of C-terminal of DNA pol  $\alpha$  from animals

tr AOA6P8GL87 AOA6P8GL87_CLUHA	DRTLLTSSSYSEVNLAKLFQAFALK-----	1470
tr AOA1S3NC92 AOA1S3NC92_SALSA	DRSLLASGYSEVNLSKLQFQAFTSLK-----	1494
tr AOA6P8T108 AOA6P8T108_GYMAC	DNALATSGYSEINLAKLFQAFSSLK-----	1469
tr AOA3P9D2Z3 AOA3P9D2Z3_9CICH	NTAMYLQTT--DTFSSIFSG-----	1326
tr AOA3P9MI28 AOA3P9MI28_ORYLA	EKALAASGYSEVNLSKLQFQAFSSLK-----	1487
tr AOA6J1UIRO AOA6J1UIRO_9SAUR	DKWLSMSSSYSEVNLGKLFQFTFSIVKSGDQSST-----	1432
tr AOA6P9D018 AOA6P9D018_PANGU	DKWLSMNSSYSEVNLGKLFQFTFSIVKSGDEESSTLKCLL	1482
sp P33609 DPOLA_MOUSE	EQFLSLWSGYSEVNLSKLQFQAFSSLK-----	1465
sp O89042 DPOLA_RAT	EHFLSLWSGG-----	1451
tr AOA7J8AWU5 AOA7J8AWU5_RHIFE	EHFLSQSGYTEVNLSKLQFADCAVRS-----	1461
tr G1T8Z5 RABIT	EQFLRLRSQGYSEVNLSKLQFADCAVKS-----	1479
tr HOUXU4 HOUXU4_CAVPO	EKFLSRSGYSEVNLSKLQFADCAMKS-----	1467
tr G3RT59 G3RT59_GORGO	EQFLSLRSQGYSEVNLSKLQFAGCAVKS-----	1458
sp P09884 DPOLA_HUMAN	EQFLSLRSQGYSEVNLSKLQFAGCAVKS-----	1462
tr AOA6D2XUY8 AOA6D2XUY8_PANTR	EQFLSLRSQGYSEVNLSKLQFAGCAVKS-----	1468
tr K7CUU4 K7CUU4_PANTR	EQFLSLRSQGYSEVNLSKLQFAGCAVKS-----	1468
tr AOA6P6ITS8 AOA6P6ITS8_PUMCO	EQFLSLRSQGYSEVNLSKLQFADCAVRP-----	1468
tr AOA6J0A827 AOA6J0A827_ACIJB	EQFLSLRSQGYSEVNLSKLQFADCAVRP-----	1468
tr AOA6P4TPK1 AOA6P4TPK1_PANPR	EQFLSLRSQGYSEVNLSKLQFADCAVRP-----	1490
tr F6ZMX7 F6ZMX7_HORSE	EQFLSLRSQGYSEVNLSKLQFADCAVKS-----	1469
tr AOA2Y9M558 AOA2Y9M558_DELLE	EQFLSLSQSGYSEVNLSKLQFADCAVRS-----	1468
tr AOA3Q1LMS8 AOA3Q1LMS8_BOVIN	-----	1424
tr AOA452EXN8 AOA452EXN8_CAPI	EQFLSLSQSGYSEVNLSRLFADCAVWS-----	1466
tr L8IUV3 L8IUV3_9CETA	EQFLSLSQSGYSEVNLSRLFADCAVWS-----	1455
tr AOA452EXY8 AOA452EXY8_CAPI	EQFLSLSQSGYSEVNLSRLFADCAVWS-----	1455
tr W5Q828 W5Q828_SHEEP	EQFLSLSQSGYSKVNLSRLFADCAVWS-----	1468
tr AOA6P3STAR3 AOA6P3STAR3_SHEEP	EQFLSLSQSGYSEVNLSRLFADCAVWS-----	1468
tr AOA1U7R9B9 AOA1U7R9B9_ALLSI	DTFLSMSSSYSEVNLGKLFQFTICTAKSGVDASNERQ-----	1488
tr AOA151N032 AOA151N032_ALLMI	DTFLSMSSSYSEVNLGKLFQFTICTAKVGVDASNERQ-----	1498
tr AOA663MCX2 AOA663MCX2_ATHCN	-----STINIEETFIR-EKMMQAERH-----	1355
tr AOA6J0I17H9 AOA6J0I17H9_9PASS	-----	1422
tr AOA7K5Y128 AOA7K5Y128_9CHAR	DKCLSTSGYSEVNGLGKLFVTs-IGRAVGESENRRQ-----	1489
tr AOA7LOMV49 AOA7LOMV49_9PSIT	DKCLSMMSGYSEVNGLGKLFVTs-IGRSVGESENRRQ-----	1487

Fig. 11 MSA of DNA polymerase  $\alpha$  catalytic subunit from animal sources

*AOA6P8GL87\_CLUHA*, *Clupea harengus*  
*AOA6P8T108\_GYMAC*, *Gymnodraco acuticeps*  
*AOA3P9MI28\_ORYLA*, *Oryzias latipes*  
*AOA6J0I7H9\_9PASS*, *Lepidothrix coronata*  
*P33609/DPOLA\_MOUSE*, *Mus musculus*  
*AOA7J8AWU5\_RHIFE*, *Rhinolophus ferrumequinum*  
*HOUXU4\_CAVPO*, *Cavia porcellus*  
*P09884/DPOLA\_HUMAN*, *Homo sapiens*  
*K7CUU4\_PANTR*, *Pan troglodytes*  
*AOA6J0A827\_ACIJB*, *Acinonyx jubatus*  
*F6ZMX7\_HORSE*, *Equus caballus*  
*AOA3Q1LMS8\_BOVIN*, *Bos tauru*  
*AOA452EXY8\_CAPI*, *Capra hircus*  
*W5Q828\_SHEEP*, *Ovis aries*  
*AOA1U7R9B9\_ALLSI*, *Alligator sinensis*  
*AOA663MCX2\_ATHCN*, *Athene cunicularia*  
*AOA7LOMV49\_9PSIT*, *Amazona guildingii*  
  
*AOA1S3NC92\_SALSA*, *Salmo salar*  
*AOA3P9D2Z3\_9CICH*, *Maylandia zebra*  
*AOA6J1UIRO\_9SAUR*, *Notechis scutatus*  
*AOA6P9D018\_PANGU*, *Pantherophis guttatus*  
*O89042/DPOLA\_RAT*, *Rattus norvegicus*  
*G1T8Z5\_RABIT*, *Oryctolagus cun*  
*G3RT59\_GORGO*, *Gorilla gorilla gorilla*  
*AOA6D2XUY8\_PANTR*, *Pan troglodytes*  
*AOA6P6ITS8\_PUMCO*, *Puma concolor*  
*AOA6P4TPK1\_PANPR*, *Panthera pardus*  
*AOA2Y9M558\_DELLE*, *Delphinapterus leucas*  
*AOA452EXN8\_CAPI*, *Capra hircus*  
*L8IUV3\_9CETA*, *Bos mutus*  
*AOA6P3STAR3\_SHEEP*, *Ovis aries*  
*AOA151N032\_ALLMI*, *Alligator mississippiensis*  
*AOA7K5Y128\_9CHAR*, *Dromas ardeola*

Figure 12 shows the MSA analysis of DNA pol  $\alpha$  from different yeasts. The N-terminal and C-terminal regions are not conserved significantly and the N-terminal region showed lots of gaps in alignment as it is not conserved (data not shown). However, after ~250 amino acids from the N-terminal, a stretch of highly conserved acidic amino acids is seen (highlighted in red) as found in animal sources, suggesting a possible role in binding to the basic histone proteins of the chromosomes. After the middle half, few conserved peptides are observed (data not shown). The invariant -SLYPS- is found around after ~850 amino acids, as found in the DNA pol  $\alpha$  from animal sources (Fig. 11). It has the typical temple-binding -YG- pair and the catalytic amino acid

-KL/v- (highlighted in yellow). Three ZBM are also found but placed apart; one in the middle and two at the C-terminal regions (highlighted in orange) and, however, the structure of ZBM, CysA and CysB are significantly different from animal sources. A highly conserved peptide with consecutive L and K residues -LLLHAKKK- is also observed as in the DNA pol  $\alpha$  from animals (Fig. 11). The two-in-one site, -YGDTS- and the Mg<sup>2+</sup>-binding (-DID-) are also found as in the DNA poly  $\alpha$  from animal sources (Fig. 11). Even though the PR activity is reported to be absent. [28] The proposed PR active site amino acids are highlighted in light blue and compared with the other PR active sites from bacteria and viruses (Table 4).

## CLUSTAL O 1.2.4) MSA analysis DNA polymerase α from different yeasts

tr AOA7D9GZJ1 AOA7D9GZJ1_DEKBR	ADADLDDDDDDDE	IVVAKR-RTTAFVDRD-V-NLSSSKAKPI-LKHPLNNSPFKDT--DV	328
tr AOA1E3NHX3 AOA1E3NHX3_9ASCO	NDFDD-DDDEDD	IIISKRKKVSAIDRK-M-KAVD-----SSPLNAKQL---DI	313
tr R9XN45 R9XN45_ASHAC	----ADDSDDEE	VYLRKAIRESAASRH-V-TFGAKFK--SVSSPAGKSIDIPHSDP	283
tr I6NDM1 I6NDM1_ERECY	SNASFDDDDSDDE	IQFGRRTVKTASDRK-V-TFTTAVP--NSSSPH----VIAHGT	312
tr AOA109UYKO AOA109UYKO_9SACH	KAEDFLDDDSDE	IQFTRPRVRSVAVNRE-V-TFHSTSA--ASQI-----VPPPSE	295
tr AOA1G4J8L9 AOA1G4J8L9_9SACH	QSESESEESEE	VKATRRTVRTAIRRA-V-NFNIGQD--AQSSPF----VTAPNTP	311
tr AOA0P1KX40 AOA0P1KX40_9SACH	-GGQSSEEDSDE	ITIRRRTIRSAASQR-A-NVDVRSS--VQSSPF----VTAPPTP	302
tr AOA1G4KH46 AOA1G4KH46_9SACH	NIDGDDDDSDDE	IQLGRRTIRTAAAARK-A-NFNMSI--TPSSPF----TTAPGTP	311
tr AOA1G4KMB7 AOA1G4KMB7_9SACH	LSDKDDEESEDE	INIGRRTVRNAAAIKRN--NINAGSNS--SSSSPF----VTAPGTP	321
tr AOA1S7HLF8 AOA1S7HLF8_9SACH	EICSTIEDDSDE	ITFGKRTVRVVAIRQ-V-NLSSKSN--PLASPV----VA-PGTP	279
tr AOA1Q3A9M9 AOA1Q3A9M9_ZYGR	ESNDNNDSEDD	IFVGRRTVRVRAAKRE-V-NLSSKSN--PLASPV----VA-PGTP	290
tr AOA5P2U3L1 AOA5P2U3L1_KLULC	SED--ESEDDEDE	IHFGRKKFKSIAAQQRQ-M-NISAQ-TAPLTSPTY----VTAPNTP	277
tr WOTGP7 WOTGP7_KLUMD	SVEDDEDDDEDE	IHFNRKVKVSAAQRQ-L-NIHAKAAPPVTSPTY----VTAPNTP	281
<b>sp P13382 DPOA_YEAST</b>	LADEDDEDDSDDE	LADEDDEDDSDDEIILKRTKMRSTVTTTRK-V-NIDSRSN--PSTSFP----VTAPGTP	314
tr AOA0L8VIK5 AOA0L8VIK5_9SACH	LADEDEDDSDDE	IILKRTKMRSTVTTTRK-V-NIDSRSN--PSTSFP----VTAPGTP	314
tr AOA7G7JEL9 AOA7G7JEL9_CANGB	STNIDSDDSDDE	VIVTKRMRTATAVRK-V-NLDKSA--LSSSPY----VTPGTP	297
tr AOA1B2JAV2 AOA1B2JAV2_PICPA	SLDDIQDEESDE	LLISRARTANTEISREINMRASVIEEK----T-----ADADP	279
tr C4R232 C4R232_KOMPG	SLDQTQDEESDE	LLISRARTANTEISREVNMRASVIEEK----A-----ADADS	280
tr AOA642UJM6 AOA642UJM6_DIURU	PATQESESSDDE	TV-VVRPRAAAVAPRKPMSTISSIKANV-----LQSSP	274
tr AOA1E4U1L7 AOA1E4U1L7_PACTA	DNEEEEDDDDDDE	IIVSRRPRAATSRVDRSV-NLAATKLP----TVKS--THLNSSP	338
tr AOA1V2L9F2 AOA1V2L9F2_CYBFA	TMDVDEEESDEDE	IVVVKRPRVSTAVNR-SVNLTAAKANE----PKEE--RQISSPP	289
tr AOA2Z1Y943 AOA2Z1Y943_9ASCO	FDDDEDVSSSEDE	IVVTRKPRAS-GAVRKTGLGVVSSVKLTP-----SAASI	293
tr AOA1D8PK28 AOA1D8PK28_CANAL	VNDNNDEDDSDDE	IVYSRKPRAA-VVKKNQVDTVSAVKANQ-----LPSSP	301
tr C4YQI4 C4YQI4_CANAW	VNDNDEDDSDDE	VNDNDEDDSDDEIVYSRKPRAA-VVKKNQVDTVSAVKANQ-----LPSSP	301
tr G8Y7L3 G8Y7L3_PICSO	EKEDASSDSDDE	EKEDASSDSDDEIVVTRPRAV-GKRNRRNDVNFTSNETK-----IKTNA	281
tr AOA0V1Q5J4 AOA0V1Q5J4_9ASCO	EDEDEIIDDSEDE	EDEDEIIDDSEDEIVVTRPRAA-SANKHKAANFSAVKASN-----IASSS	299
tr ASDC55 ASDC55_PICGU	EPKSQPESDSEDE	EPKSQPESDSEDEVVVARRPKSA--VSRTTMSTLTAVRTSE-----LPSSS	271
tr AOA512UAU0 AOA512UAU0_9ASCO	EPVSEDDSESED	EPVSEDDSESEDIVMVRQRRT-TVAAPKTAANISVKTAS-----LPAKS	305
tr AOA1L0D0V4 AOA1L0D0V4_9ASCO	-LKEESESD-DE	-LKEESESD-DEVVVRPRP-SAPVRETATISTVKAAD-----L---	280
tr AOA0L0NRD1 AOA0L0NRD1_CANAR	EPASESDSED-DE	EPASESDSED-DEIVVAKRRAAAAPARNTGATISAVKASG-----L---	278
... : :			

Zn1	QGRLLCDIANEMGQSLTPK	QSWDLSEEMYQVTC	KKYKTPMEVNLANPAFGEDAGRLLAAF	793
	AGRLLCDIANEMGQSLTPK	QSWELABEMYEVV	CDKYSPEMVNLANPLVAEDANKLLAAI	775
	SGRLLCDISNEMGQSLTPK	QSWELGEEMYQVQ	QREYKPLEINFNDARYHEDANFLVLAL	720
	SGRLLCDISNEMGQSLTPK	QNWELTEEMYQVA	QKDHKAIEINFNDVRFQEDVSLFVLAL	750
	SGRLLCDISNEMGQSLTPK	QSWELGEEMYQVTC	QKDHFLDWDYSDAQYHDNANCILVLA	733
	SGRLLCDISNEMGQSLTPK	QNWELPEMYQVTC	QKDHKLSDINFQASQYQSDVNAFILAL	753
	SGRLLCDISNEMGQSLTPK	QSWDLSEEMYQVTC	QKEQTLEINYQNPQYQNDVNSMIMAL	734
	AGRLLCDIANEMGQSLTPK	QSWDLSEEMYQVTC	QKEHKSLEVNYQSPQYQNDVNSMVLAL	744
	GGRLLCDISNEMGQSLTPK	QSWELABEMYQVTC	QKEHKAMEINYQSPQYQSDVNSMILAL	754
	SGRLLCDIANDMGQSLTPK	QTWDLSEEMYEVTC	QKEHKSLDINYQNPQYQDAETMTMAL	718
	AGRLLCDIANEMGQSLTPK	QTWDLPEMYLVA	QKETTPLDINYQNPHYREDADTMAMAL	728
	AGRLLCDIANEMGQSLTPK	QNWELNEQYQVSG	GIERKSLEINYLPNIYQEDESQMILLAL	711
	AGRLLCDIANEMGQSLTPK	QNWELNEQYQVSG	GIERKSLEINYLSP1YQEDESQMILAL	715
	SGRLLCDIANEMGQSLTPK	QSWDLSEEMYQVTC	EKEHKPLDIDYQNPQYQNDVNSMTMAL	751
	SGRLLCDIANEMGQSLTPK	QSWDLSEEMYQVTC	EKEHKPLDIDYQNPQYQNDVNSMTMAL	751
	AGRLLCDIANEMGQSLTPK	QNWDLABEMYQVTC	EKEHKPLDIDYQNPQYQNDVNSMTMAL	737
	QGRLLCDIGNDMGQSLTTK	QSWDLPEMYQSVG	QVDPHKEPLEINYQNTQYQDGVNLLTMAL	717
	HGRLLCDIGNDMGQSLTTK	QSWDLPEMYQSVG	QGKFKPSEIALNNPQLSENVNLLLTVA	718
	AGRLLCDIAGDLGRSLSTK	QSWLSEMYRVVG	QGKQYAGMEVNWNAMS--NPDYFMTVL	694
	QGRLLCDIANEMGQSLTTK	QSWDLSEMYQVTC	QKWLTDKTSIDEINLSNPKILEDANTILMINI	791
	AGRLLCDIANEMGQSLTTK	QSWELHEMYEVV	QKWLTDKTSIDEINLSNPKILEDANTILMINI	732
	QGRLLCDIANEMGQSLTPK	QSWDLVEMYDIVC	QVDPHKEPLEINYQEPKYAEADASFMMAL	732
	QGRLLCDIANEMGQSLTPK	QSWDLHEMYDVVG	QKWLTDKTSIDEINLSNPKILEDANTILMINI	741
	QGRLLCDIANEMGQSLTPK	QSWDLHEMYDVVG	QVDPHKEPLEINYQEPKYAEADASFMMAL	741
	QGRLLCDIANEMGQSLTPK	QSWDLHEMYDVVG	QKWLTDKTSIDEINLSNPKILEDANTILMINI	724
	QGRLLCDIANEMGQSLTPK	QSWDLNEMYEVV	QVDPHKEPLEINYQEPKYAEADASFMMAL	745
	QGRLLCDIANEMGQSLTTK	QSWDLSEMLDVG	QVDPHKEPLEINYQEPKYAEADASFMMAL	712
	QGRLLCDIANELGQSLTPK	QSWDLPEMYEVV	QVDPHKEPLEINYQEPKYAEADASFMMAL	747
	QGRLLCDIANELGQSLTPK	QSWDLNEMYEVV	QVDPHKEPLEINYQEPKYAEADASFMMAL	719
	QGRLLCDIANELGQSLTTK	QSWDLNEMYEVV	QVDPHKEPLEINYQEPKYAEADASFMMAL	718
*** : * : . : * : *** : * : . : * : * : :				

tr A0A7D9GZJ1 A0A7D9GZJ1_DEKBR	MDFN SLYPS  IQE FNICFTTVDRSKLTE----DELP EPVSPSPEIPRGVLPKLLHELVSR	957
tr A0A1E3NHX3 A0A1E3NHX3_9ASCO	MDFN SLYPS  IQE FNICFTTIERNLGE----DELP KVPTK-SAMGVLPKLLQLQVLVTR	947
tr R9XN45 R9XN45_ASHAC	MDFN SLYPS  IQE FNICFTTVERNPHD----DEL PPVTE-AKAGVLPRLLANLVERR	891
tr I6NDM1 I6NDM1_ERECY	MDFN SLYPS  IQE FNICFTTVERDPGD----EQLE PEPVPSGIVQGVLPKLLANLVQRR	922
tr A0A109YUK0 A0A109YUK0_9SACH	MDFN SLYPS  IQE YNICFTTVDRNPN----DELP VPVSSDMAQGVLPKLLANLVQRR	905
tr A0A1G4J8L9 A0A1G4J8L9_9SACH	MDFN SLYPS  IQE FNICFTTVDRDPNN----EELP VPVSTEEAQGVLPKLLANLVQRR	922
tr A0A0P1KX40 A0A0P1KX40_9SACH	MDFN SLYPS  IQE YNICFTTVERDPNSN----DEL DPVPSSEMSQGVLPRLANLVQKR	906
tr A0A1G4KH46 A0A1G4KH46_9SACH	MDFN SLYPS  IQE FNICFTTVDRDPNN----EELP VPVPSDSKQGVLPRLANLVDRR	916
tr A0A1G4KMB7 A0A1G4KMB7_9SACH	MDFN SLYPS  IQE YNICFTTVERDPNN----EELP VPVPSDEIKQGVLPRLANLVQRR	926
tr A0A1S7HLF8 A0A1S7HLF8_9SACH	MDFN SLYPS  IQE FNICFTTVDRRNLED----DEL PKVPSGDEDYGVLPRLANLVQRR	890
tr A0A1Q3A9M9 A0A1Q3A9M9_ZYGR0	MDFN SLYPS  IQE FNICFTTVDRNL----DELP EPVPSNGSMGVLPRLANLVQRR	901
tr A0A5P2U3L1 A0A5P2U3L1_KULUC	MDFN SLYPS  IQE FNICFTTVDRNPQDI----EELP DVPSREKSQGVLPRLANLVQRR	877
tr WOTGP7 WOTGP7_KLUMD	MDFN SLYPS  IQE FNICFTTVDRPNNI----EELP DVPSQDNPGVLPRLANLVQRR	881
sp P13382 DPOA_YEAST	MDFN SLYPS  IQE FNICFTTVDRNKED----DEL PSVPSEVDQGVLPRLANLVDR	917 SMDM
tr A0A0L8VIK5 A0A0L8VIK5_9SACH	MDFN SLYPS  IQE FNICFTTVDRNKED----DEL PSVPSEVDQGVLPRLANLVDR	917
tr A0A7G7JEL9 A0A7G7JEL9_CANGB	MDFN SLYPS  IQE FNICFTTVDRDVENI----DELP EPVSPSADKAQGVLPRLANLVERR	903
tr A0A1B2JAV2 A0A1B2JAV2_PICPA	LDFN SLYPS  IQE FNICFTTVERESGNV----DTIPP VPVNQIQQGVLPKLNLNVSR	888
tr C4R232 C4R232_KOMP	LDFN SLYPS  IQE FNICFTTVERESGNV----DTIPP VPVNQNSQQGVLPRLANLNVSR	889
tr A0A642UJM6 A0A642UJM6_DIURU	MDFN SLYPS  IQE FNICFTTVARDNDYNRTHDEQNDMFSYPERDQNGFLPKLNLTVSR	868
tr A0A1E4U1L7 A0A1E4U1L7_PACTA	MDFN SLYPS  IQE FNICFTTVNRSLSNQ----DEIPS VPVPLGIEQGVLPRLLATLNVRR	964
tr A0A1V2L9F2 A0A1V2L9F2_CYBFA	MDFN SLYPS  IQE FNICFTTVRDASDI----DALPE PVATSVSQGVLPRLNLTVTR	907
tr A0A2Z1Y943 A0A2Z1Y943_9ASCO	MDFN SLYPS  IQE FNICFTTIERTQFNITHDEHDHMPOLPEKDLMNGVLPKLLNLSVSR	909
tr A0A1D8PK28 A0A1D8PK28_CANAL	MDFN SLYPS  IQE FNICFTTVNRDRFVNTHDENKDMPILPERDTEGVLPRLNLTVSR	918
tr C4YQI4 C4YQI4_CANAW	MDFN SLYPS  IQE FNICFTTVNRDRFVNTHDENKDMPILPERDTEGVLPRLNLTVSR	918
tr G8Y7L3 G8Y7L3_PICSO	MDFN SLYPS  IQE FNICFTTVDRDQNYNLTD EDRDLFLPDGDSQQGVLPRLNLTVSR	901
tr A0A0V1Q5J4 A0A0V1Q5J4_9ASCO	MDFN SLYPS  IQE YNICFTTVDRDNYNAAKIDEGLP DIPD PRDVAQGVLPRLNALVSR	922
tr A5DC55 A5DC55_PICGU	MDFN SLYPS  IQE YNICFTTVDRDFNVSHDEEKDMFVLPDRSEVQGVLPRLNLTVSR	889
tr A0A512UAU0 A0A512UAU0_9ASCO	MDFN SLYPS  IQE FNICFTTVMRDNYNITHDEDKDMP EIPERDS DAGVLPRLLNLTVSR	924
tr A0A1L0D0V4 A0A1L0D0V4_9ASCO	MDFN SLYPS  IQE FNICFTTVERDSYNI SHDEERDLPSI PERDSEAGVLPRLLNLTVSR	896
tr A0A0L0NRD1 A0A0L0NRD1_CANAR	MDFN SLYPS  IQE FNICFTTVERTAYNI SHDEERDLPSI PERDSDAGVLPRLLNALVSR	895

tr A0A7D9GJZ1 A0A7D9GJZ1_DEKBR	REVKKLLKDPRISTQERIYDIDRQLAKTANSMYCCLGYANSRFYAKPLAMLVTNKGRE	1017
tr A0A1E3NHX3 A0A1E3NHX3_9ASCO	REVKSLSMKDPKLSAIEKAQLDICKOMALKTANSMYCCLGYVNSRFYAKPLAMLVTNKGRE	1007
tr R9XN45 R9XN45_ASHAC	REVKKLMNAE-ADPHKKAECDIRQQAALKLTANSMYCCLGYVNSRFYAKPLAMLVTNKGRE	950
tr I6NDM1 I6NDM1_ERECY	REVKKLMKTE-IDPKQKAECDIRQQAALKLTANSMYCCLGYVNSRFYAKPLAMLVTNKGRE	981
tr A0A109UYKO A0A109UYKO_9SACH	YEVKKLMKTE-TDPHRAECDIRQQAALKLTANSMYCCLGYVNSRFYAKPLAMLVTNKGRE	964
tr A0A1G4J8L9 A0A1G4J8L9_9SACH	REVKKIMSTE-TNLEKKTQCDIRQQAALKLTANSMYCCLGYVNSRFYAKPLAMLVTNKGRE	981
tr A0A0P1KX40 A0A0P1KX40_9SACH	REVKKLLKTE-SDPHKRTQCDIRQQAALKLTANSMYCCLGYVNSRFYAKPLAMLVTNKGRE	965
tr A0A1G4KH46 A0A1G4KH46_9SACH	REVKKLLKAE-TDPHKKRVQCDIRQQAALKLTANSMYCCLGYVNSRFYAKPLAMLVTNKGRE	975
tr A0A1G4KMB7 A0A1G4KMB7_9SACH	REVKKLLKTE-SDPHKRVQCDIRQQAALKLTANSMYCCLGYVNSRFYAKPLAMLVTNKGRE	985
tr A0A1S7HLF8 A0A1S7HLF8_9SACH	REVKRIMATE-TDPHKTQCDIRQQAALKLTANSMYCCLGYVNSRFYAKPLAMLVTNKGRE	949
tr A0A1Q3A9M9 A0A1Q3A9M9_ZYGR0	REVKKIMTE-KDPHKKRVQCDIRQQAALKLTANSMYCCLGYVNSRFYAKPLAMLVTNKGRE	960
tr A0A5P2U3L1 A0A5P2U3L1_KLULC	REVKNLMIKTE-NDASKAQAYDIDRQLAKLTANSMYCCLGYVNSRFYAKPLAMLVTNKGRE	936
tr W0TGP7 W0TGP7_KLUMD	REVKNLMKNE-NDPAKRAQYDIDRQLAKLTANSMYCCLGYVNSRFYAKPLAMLVTNKGRE	940
sp P13382 DPOA_YEAST	REVKKVMKTE-TDPHKKRVQCDIRQQAALKLTANSMYCCLGYVNSRFYAKPLAMLVTNKGRE	976
tr A0A0L8V1K5 A0A0L8V1K5_9SACH	REVKKVMKTE-TDPHKKRVQCDIRQQAALKLTANSMYCCLGYVNSRFYAKPLAMLVTNKGRE	976
tr A0A7G7JEL9 A0A7G7JEL9_CANGB	REVKKIMKTE-TDPHKKRVQCDIRQQAALKLTANSMYCCLGYVNSRFYAKPLAMLVTNKGRE	962
tr A0A1B2JAV2 A0A1B2JAV2_PICPA	REVKKLLKDPKIGSIEKAQFDIKOQALKLTANSMYCCLGYVNSRFYARPLAMLVT_SKGRE	948
tr C4R232 C4R232_KOMP_G	REVKKLLKDPSK1GVEKAQFDIKOQALKLTANSMYCCLGYVNSRFYARPLAMLVT_SKGRE	949
tr A0A642UJM6 A0A642UJM6_DIURU	REVKKLLKDPLTPQQRAQFDWKOTALKLTANSMYCCLGYVQSRFYAKPLAMLVTNKGRE	928
tr A0A1E4U1L7 A0A1E4U1L7_PACTA	REVKKLLKDKSTQAEKIYDIDRQLAKLTANSMYCCLGYVNSRFYAKPLAMLVTNKGRE	1024
tr A0A1V2L9F2 A0A1V2L9F2_CYBFA	REVKKLLKDPSTSPIERAQYDIDRQLAKLTANSMYCCLGYVNSRFYAKPLAMLVTNKGRE	967
tr A0A2Z1Y943 A0A2Z1Y943_9ASCO	REVKKLLKAPSLTNFRQAFDIDRQLAKLTANSMYCCLGYVNSRFYAKPLAMLVTNKGRE	969
tr A0A1D8PK28 A0A1D8PK28_CANAL	REVKKLLKDPNTPFQKAOYDIDRQLAKLTANSMYCCLGYVNSRFYAKPLAMLVTNKGRE	978
tr C4YQI4 C4YQI4_CANAW	REVKKLLKDPKNTPFQKAOYDIDRQLAKLTANSMYCCLGYVNSRFYAKPLAMLVTNKGRE	978
tr G8Y7L3 G8Y7L3_PICSO	REVKKLLKDPKITSEKAQYDIDRQLAKLTANSMYCCLGYVNSRFYAKPLAMLVTNKGRE	961
tr A0A0V1Q5J4 A0A0V1Q5J4_9ASCO	REVKKLLKDPKNTAERAQYDIDRQLAKLTANSMYCCLGYVNSRFYAKPLAMLVTNKGRE	982
tr A5DC55 A5DC55_PICGU	REVKKLLKDPKNSLVERAQYDIDRQLAKLTANSMYCCLGYVNSRFYAKPLAMLVTNKGRE	949
tr A0A512UAU0 A0A512UAU0_9ASCO	REVKKLLKDSRSISSEERAQYDIDRQLAKLTANSMYCCLGYVNSRFYAKPLAMLVTNKGRE	984
tr A0A1L0D0V4 A0A1L0D0V4_9ASCO	REVKKLLKPNPKITPEERAQYDIDRQLAKLTANSMYCCLGYVNSRFYAKPLAMLVTNKGRE	956
tr A0A0L0NRD1 A0A0L0NRD1_CANAR	REVKKLLKPNPKNTPHERAQYDIDRQLAKLTANSMYCCLGYVNSRFYAKPLAMLVTNKGRE	955

tr A0A7D9GZJ1 A0A7D9GZJ1_DEKBR	ILMDTRHLAESLGMRVVYGDTSVMVDTGCDTMKEALKFAEDFRKRVNDRYKLLEIDDN	1077
tr A0A1E3NHX3 A0A1E3NHX3_9ASCO	ILLDTRQLAESELLVYGDTSVMIDTRCDSFAEAIIKGNGFKEKVNERYLLEIDDN	1067
tr R9XN45 R9XN45_ASHAC	ILMNTRQLAESELLVYGDTSVMIDTGCNDYKEALNIGENFKKLVNERYLLEIDDN	1010
tr I6NDM1 I6NDM1_ERECY	ILMNTRQLAESELLVYGDTSVMIDTGCNDYKESVNIGENFKKLVNERYLLEIDDN	1041
tr A0A109UYK0 A0A109UYK0_9SACH	ILMNTRQLAESELLVYGDTSVMIDTGCNDYKDAFNIGENFKKLVNERYLLEIDDN	1024
tr A0A1G4J8L9 A0A1G4J8L9_9SACH	ILMNTRQLAESELLVYGDTSVMIDTGCNDYKDAFNIGENFKKLVNERYLLEIDDN	1041
tr A0A0P1KX40 A0A0P1KX40_9SACH	ILMNTRQLAESELLVYGDTSVMIDTGCNDYKDAFNIGENFKKLVNERYLLEIDDN	1025
tr A0A1G4KH46 A0A1G4KH46_9SACH	ILMNTRQLAESELLVYGDTSVMIDTGCNDYKDAFNIGENFKKLVNERYLLEIDDN	1035
tr A0A1G4KMB7 A0A1G4KMB7_9SACH	ILMNTRQLAESELLVYGDTSVMIDTGCNDYKDAFNIGENFKKLVNERYLLEIDDN	1045
tr A0A1S7HFL8 A0A1S7HFL8_9SACH	ILMNTRQLAESELLVYGDTSVMIDTGCNDYKDAFNIGENFKKLVNERYLLEIDDN	1009
tr A0A1Q3A9M9 A0A1Q3A9M9_ZYGR	ILMNTRQLAESELLVYGDTSVMIDTGCNDYKDAFNIGENFKKLVNERYLLEIDDN	1020
tr A0A5P2U3L1 A0A5P2U3L1_KLULC	ILMNTRQLAESELLVYGDTSVMIDTGCNDYKDAFNIGENFKKLVNERYLLEIDDN	996
tr WOTGP7 WOTGP7_KLUMD	ILMNTRQLAESELLVYGDTSVMIDTGCNDYKDAFNIGENFKKLVNERYLLEIDDN	1000
sp P13382 DPOA YEAST	ILMNTRQLAESELLVYGDTSVMIDTGCNDYKDAFNIGENFKKLVNERYLLEIDDN	1036_SDM
tr A0A0L8VIK5 A0A0L8VIK5_9SACH	ILMNTRQLAESELLVYGDTSVMIDTGCNDYKDAFNIGENFKKLVNERYLLEIDDN	1036
tr A0A7G7JEL9 A0A7G7JEL9_CANGB	ILMNTRQLAESELLVYGDTSVMIDTGCNDYKDAFNIGENFKKLVNERYLLEIDDN	1022
tr A0A1B2JAV2 A0A1B2JAV2_PICPA	ILMDTRQLAESELLVYGDTSVMIDTGCNDYKDAFNIGENFKKLVNERYLLEIDDN	1008
tr C4R232 C4R232_KOMPG	ILMDTRQLAESELLVYGDTSVMIDTGCNDYKDAFNIGENFKKLVNERYLLEIDDN	1009
tr A0A642UJM6 A0A642UJM6_DIURU	ILMDTRQLAESELLVYGDTSVMIDTGCNDYKDAFNIGENFKKLVNERYLLEIDDN	988
tr A0A1E4U1L7 A0A1E4U1L7_PACTA	ILMDTRQLAESELLVYGDTSVMIDTGCNDYKDAFNIGENFKKLVNERYLLEIDDN	1084
tr A0A1V2L9F2 A0A1V2L9F2_CYBFA	ILMDTRQLAESELLVYGDTSVMIDTGCNDYKDAFNIGENFKKLVNERYLLEIDDN	1027
tr A0A2Z1Y943 A0A2Z1Y943_9ASCO	ILMDTRQLAESELLVYGDTSVMIDTGCNDYKDAFNIGENFKKLVNERYLLEIDDN	1029
tr A0A1D8PK28 A0A1D8PK28_CANAL	ILMDTRQLAESELLVYGDTSVMIDTGCNDYKDAFNIGENFKKLVNERYLLEIDDN	1038
tr C4YQI4 C4YQI4_CANAW	ILMDTRQLAESELLVYGDTSVMIDTGCNDYKDAFNIGENFKKLVNERYLLEIDDN	1038
tr G8Y7L3 G8Y7L3_PICSO	ILMDTRQLAESELLVYGDTSVMIDTGCNDYKDAFNIGENFKKLVNERYLLEIDDN	1021
tr A0A0V1Q5J4 A0A0V1Q5J4_9ASCO	ILMDTRQLAESELLVYGDTSVMIDTGCNDYKDAFNIGENFKKLVNERYLLEIDDN	1042
tr A5DC55 A5DC55_PICGU	ILMDTRQLAESELLVYGDTSVMIDTGCNDYKDAFNIGENFKKLVNERYLLEIDDN	1009
tr A0A512UAU0 A0A512UAU0_9ASCO	ILMDTRQLAESELLVYGDTSVMIDTGCNDYKDAFNIGENFKKLVNERYLLEIDDN	1044
tr A0A1L0DOV4 A0A1L0DOV4_9ASCO	ILMDTRQLAESELLVYGDTSVMIDTGCNDYKDAFNIGENFKKLVNERYLLEIDDN	1016
tr A0A0L0NRD1 A0A0L0NRD1_CANAR	ILMDTRQLAESELLVYGDTSVMIDTGCNDYKDAFNIGENFKKLVNERYLLEIDDN	1015
***:***:*** : *:*****:***: : .. . : : ***:***:***: **		
tr A0A7D9GZJ1 A0A7D9GZJ1_DEKBR	VFKKMLLHAKKKYAAVNCTINA-DGKESTSLEVKGGLDMKRREYCPLSKEVSTFILEVLG	1136
tr A0A1E3NHX3 A0A1E3NHX3_9ASCO	VFKRLLLHSKKKYAAALNSINE-QNEEVTSLEVKGGLDMKRREYCPLSKELSIFVNNILH	1126
tr R9XN45 R9XN45_ASHAC	VFKLLLHAKKKYAAALNSFDK-AGNETTLEVKGLDMRRREYCPLSKEVSTKVLNTILS	1069
tr I6NDM1 I6NDM1_ERECY	VFKLLLHAKKKYAAALNVGFDK-LGNETSSLEVKGGLDMRRREYCPLSKEVSTQVLNTILS	1100
tr A0A109UYK0 A0A109UYK0_9SACH	LFKLLLHAKKKYAAALASFSDK-AGNERTLEVKGLDMRRREYCPLSKEVSTVNLNTILS	1083
tr A0A1G4J8L9 A0A1G4J8L9_9SACH	VYKLLLHAKKKYAAALNFLDK-SGETHTSLEVKGGLDMKRREYCPLSKEVSTHVLNTILS	1100
tr A0A0P1KX40 A0A0P1KX40_9SACH	VFRFDLHAKKKYAAALTVVFDK-SGAEQTLLEVKGGLDMRRREYCPLSKEVSTHVLNTILS	1084
tr A0A1G4KH46 A0A1G4KH46_9SACH	VFRFDLHAKKKYAAALTVSFDK-QGVEQTLEVKGGLDMRRREYCPLSKEVSVHVLNTILS	1094
tr A0A1G4KMB7 A0A1G4KMB7_9SACH	VFRFDLHAKKKYAAALTVVFDK-DGNEGTALEVKGLDMRRREYCPLSKEVTHVLNTILS	1104
tr A0A1S7HFL8 A0A1S7HFL8_9SACH	VFRFDLHAKKKYAAALVSLDM-DGKERVVLEVKGGLDMKRREYCPLSRDISIHVLNKLS	1068
tr A0A1Q3A9M9 A0A1Q3A9M9_ZYGR	VFKLLLHAKKKYAAALTVSLDK-EGNERTVLEVKGGLDMKRREYCPLSRDVSIHLNTILS	1079
tr A0A5P2U3L1 A0A5P2U3L1_KLULC	VFKRLLLHAKKKYAAALVTFDK-QNNEKTTLLEVKGGLDMKRREYCPLSKEVSTRVNLNTILS	1055
tr WOTGP7 WOTGP7_KLUMD	VFKRLLLHAKKKYAAALVTFDK-QNNEKTTLLEVKGGLDMKRREYCPLSKEVSTRVNLNTILS	1059
sp P13382 DPOA YEAST	VFKRLLLHAKKKYAAALTVNLDK-NGNGTTLEVKGGLDMKRREYCPLSRDVSIHLNTILS	1095
tr A0A0L8VIK5 A0A0L8VIK5_9SACH	VFKRLLLHAKKKYAAALTVNLDK-NGNGTTLEVKGGLDMKRREYCPLSRDVSIHLNTILS	1095
tr A0A7G7JEL9 A0A7G7JEL9_CANGB	VFKRLLLHAKKKYAAALTVNLDK-NGNGTTLEVKGGLDMKRREYCPLSRDVSIHLNTILS	1081
tr A0A1B2JAV2 A0A1B2JAV2_PICPA	VFKRILLHAKKKYAAAMNLSQL--NGKMSOLEVKGLDMRRREYCQLSKIEISTHVLQILG	1066
tr C4R232 C4R232_KOMPG	VFKRILLHAKKKYAAAMNLSQL--NGKMSOLEVKGLDMRRREYCQLSKIEISTHVLQILG	1067
tr A0A642UJM6 A0A642UJM6_DIURU	VFKRILLHAKKKYAAAMVITYDKESDTENLKEVKGLDIVRREYCQLSKDVSTFILGKILG	1048
tr A0A1E4U1L7 A0A1E4U1L7_PACTA	VFKRILLHAKKKYAAAMVVSF-ENNTEVLNLEVKGGLDMRRREYCQLSKEVSTVNLKLS	1143
tr A0A1V2L9F2 A0A1V2L9F2_CYBFA	VFKRILLHAKKKYAAAMVVTDFK-AGNETQLEVKGGLDMRRREYCQLSKEVSEVNLKLS	1086
tr A0A2Z1Y943 A0A2Z1Y943_9ASCO	VFKRILLHAKKKYAAAMNASINKTGEVEASLEVKGGLDMRRREYCQLSKDISTFVLEKILS	1089
tr A0A1D8PK28 A0A1D8PK28_CANAL	VFKRILLHAKKKYAAAMNASINKATNEETATLEVKGGLDMRRREYCQLSKDISTFVLTKILS	1098
tr C4YQI4 C4YQI4_CANAW	VFKRILLHAKKKYAAAMNASINKATNEETATLEVKGGLDMRRREYCQLSKDISTFVLTKILS	1098
tr G8Y7L3 G8Y7L3_PICSO	VFKRILLHAKKKYAAAMVNSIDKTTGKEITILEVKGLDMRRREYCQLSKDISTYVLEKILS	1081
tr A0A0V1Q5J4 A0A0V1Q5J4_9ASCO	VFKRILLHAKKKYAAAMVNSIDRKTGQEVKLEVKGGLDMRRREYCQLSKDISTFVLEKILS	1102
tr A5DC55 A5DC55_PICGU	VFKRILLHAKKKYAAAMVNSIDRKTGEEKVVLEVKGGLDMRRREYCQLSKDISTFILEKILA	1069
tr A0A512UAU0 A0A512UAU0_9ASCO	VFKRILLHAKKKYAAAMVNSFDKTTGKEVLSLEVKGGLDMRRREYCQLSKDISTFVLEKILS	1104
tr A0A1L0DOV4 A0A1L0DOV4_9ASCO	VFKRILLHAKKKYAAAMVNSLDRRTGKEVTALEVKGGLDMRRREYCQLSKEVSTFVLEKILS	1076
tr A0A0L0NRD1 A0A0L0NRD1_CANAR	VFKRLLLHAKKKYAAAMVNSLDRRTGKEISSLEVKGGLDMRRREYCQLSKEVSTFVLEKILS	1075
***:***:*** : *:*****:***: : .. . : : ***:***:***: **		

		ZBMx2		
tr A0A7D9GZJ1 A0A7D9GZJ1_DEKBR		GLQPLESTVTDSERFRDAKDFLINCPK	QASIKFGGIQSSQQYRVTSGIK	GEODSIIP
tr A0A1E3NHX3 A0A1E3NHX3_9ASCO		ELQPLESTIPDSERFRMDKDLKLCTTNSC	SEFLFGGIQPSADYQMTFSGIK	KQOEHNHLH
tr R9XN45 R9XN45_ASHAC		ELQPLESTISDEERFKDATVLRVTC	CPSQIKPFPGGIVASNHYHFTYSGLO	KAONHIFA
tr I6NDM1 I6NDM1_ERECY		NLQPLESTISDEERFKDTEWLDITCS	SQHQHKFPFGGITASNSHYQWTYNGLO	KHODHLFS
tr A0A109UYK0 A0A109UYK0_9SACH		ALQPLESTISDEERFKDTSKLLVSCPS	QNKFPGGIVASNHYIFSYNGLQ	KHONHLF
tr A0A1G4J8L9 A0A1G4J8L9_9SACH		SLEPLESSIISDEERFKNSAQQLTLRC	PNCNEFLFGGIVASNHYHTICYNGLK	RNDHLYS
tr A0A0P1KX40 A0A0P1KX40_9SACH		SLQPLESSIISDEERFKDVASLTLQC	PCSQQKFLFGGIVASKRYQMSYSGLQ	SSCQHTFA
tr A0A1G4KH46 A0A1G4KH46_9SACH		NLQPLESAISDEERFKGVASLTLQC	PGSQQRFPFGGIVASKYQMNFGNLQ	SAHHHQFS
tr A0A1G4KMB7 A0A1G4KMB7_9SACH		GLOPLESSIISDEERFKNAAPLILEC	CSSHQKFPGGIVASKRYRMSYTGIC	EIOQHQFT
tr A0A1S7HLF8 A0A1S7HLF8_9SACH		SLQPLETTISDEERFRKDADVPLEMNC	PCQGQKFPFGGIVASKYHSMCYNGLQ	KHONHFLA
tr A0A1Q3A9M9 A0A1Q3A9M9_ZYGR		SLQPLETTISDEERFRDVLQI	RCPCQDQEFPFGGIVSSNHYMMCYSGLQ	KHONHFLT
tr A0A5P2U3L1 A0A5P2U3L1_KLULC		TLQPLETTISDEERFRDVSVPLETC	PNCDTKFVYGGIVASKYRFVYSGIQ	CSNHNIMT
tr WOTGP7 WOTGP7_KLUMD		NLQPLETTISDEERFGDSVSL	ELACPSOSAKFAGGGIVASNDYRVVYSGIQ	TKOSHILT
sp P13382 DPOA YEAST		NLQPLETTITDVERFKDTVTLELSC	PSCDKRFPPFGGIIVSSNNYRVSYNGLQ	CKHCEQLFT
tr A0A0L8VIK5 A0A0L8VIK5_9SACH		NLQPLETTITDVERFKDTVTLELSC	PSCDKRFPPFGGIIVSSNNYRVSYNGLQ	CKHCEQLFT
tr A0A7G7JEL9 A0A7G7JEL9_CANGB		ELEPLETTISDEERFKDSSNLVNC	PNCNCKKEPYGGIVASNNYQHCYNGLQ	CKSODYLFS
tr A0A1B2JAV2 A0A1B2JAV2_PICPA		LLRPLESCISDEERFKGAATLELTCT	CGAHTFTGGIQESPSPSVTSFTSGI	HNGOTELS
tr C4R232 C4R232_KOMPG		LLQPLESCISDEERFKEAATLELTCT	CGAHTFAFGGIQESSYYTVTSGI	CGSSQGTQLS
tr A0A642UJM6 A0A642UJM6_DIURU		AYVPMENALKDEERFSQCSHFVVTCT	CGKRFRGGIQSSPDYKVGVNGV	CASQNQSI
tr A0A1E4U1L7 A0A1E4U1L7_PACTA		ELQPLEMSISDEERFKNSEYTLQCS	CGSSFKFGGIQASKDMITHHGVC	QNOSTNFS
tr A0A1V2L9F2 A0A1V2L9F2_CYBFA		ILQPLESTISDEERFKSSVPLKINCH	CGASFKFGGIVASHDMMTFNGVQ	KKOSKTLS
tr A0A2Z1Y943 A0A2Z1Y943_9ASCO		EIAPILESGISDEERFRQSSYFV	ICPSCEKAFRFGGILASKD	YKVTFNGI
tr A0A1D8PK28 A0A1D8PK28_CANAL		EILPIESNISDEERFRATSYLV	LHCKCGSSFRFGGIMAETYK	VTFNGVCTKCN
tr C4YQI4 C4YQI4_CANAW		EILPIESNISDEERFRATSYLV	LHCKCGSSFRFGGIMAETYK	VTFNGVCTKCN
tr G8Y7L3 G8Y7L3_PICSO		DIVPIESSISDEERFRQASYLV	LECKCGAKRFRGGIMASTD	YQTVTNGIK
tr A0A0V1Q5J4 A0A0V1Q5J4_9ASCO		EILPLESNISDEERFRQSSYLV	LECKCGSKFRFGGIMASNEYK	VTFNGVQ
tr A5DC55 A5DC55_PICGU		EIVPLESNISDEERFRQCSYLV	LHCKCGVNFRFGGIVASNDYK	VTFNGVQ
tr A0A512UAU0 A0A512UAU0_9ASCO		DVLPLETNVSDEERFRQTSYLV	LKCKCGHHFRFGGIMASPDYNI	KFHGISGKCN
tr A0A1L0D0V4 A0A1L0D0V4_9ASCO		DISPLESNITDIEFRQSSYLV	LECKCGAKFRFGGILLASKD	YKTFNGVQ
tr A0A0L0NRD1 A0A0L0NRD1_CANAR		EIAPILESKITDIEFRQSSYFL	VLRCRGANFRFGGLQPSRDY	LITFNGIK
		Znx3 (4Fe-4S)		
tr A0A7D9GZJ1 A0A7D9GZJ1_DEKBR		VLSVTAOLEHFIRHEISLYYAGYL	TSD-OGL-TTRQSVSVYGR	CGIGGQGTARS
tr A0A1E3NHX3 A0A1E3NHX3_9ASCO		PLSIAAQALELFIRTELSIYYQCWL	VODD-OGI-QTRQSVYGR	KGVM
tr R9XN45 R9XN45_ASHAC		ALQLTSQLEFAIRAHISLYAAWMV	ODDPCTGT-VTRQLSVFGKR	CLN-----EAK
tr I6NDM1 I6NDM1_ERECY		LIELTCQLERSIRSHISLYYAGWL	I	TGVM
tr A0A109UYK0 A0A109UYK0_9SACH		PIALSCQLERAIRTHISLYYAGWL	CDPDTGNN-LTRQISVFGR	CLN-----EN
tr A0A1G4J8L9 A0A1G4J8L9_9SACH		PIQLSCQLERSIRAQISLYYAGWV	QDDSTCGN-QTRSISVFGR	CLN-----EE
tr A0A0P1KX40 A0A0P1KX40_9SACH		PLQMSAQLESAIRSHISMYYAGYL	QDDSTCGN-NTRQISVFGR	CLS-----EN
tr A0A1G4KH46 A0A1G4KH46_9SACH		PLQLSCQLESHIRSHISRYYAGYL	QDDSTCGV-VTRQISVFGR	CLS-----EN
tr A0A1G4KMB7 A0A1G4KMB7_9SACH		PLQVSSQLERAIRSHISMYYAGNL	QDDSTCGI-QTRQISVFGR	CLS-----ED
tr A0A1S7HLF8 A0A1S7HLF8_9SACH		PLQLTCQLERSIRSHISLYYAGWL	QDDPTGNN-VCRQISVFGR	CLN-----EG
tr A0A1Q3A9M9 A0A1Q3A9M9_ZYGR		PLQLTCQLERSIRLHISLYYAGWL	QDDRTCCNNNSIRQISVFGR	CLN-----DG
tr A0A5P2U3L1 A0A5P2U3L1_KLULC		PIQVTSQLERIIHISLYYAGWL	QDDSTCGN-VTRQISVFGR	CLN-----DG
tr WOTGP7 WOTGP7_KLUMD		PIQVTSQLERVIRAHISLYYASWV	QDDSTCGN-VTRQISVFGR	CLM-----DG
sp P13382 DPOA YEAST		PLQLTSQIEHSIRAHISLYYAGWL	QCDSTCGI-VTRQISVFGR	CLN-----DG
tr A0A0L8VIK5 A0A0L8VIK5_9SACH		PLQLTSQIEHSIRAHISLYYAGWL	QCDSTCGI-VTRQISVFGR	TGVM
tr A0A7G7JEL9 A0A7G7JEL9_CANGB		PLQMTCQLERAIRSHISLYYAGWL	QCDPTCGN-VTRQLSVFGKR	CLN-----DG
tr A0A1B2JAV2 A0A1B2JAV2_PICPA		SLRITSQLERLIRKHISLYYAGWL	VODDVSOGV-HTRQSVYGR	CGINTGRSVC
tr C4R232 C4R232_KOMPG		SLRVASQLERLIRHISLYYAGWL	VODDVSOGV-HTRQSVYGR	CGINTGRSVC
tr A0A642UJM6 A0A642UJM6_DIURU		GYTLAAQLEYAIRRLAAYYAGWL	VONDASCHN-RTRQISVYGR	CGHGT-----VCG
tr A0A1E4U1L7 A0A1E4U1L7_PACTA		NLRVTSQLEKIRANISLYYSGWL	YGR	G
tr A0A1V2L9F2 A0A1V2L9F2_CYBFA		CLQLSSQLERTIRGFISQYYQGV	VODDVTGNN-TRQISVYGR	CGHSGKAHCGKGIM
tr A0A2Z1Y943 A0A2Z1Y943_9ASCO		VLEINSQLEVTLRQHISLYYAGWL	QDDSTCGM-TTKQISVYGR	CLGSSGKAHCGKGIM
tr A0A1D8PK28 A0A1D8PK28_CANAL		AIKLTSQLSMIRKHALYYAGWL	QDDPAOGI-TTRQISVYGR	CGNSGRAMDGKGVM
tr C4YQI4 C4YQI4_CANAW		AIKLTSQLSMIRKHALYYAGWL	QDDPAOGI-TTRQISVYGR	CGNSGRAMDGKGVM
tr G8Y7L3 G8Y7L3_PICSO		LIKLTSQLEVAIRKHISLYYCGWL	QDDSAOGI-ETRQISVYGR	CLGASGKAYCGKGIM
tr A0A0V1Q5J4 A0A0V1Q5J4_9ASCO		ILRITSQVEYTLRNHISLYYAGWL	QDDSAOGI-TTRQISVYGR	CGTSGKAHCGKGVM
tr A5DC55 A5DC55_PICGU		TFRIAQLENMIRRHISLYYAGWL	VODDGAOGI-TTRQISVYGR	CGASGTTATCGKGVM
tr A0A512UAU0 A0A512UAU0_9ASCO		PLKIASQLEVEIRKHLALYYAGWL	VODDGAOGI-KTRQISVYGR	CGASGTAHCGKGVM
tr A0A1L0D0V4 A0A1L0D0V4_9ASCO		CLKLTSQLESIRKHLALYYAGWL	VODDGAOGV-KTRQISVYGR	CGSSGKAHCGKGVM
tr A0A0L0NRD1 A0A0L0NRD1_CANAR		TLKVTSQLENSIRKHLALYYAGWL	VODDGAOGV-KTRQISVYGR	CGNSGKAHCGKGVM

tr A0A7D9GZJ1 A0A7D9GZJ1_DEKBR	RMLYNQLLYLQSLFDVEKAKKRTLKPLYDITTONKD---DLPPELNQGELDAL	1467
tr A0A1E3NHX3 A0A1E3NHX3_9ASCO	R <sup>Y</sup> YS <sup>T</sup> DKM <sup>I</sup> YNQLLYLQ <sup>T</sup> LF <sup>D</sup> I <sup>E</sup> KTKQQQLKPLPS---DK---PLPVTLNKGEIDAL	1458
tr R9XN45 R9XN45_ASHAC	H <sup>Y</sup> RS <sup>D</sup> KQLYNQLLYFDSLFDIVKT <sup>K</sup> NKDLKPL <sup>P</sup> LYLPGDVD---MFAEP <sup>L</sup> SSSQVAAL	1398
tr I6NDM1 I6NDM1_ERECY	F <sup>Y</sup> RYSD <sup>D</sup> KQLYNQLLYFDSLFD <sup>I</sup> EKNKQ <sup>K</sup> QELKPL <sup>P</sup> LYLPNDAD---IPTD <sup>L</sup> NNSSQIAAL	1437
tr A0A109UYK0 A0A109UYK0_9SACH	F <sup>Y</sup> RYSD <sup>D</sup> KQLYNQLLYFDSLFD <sup>I</sup> INKNAQ <sup>A</sup> LLR <sup>P</sup> LYLPGD <sup>D</sup> ---EPKEPLSSSQIAAL	1416
tr A0A1G4J8L9 A0A1G4J8L9_9SACH	F <sup>Y</sup> RYSD <sup>D</sup> KQLYNQLLYFESLFD <sup>V</sup> SKNKSQ <sup>L</sup> KPL <sup>P</sup> LYSPDSD---KPTHKL <sup>P</sup> TSQVAAL	1435
tr A0A0P1KX40 A0A0P1KX40_9SACH	F <sup>Y</sup> RYTD <sup>D</sup> KQLYNQLLYFESLFD <sup>V</sup> DNKKA <sup>L</sup> Q <sup>L</sup> R <sup>P</sup> LYLPDHPD---APEAKIPTSQLLAL	1418
tr A0A1G4KH46 A0A1G4KH46_9SACH	S <sup>Y</sup> S <sup>T</sup> D <sup>D</sup> KQLYNQLLQYFGSLFDV <sup>V</sup> VNNKRELLKPL <sup>P</sup> LYRPDH <sup>D</sup> ---AAETK <sup>I</sup> PTCQVLAL	1429
tr A0A1G4KMB7 A0A1G4KMB7_9SACH	S <sup>Y</sup> LYT <sup>D</sup> D <sup>D</sup> KQLYNQLLYFESLFD <sup>V</sup> D <sup>V</sup> DNK <sup>N</sup> KELV <sup>K</sup> PL <sup>P</sup> LYRPDH <sup>D</sup> ---IPETKL <sup>L</sup> STSQVLAL	1439
tr A0A1S7HLF8 A0A1S7HLF8_9SACH	H <sup>Y</sup> RYSD <sup>D</sup> KQLYNQLLYFESLFD <sup>V</sup> DNKCE <sup>N</sup> KNQGLKPL <sup>P</sup> LYHEGDAD---FFGETMTDSSVKAL	1403
tr A0A1Q3A9M9 A0A1Q3A9M9_ZYGRO	H <sup>Y</sup> RYSD <sup>D</sup> KQLYNQLLYFESLFD <sup>V</sup> DNKCE <sup>N</sup> KNQGLKPL <sup>P</sup> LYVEGDTD---FPRELLTDSSVKAL	1415
tr A0A5P2U3L1 A0A5P2U3L1_KLULC	S <sup>Y</sup> YKS <sup>D</sup> D <sup>D</sup> KQLYNQMLYFQSLFDLQ <sup>N</sup> KQ <sup>L</sup> T <sup>L</sup> KPL <sup>P</sup> LYHGS <sup>D</sup> N-----APEKSLSHSEIQAL	1391
tr WOTGP7 WOTGP7_KLUMD	S <sup>Y</sup> YKS <sup>D</sup> D <sup>D</sup> KR <sup>L</sup> YNQMLYFQSLFDL <sup>N</sup> LNKNT <sup>N</sup> QALKPL <sup>P</sup> LYHPGDKN-----APEKP <sup>L</sup> SHSEIQAL	1395
sp P13382 DPOA_YEAST	R <sup>Y</sup> Y <sup>K</sup> SD <sup>D</sup> KQ <sup>L</sup> YNQMLYFDSLFD <sup>C</sup> ENK <sup>N</sup> KQ <sup>L</sup> KPL <sup>P</sup> LYP <sup>D</sup> LD <sup>D</sup> ---Y <sup>P</sup> KEQLTESSIKAL	1430
tr A0A0L8VIK5 A0A0L8VIK5_9SACH	F <sup>Y</sup> Y <sup>K</sup> SD <sup>D</sup> KQLYNQLLYFDSLFD <sup>C</sup> ENK <sup>N</sup> KQ <sup>L</sup> KPL <sup>P</sup> LYP <sup>D</sup> LD <sup>D</sup> ---Y <sup>P</sup> KEQLTESSIKAL	1430
tr A0A7G7JEL9 A0A7G7JEL9_CANGB	H <sup>Y</sup> RT <sup>D</sup> D <sup>D</sup> KQLYNQLLYFDSLFD <sup>C</sup> ENK <sup>N</sup> KER <sup>L</sup> KPL <sup>P</sup> LYLEGDOD---Y <sup>P</sup> SEKLSDT <sup>S</sup> INAL	1416
tr A0A1B2JAV2 A0A1B2JAV2_PICPA	S <sup>Y</sup> Y <sup>T</sup> D <sup>D</sup> RELYNQLLYLES <sup>L</sup> FD <sup>V</sup> DKARR <sup>Q</sup> ELR <sup>L</sup> PLQDD <sup>R</sup> DKSSPD <sup>G</sup> EMKKG <sup>I</sup> SQSTLDAL	1405
tr C4R232 C4R232_KOMPG	S <sup>Y</sup> Y <sup>T</sup> D <sup>D</sup> RELYNQLLYLES <sup>L</sup> FD <sup>V</sup> DKARR <sup>Q</sup> ELR <sup>L</sup> PLQDD <sup>R</sup> ENRIRSDGET <sup>K</sup> KG <sup>I</sup> SQSTLDAL	1406
tr A0A642UJM6 A0A642UJM6_DIURU	R <sup>Y</sup> ET <sup>D</sup> D <sup>D</sup> KR <sup>L</sup> YNQLLYFSSLFDV <sup>V</sup> D <sup>V</sup> KAK <sup>K</sup> PL <sup>K</sup> S <sup>L</sup> D <sup>I</sup> PG-----VEALPHGS <sup>I</sup> AL	1370
tr A0A1E4U1L7 A0A1E4U1L7_PACTA	R <sup>Y</sup> Y <sup>S</sup> D <sup>D</sup> KQLYNQLLYFSSLFDV <sup>V</sup> D <sup>V</sup> KAK <sup>K</sup> RS <sup>L</sup> KPL <sup>P</sup> LYEEQ <sup>G</sup> EDK---ENV <sup>K</sup> PTIQGA <sup>L</sup> DAL	1482
tr A0A1V2L9F2 A0A1V2L9F2_CYBFA	R <sup>Y</sup> Y <sup>K</sup> SD <sup>D</sup> KQLYNQLLYFDSLFD <sup>V</sup> EKNKSQ <sup>K</sup> IKPL <sup>P</sup> LYD <sup>P</sup> ED---K---TKPEPLIKGQ <sup>I</sup> DAL	1406
tr A0A2Z1Y943 A0A2Z1Y943_9ASCO	R <sup>Y</sup> Y <sup>K</sup> SD <sup>D</sup> KR <sup>L</sup> YNQLLYFDSLFD <sup>V</sup> DA <sup>F</sup> KAK <sup>K</sup> QL <sup>L</sup> R <sup>P</sup> LYD <sup>P</sup> FA---A---NPSPSKL <sup>F</sup> GGQ <sup>I</sup> DAL	1427
tr A0A1D8PK28 A0A1D8PK28_CANAL	R <sup>Y</sup> Y <sup>N</sup> D <sup>D</sup> KQLYNQLLYFQ <sup>S</sup> IF <sup>D</sup> VD <sup>V</sup> DKTR <sup>G</sup> E <sup>L</sup> R <sup>L</sup> PL <sup>V</sup> DALEESKD---KQLP <sup>L</sup> PSG <sup>Q</sup> VEAL	1432
tr C4YQI4 C4YQI4_CANAW	R <sup>Y</sup> Y <sup>N</sup> D <sup>D</sup> KQLYNQLLYFQ <sup>S</sup> IF <sup>D</sup> VD <sup>V</sup> DKTR <sup>G</sup> E <sup>L</sup> R <sup>L</sup> PL <sup>V</sup> DALEESKD---KQLP <sup>L</sup> PSG <sup>Q</sup> VEAL	1432
tr G8Y7L3 G8Y7L3_PICSO	R <sup>Y</sup> Y <sup>S</sup> D <sup>D</sup> KALYNQLLYFDSLFD <sup>V</sup> VNKAK <sup>K</sup> Q <sup>K</sup> L <sup>K</sup> PL <sup>P</sup> LYESHENI-E---NIPP <sup>K</sup> PSAQ <sup>V</sup> AL	1412
tr A0A0V1Q5J4 A0A0V1Q5J4_9ASCO	R <sup>Y</sup> Y <sup>K</sup> SD <sup>D</sup> KALYNQLLYFDSLFD <sup>V</sup> VNKAK <sup>K</sup> Q <sup>K</sup> L <sup>K</sup> PL <sup>P</sup> LYD <sup>V</sup> GA-----PLPLASGQ <sup>V</sup> DAL	1430
tr A5DC55 A5DC55_PICGU	R <sup>Y</sup> Y <sup>K</sup> SD <sup>D</sup> KALYNQLLYFSSLFDV <sup>V</sup> DKAR <sup>K</sup> TEL <sup>K</sup> PL <sup>P</sup> V <sup>F</sup> DAFD---KG---SEPAKIASLGQ <sup>I</sup> AL	1401
tr A0A512UAU0 A0A512UAU0_9ASCO	R <sup>Y</sup> Y <sup>K</sup> SD <sup>D</sup> KALYNQLLYFSSLFDV <sup>V</sup> DKAMKKL <sup>K</sup> LPK <sup>I</sup> SE-----E---SKVPEMTEGE <sup>I</sup> NAL	1432
tr A0A1L0D0V4 A0A1L0D0V4_9ASCO	R <sup>Y</sup> Y <sup>T</sup> D <sup>D</sup> KALYNQLLYFDSLFD <sup>V</sup> VNKAK <sup>K</sup> LR <sup>P</sup> V <sup>S</sup> -----D---SALENLSEGQLNAL	1401
tr A0A0L0NRD1 A0A0L0NRD1_CANAR	R <sup>Y</sup> Y <sup>S</sup> D <sup>D</sup> KALYNQLLYFDSLFD <sup>V</sup> VNKAK <sup>K</sup> LR <sup>P</sup> I <sup>Y</sup> D---TD---VVAPELSTGQ <sup>I</sup> NAL	1405
	*.* : ***: :: :** * : : : : : : : :	**

//End of C-terminal of DNA pol  $\alpha$  from yeasts

tr A0A7D9GZJ1 A0A7D9GZJ1_DEKBR	VEQSRPEFGVIQS <sup>V</sup> VTKY <sup>L</sup> DDCGRRYYDMGS <sup>I</sup> FGFLK-----	1504
tr A0A1E3NHX3 A0A1E3NHX3_9ASCO	AEQNRKQFS <sup>L</sup> LN <sup>R</sup> RV <sup>V</sup> KY <sup>L</sup> DV <sup>N</sup> GRRYY <sup>V</sup> DLAS <sup>I</sup> FRVN-----	1494
tr R9XN45 R9XN45_ASHAC	AEQNRD <sup>L</sup> LELN <sup>R</sup> RAV <sup>V</sup> VQ <sup>K</sup> Y <sup>L</sup> SQCGR <sup>R</sup> YYDMG <sup>E</sup> I <sup>F</sup> FMGQ-----	1436
tr I6NDM1 I6NDM1_ERECY	SEQNRET <sup>N</sup> INR <sup>S</sup> V <sup>V</sup> VQ <sup>K</sup> Y <sup>L</sup> DQ <sup>C</sup> GR <sup>R</sup> YYDMG <sup>D</sup> I <sup>F</sup> DFMKA-----	1475
tr A0A109UYK0 A0A109UYK0_9SACH	SEQNREL <sup>F</sup> TSQAVVQ <sup>V</sup> Q <sup>K</sup> YLNQ <sup>C</sup> GR <sup>R</sup> YYD <sup>I</sup> RA <sup>F</sup> DFMV-----	1453
tr A0A1G4J8L9 A0A1G4J8L9_9SACH	SEQNREL <sup>F</sup> EIMNSV <sup>V</sup> VQ <sup>K</sup> Y <sup>L</sup> SDCGRRYYDMKS <sup>I</sup> FS <sup>F</sup> MGPP-----	1474
tr A0A0P1KX40 A0A0P1KX40_9SACH	GEQNREL <sup>F</sup> QVMSSV <sup>V</sup> VQ <sup>K</sup> YLANCARR <sup>F</sup> VD <sup>M</sup> DA <sup>I</sup> FAF <sup>A</sup> VAN-----	1457
tr A0A1G4KH46 A0A1G4KH46_9SACH	GEQNRLDFQT <sup>M</sup> MESV <sup>V</sup> VQ <sup>K</sup> YLGNC <sup>R</sup> ARS <sup>I</sup> FD <sup>F</sup> FMNPE-----	1468
tr A0A1G4KMB7 A0A1G4KMB7_9SACH	GEQNREL <sup>F</sup> QT <sup>M</sup> KEV <sup>V</sup> VQ <sup>K</sup> YLDN <sup>C</sup> CR <sup>R</sup> YYDMG <sup>T</sup> I <sup>F</sup> DFMVTH-----	1478
tr A0A1S7HLF8 A0A1S7HLF8_9SACH	SEQNRDYLGI <sup>A</sup> RG <sup>V</sup> VQ <sup>K</sup> Y <sup>L</sup> SDCGRRYYDM <sup>S</sup> V <sup>I</sup> FD <sup>F</sup> MTAAK-----	1443
tr A0A1Q3A9M9 A0A1Q3A9M9_ZYGRO	SEQNRDYL <sup>D</sup> V <sup>G</sup> RV <sup>V</sup> Q <sup>K</sup> Y <sup>L</sup> GD <sup>C</sup> GR <sup>R</sup> YYNM <sup>S</sup> N <sup>I</sup> FD <sup>F</sup> MTPAK-----	1455
tr A0A5P2U3L1 A0A5P2U3L1_KLULC	AEQNRL <sup>F</sup> EV <sup>V</sup> VQ <sup>K</sup> Y <sup>L</sup> SDCG <sup>R</sup> AY <sup>V</sup> NMGS <sup>I</sup> FD <sup>F</sup> MLSQ-----	1430
tr WOTGP7 WOTGP7_KLUMD	AEQNRLSFETAKQ <sup>V</sup> V <sup>E</sup> KY <sup>L</sup> SDCG <sup>R</sup> AY <sup>V</sup> NMGS <sup>I</sup> FD <sup>F</sup> MQO-----	1433
sp P13382 DPOA_YEAST	TEQNREL <sup>M</sup> ETGR <sup>S</sup> V <sup>V</sup> VQ <sup>K</sup> Y <sup>L</sup> ND <sup>C</sup> GR <sup>R</sup> YYDM <sup>T</sup> S <sup>I</sup> FD <sup>F</sup> MLN-----	1468
tr A0A0L8VIK5 A0A0L8VIK5_9SACH	TEQNREL <sup>M</sup> ETGR <sup>S</sup> V <sup>V</sup> VQ <sup>K</sup> Y <sup>L</sup> ND <sup>C</sup> GR <sup>R</sup> YYDM <sup>T</sup> S <sup>I</sup> FD <sup>F</sup> MLN-----	1468
tr A0A7G7JEL9 A0A7G7JEL9_CANGB	TEQNRLD <sup>L</sup> LELN <sup>R</sup> HISV <sup>V</sup> VQ <sup>K</sup> Y <sup>L</sup> SDCG <sup>R</sup> YYDM <sup>T</sup> AT <sup>I</sup> EF <sup>M</sup> NN-----	1454
tr A0A1B2JAV2 A0A1B2JAV2_PICPA	AEQNRTDFRAFQSV <sup>V</sup> VQ <sup>K</sup> Y <sup>L</sup> FEQ <sup>C</sup> GR <sup>R</sup> YYDM <sup>S</sup> LFNFS-----	1441
tr C4R232 C4R232_KOMPG	AEQNRTDFGAFQSV <sup>V</sup> VQ <sup>K</sup> Y <sup>L</sup> FEQ <sup>C</sup> GR <sup>R</sup> YYDM <sup>S</sup> LFNFS-----	1442
tr A0A642UJM6 A0A642UJM6_DIURU	AEQNREL <sup>F</sup> TRCT <sup>S</sup> V <sup>P</sup> E <sup>K</sup> Y <sup>L</sup> SNCGR <sup>R</sup> YYDM <sup>G</sup> S <sup>I</sup> FN <sup>F</sup> LQV-----	1408
tr A0A1E4U1L7 A0A1E4U1L7_PACTA	CEQNRFDD <sup>F</sup> DICRGV <sup>V</sup> IN <sup>R</sup> Y <sup>L</sup> NN <sup>S</sup> GR <sup>R</sup> YYDM <sup>G</sup> S <sup>I</sup> FEF <sup>M</sup> V <sup>P</sup> E <sup>Q</sup> KAVN	1526
tr A0A1V2L9F2 A0A1V2L9F2_CYBFA	SEQNRD <sup>L</sup> FEVQKV <sup>V</sup> VQ <sup>K</sup> Y <sup>L</sup> SDCG <sup>R</sup> YYNM <sup>S</sup> IF <sup>D</sup> FMK-----	1443
tr A0A2Z1Y943 A0A2Z1Y943_9ASCO	AEQNRIY <sup>F</sup> GY <sup>S</sup> KEV <sup>V</sup> VQ <sup>K</sup> Y <sup>L</sup> LGSC <sup>R</sup> YYNM <sup>G</sup> AL <sup>F</sup> EF <sup>M</sup> S-----	1464
tr A0A1D8PK28 A0A1D8PK28_CANAL	AEQNREL <sup>F</sup> ICQEV <sup>V</sup> VQ <sup>K</sup> Y <sup>L</sup> GE <sup>C</sup> GR <sup>R</sup> YYNM <sup>G</sup> S <sup>I</sup> FD <sup>F</sup> FI <sup>R</sup> N-----	1470
tr C4YQI4 C4YQI4_CANAW	AEQNREL <sup>F</sup> ICQEV <sup>V</sup> VQ <sup>K</sup> Y <sup>L</sup> GE <sup>C</sup> GR <sup>R</sup> YYNM <sup>G</sup> S <sup>I</sup> FD <sup>F</sup> FI <sup>R</sup> N-----	1470
tr G8Y7L3 G8Y7L3_PICSO	AEQNRENFSACQDV <sup>V</sup> VQ <sup>K</sup> Y <sup>L</sup> R <sup>D</sup> SG <sup>R</sup> YYDM <sup>G</sup> <sup>I</sup> FD <sup>F</sup> MGSI <sup>K</sup> S-----	1453
tr A0A0V1Q5J4 A0A0V1Q5J4_9ASCO	AEQNREL <sup>F</sup> SICQNV <sup>V</sup> VQ <sup>K</sup> Y <sup>L</sup> SECGR <sup>R</sup> YYNM <sup>S</sup> IF <sup>D</sup> FMSSK-----	1470
tr A5DC55 A5DC55_PICGU	AEQNREHFVFH <sup>V</sup> Q <sup>G</sup> Y <sup>I</sup> V <sup>D</sup> RYLA <sup>E</sup> C <sup>G</sup> RR <sup>R</sup> YYNM <sup>G</sup> S <sup>I</sup> FD <sup>F</sup> MPK-----	1438
tr A0A512UAU0 A0A512UAU0_9ASCO	AEQNRLVLF <sup>D</sup> Q <sup>C</sup> Y <sup>I</sup> V <sup>D</sup> Y <sup>L</sup> KDCG <sup>R</sup> YYNM <sup>K</sup> A <sup>I</sup> FD <sup>F</sup> MP <sup>T</sup> PPM-----	1472
tr A0A1L0D0V4 A0A1L0D0V4_9ASCO	AEQNRLFGNCREV <sup>V</sup> D <sup>R</sup> Y <sup>L</sup> KDCG <sup>R</sup> YYNM <sup>K</sup> AI <sup>F</sup> EF <sup>M</sup> A-----	1438
tr A0A0L0NRD1 A0A0L0NRD1_CANAR	AEQNRAMFDHC <sup>R</sup> EV <sup>V</sup> N <sup>K</sup> Y <sup>L</sup> ADCGR <sup>R</sup> YYNM <sup>R</sup> S <sup>I</sup> FD <sup>F</sup> MV-----	1442
	*.* : : : * : : : * : : : * : :	*

Fig. 12 MSA analysis of DNA polymerase  $\alpha$  catalytic subunit from different yeasts

<i>A0A7D9GZJ1_DEKBR, Dekkera bruxellensis</i>	<i>AOA1E3NHX3_9ASCO, Pichia membranifaciens</i>
<i>R9XN45_ASHAC, Ashbya aceri</i>	<i>I6NDM1_ERECY, Eremothecium cymbalariae</i>
<i>AOA109UYK0_9SACH, Eremothecium sinecaudum</i>	<i>AOA1G4KH46_9SACH, Lachancea meyersii</i>
<i>AOA1G4KMB7_9SACH, Lachancea nothofagi</i>	<i>AOA0P1KX40_9SACH, Lachancea quebecensis</i>
<i>AOA1G4J8L9_9SACH, Lachancea mirantina</i>	<i>AOA1S7HLF8_9SACH, Zygosaccharomyces parabailii</i>
<i>AOA1Q3A9M9_ZYGRO, Zygosaccharomyces rouxi</i>	<i>AOA5P2U3L1_KLULC, Kluyveromyces lactis</i>
<i>WOTGP7_KLUMD, Kluyveromyces marxianus</i>	<i>P13382/DPOA_YEAST, Saccharomyces cerevisiae</i>
<i>AOA0L8VIK5_9SACH, Saccharomyces boulardii</i>	<i>AOA7G7JEL9_CANGB, Candida glabrata</i>
<i>AOA1B2JAV2_PICPA, Komagataella pastoris</i>	<i>C4R232_KOMPG, Komagataella phaffii</i>
<i>AOA642UJM6_DIURU, Diutina rugosa</i>	<i>AOA1E4U1L7_PACTA, Pachysolen tannophilus</i>
<i>AOA1V2L9F2_CYBFA, Cyberlindnera fabianii</i>	<i>AOA2Z1Y943_9ASCO, Scheffersomyces spartinae</i>
<i>AOA1D8PK28_CANAL, Candida albicans</i>	<i>C4YQI4_CANAW, Candida albicans</i>
<i>G8Y7L3_PICSO, Pichia sorbitophila</i>	<i>A0A0V1Q5J4_9ASCO, Debaryomyces fabryi</i>
<i>A5DC55_PICGU, Meyerozyma guilliermondii</i>	<i>AOA512UAU0_9ASCO, Metschnikowia sp.</i>
<i>AOA1L0D0V4_9ASCO, Candida intermedia</i>	<i>AOA0L0NRD1_CANAR, Candida auris</i>

## Active Site Analyses of DNA polymerase $\alpha$

Human and yeast (*S. cerevisiae*) DNA polymerases  $\alpha$  are extensively studied by both SDM and X-ray crystallographic techniques. The crystal structure of the CTD of the catalytic subunit (p180c) of the human DNA pol  $\alpha$ , in complex with the B subunit was reported by Suwa et al. [26] They found that the p180c adopted an elongated asymmetric saddle shape, with a three-helix bundle in the middle and two ZBM (Zn1 and Zn2) on each side. The invariant 8 Cys residues form the tetrahedral coordination geometries with zinc ions of the ZBM; (Zn1: Cys-1283, Cys-1286, Cys-1310, and Cys-1315 and Zn2: Cys-1348, Cys-1353, Cys-1371, and Cys-1374, numbering from human DNA pol  $\alpha$ ). [26]

It is interesting to note, all the 8 conserved Cys residues are also present in the CTD of the catalytic subunits of all four eukaryotic B-family DNA polymerases (pol  $\alpha$ ,  $\delta$ ,  $\epsilon$  and  $\zeta$ ) suggesting their importance in all B-family DNA polymerases. The 8 Cys residues form two ZBM and are known as CysA (Zn1) and CysB, (Zn2). Netz et al. demonstrated by *in vivo* and *in vitro* experiments that CysB motif was bound to [4Fe-4S] cluster in all the four B-family polymerases. [29]

This finding suggests an important physiological role of the [4Fe-4S] cluster in the eukaryotic B-family polymerases. Furthermore, they have also shown that the Zn-binding CysA motif was required for DNA pol processivity and CysB for interaction with the regulatory subunit B of DNA pol  $\alpha$ . Interestingly, the CTD of the non-catalytic subunit (p68) of the human primase, also contains a [4Fe-4S] cluster instead of a Zn-finger and was shown to interact with a primed-template, only if the primer strand contains a 5'-terminal triphosphates as in the eukaryotic RNA primases. [30]

The MSA analysis also confirms the two ZBM (marked) at the C-terminal of the polymerase  $\alpha$ . (Figs. 11, 12). Baranovskiy et al. have shown that in the primase complex, both subunits are necessary for the initial dinucleotide formation, but the extension of the primer

corroborating that the ZBM in the regulatory subunit is essential for the first dinucleotide formation. [17]

Further, SDM analysis has shown that L<sup>868</sup>→M of yeast DNA pol  $\alpha$  exhibited normal catalytic efficiency and processivity, but reduced fidelity. The *pol1-L868M* allele conferred a mutator phenotype *in vivo*, which was strongly enhanced upon inactivation of the 3'→5' exonuclease of DNA pol  $\delta$  but not that of DNA pol  $\epsilon$ . [28] DePamphilis have determined crystal structures of the catalytic core (349–1258; 910 amino acids) of yeast DNA pol  $\alpha$  in an unliganded form (apo form), bound to an RNA primer/DNA template duplex (binary complex) and in a productive complex with RNA primer/DNA template and incoming dGTP (ternary complex). [31] For structural studies of the ternary complex, they used a polymerase mutant (D<sup>998</sup>→N) (highlighted in dark blue) with attenuated catalytic activity. They found that the catalytic region of DNA pol  $\alpha$ , like other DNA polymerases, adopted a cupped 'right-hand' fold consisting of a palm domain harboring the active site, a fingers domain interacting with the incoming nucleotides and a thumb domain that grips the primer-template duplex. Even though the PR activity is reported to be absent in DNA pol  $\alpha$ , by sequence similarity the proposed amino acids that could make a PR active site in DNA pol  $\alpha$  is highlighted in light blue and compared with well documented PR active site amino acids (Table 4). [28] Pavlov et al. found that although DNA pol  $\alpha$  contains an exonuclease domain, no proofreading activity was observed and the errors made by DNA pol  $\alpha$  were corrected by the next PR-proficient enzyme, DNA pol  $\delta$  in yeast. [28] It is clear that in eukaryotic genome replication, an additional checkpoint is introduced with the second primase, i.e., the DNA pol  $\alpha$ . This possibly ensures that the replicative polymerases, viz. the pol  $\delta$  and pol  $\epsilon$  do not function like other DdDps, but functional only during cell divisions with an absolute requirement for DNA primers.

Table 3 Proposed catalytic regions and MBSs of DNA polymerases  $\alpha$  of eukaryotic primosomesFrom animals (Human)

Zn <sup>2+</sup>	Catalytic regions	Mg <sup>2+</sup>
1) -P <sup>1</sup> CPTC <sup>2</sup> GTE <sup>3</sup> NIYDNVFDGSGTDMEPSLYR <sup>4</sup> C <sup>5</sup> SNID <sup>6</sup> C <sup>7</sup> K <sup>1316</sup> - CX2C CX4C	-945R <sup>5</sup> QKAL <sup>6</sup> K <sup>7</sup> L <sup>8</sup> TANS <sup>9</sup> M <sup>10</sup> Y <sup>11</sup> G <sup>12</sup> CLGF-	1)- <sup>1036</sup> EIDIDG-
2) -C <sup>1</sup> EEPTC <sup>2</sup> RNRTRHLPLQFSRTGPL <sup>3</sup> C <sup>4</sup> PACM <sup>5</sup> 1375- CX4C CX2C	-981R <sup>78</sup> EILMHT <sup>81</sup> K <sup>84</sup> E <sup>85</sup> MVQKMNLEVI <sup>93</sup> Y <sup>94</sup> D <sup>1002</sup> TDS-	2) -D <sup>1002</sup> TDS-

From yeasts (*S. cerevisiae*)

Zn <sup>2+</sup>	Catalytic regions	Mg <sup>2+</sup>
1) -C <sup>1</sup> DIANEMGQSLTPK <sup>2</sup> CQS <sup>3</sup> WDLS <sup>4</sup> EMYQVT <sup>5</sup> C <sup>6</sup> E <sup>725</sup> -	-937D <sup>7</sup> I <sup>8</sup> R <sup>9</sup> QQAL <sup>10</sup> K <sup>11</sup> L <sup>12</sup> TANS <sup>13</sup> M <sup>14</sup> Y <sup>15</sup> G <sup>16</sup> CLGY-	1)- <sup>1030</sup> EIDIDN-
2) -S <sup>1</sup> C <sup>2</sup> PSC <sup>3</sup> DKRFPFGGISSNNYRVSYNGLQ <sup>4</sup> C <sup>5</sup> H <sup>6</sup> C <sup>7</sup> E <sup>1318</sup> - CX2C CX2C	-975R <sup>7</sup> EILMDT <sup>8</sup> R <sup>9</sup> Q <sup>10</sup> LAELIGLRVV <sup>11</sup> Y <sup>12</sup> G <sup>13</sup> D <sup>14</sup> 996TDS-	2)-D <sup>996</sup> TDS-
3) -Q <sup>1</sup> C <sup>2</sup> DDST <sup>3</sup> C <sup>4</sup> GI-VTRQVSFGKR <sup>5</sup> C <sup>6</sup> LNDG <sup>7</sup> C <sup>8</sup> T <sup>1372</sup> - CX4C CX4C		

Based on the SDM, X-ray crystallographic data and MSA analysis, the proposed amino acids at the active site of the DNA

polymerase  $\alpha$  and the priming reactions are shown in Fig. 13.

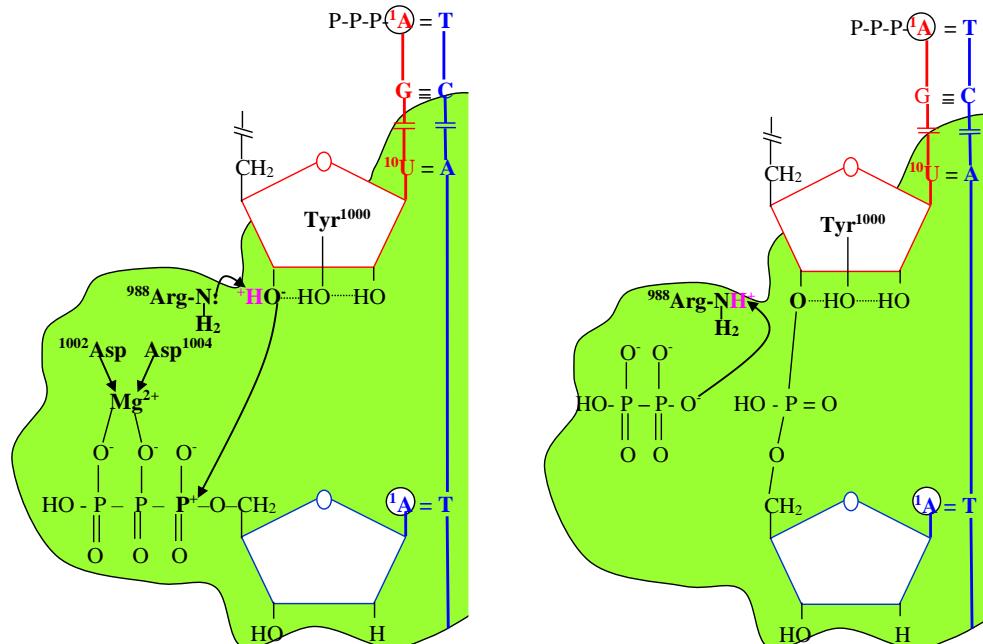


Fig. 13 The proposed active site structure of DNA polymerase  $\alpha$  and the DNA priming reactions (numbering from DNA pol  $\alpha$  of *H. sapiens*).

**CLUSTAL O (1.2.4) MSA of the regulatory subunit B of DNA polymerase  $\alpha$**   
**(Only the required sequences are shown)**

tr A0A1D5QNG5 AOA1D5QNG5_MACMU	MSASPPQQLAEEELQIFGIDCCE	EA	LIEK1KVELCVQYQGNQNEEGMVGELIAFCTSTHKVGLTSE	60
tr A0A2K5KX68 AOA2K5KX68_CERAT	MSASPPQQLAEEELQIFGIDCCE	EA	LIEK1KVELCVQYQGNQNEEGMVGELIAFCTSTHKVGLTSE	60
tr A0A2I3SDP8 AOA2I3SDP8_PANTR	MSASAPQQQLAEEELQIFGIDCCE	EA	LIEK1KVELCVQYQGNQNEEGMVGELIAFCTSTHKVGLTSE	60
tr A0A2J8TZ97 AOA2J8TZ97_PONAB	MSASAAQQQLAEEELQIFGIDCCE	EA	LIEK1KVELCVQYQGNQNEEGMVGELIAFCTSTHKVGLTSE	60
tr G3QKNO G3QKNO_GORGO	MSASAAQQQLAEEELQIFGIDCCE	EA	LIEK1KVELCVQYQGNQNEEGMVGELIAFCTSTHKVGLTSE	60
tr AOA2R9BL46 AOA2R9BL46_PANPA	MSASAAQQQLAEEELQIFGIDCCE	EA	LIEK1KVELCVQYQGNQNEEGMVGELIAFCTSTHKVGLTSE	60
<b>sp Q14181 DPOA2_HUMAN</b>	<b>MSASAAQQQLAEEELQIFGIDCCE</b>	<b>EA</b>	<b>LIEK1KVELCVQYQGNQNEEGMVGELIAFCTSTHKVGLTSE</b>	<b>60</b>
tr K7B9Q6 K7B9Q6_PANTR	MSASAAQQQLAEEELQIFGIDCCE	EA	LIEK1KVELCVQYQGNQNEEGMVGELIAFCTSTHKVGLTSE	60
tr A0A2K6NHA0 AOA2K6NHA0_RHIRO	MSASPQQQLAEEELQIFGIDCCE	EA	LIEK1KVELCVQYQGNQNEEGMVGELIAFCTSTHKVGLTSE	60
tr A0A2K5ZN81 AOA2K5ZN81_MANLE	MSASPQQQLAEEELQIFGIDCCE	EA	LIEK1KVELCVQYQGNQNEEGMVGELIAFCTSTHKVGLTSE	60
tr A0A2K6EC81 AOA2K6EC81_MACNE	MSASPQQQLAEEELQIFGIDCCE	EA	LIEK1KVELCVQYQGNQNEEGMVGELIAFCTSTHKVGLTSE	60
tr A0A2K5KER4 AOA2K5KER4_COLAP	MSASPQQQLAEEELQIFGIDCCE	EA	LIEK1KVELCVQYQGNQNEEGMVGELIAFCTSTHKVGLTSE	60
tr A0A2K6NHC8 AOA2K6NHC8_RHIRO	MSASPQQQLAEEELQIFGIDCCE	EA	LIEK1KVELCVQYQGNQNEEGMVGELIAFCTSTHKVGLTSE	60
tr AOA096MMDO AOA096MMDO_PAPAN	MSASPQQQLAEEELQIFGIDCCE	EA	LIEK1KVELCVQYQGNQNEEGMVGELIAFCTSTHKVGLTSE	60
*****	*****	*****	*****	*****
tr A0A1D5QNG5 AOA1D5QNG5_MACMU	TPLLAPAPAQEPTVTLGGQICCD	DS	NGKLNNSKSVIDLEGDRHSSGAQIPVDSLSELKEYSLFFGQ	300
tr A0A2K5KX68 AOA2K5KX68_CERAT	TPLLAPAPAQEPTVTLGGQICCD	DS	NGKLNNSKSVIDLEGDRHSSGAQIPVDSLSELKEYSLFFGQ	300
tr A0A2I3SDP8 AOA2I3SDP8_PANTR	TPLLAPAPAQEPTVTLGGQICCD	DS	NGKLNNSKSVIDLEGDRHSSGAQIPVDSLSELKEYSLFFGQ	300
tr A0A2J8TZ97 AOA2J8TZ97_PONAB	TPLLAPAPAQEPTVTLGGQICCD	DS	NGKLNNSKSVIDLEGDRHSSGAQIPVDSLSELKEYSLFFGQ	300
tr G3QKNO G3QKNO_GORGO	TPLLAPAPAQEPTVTLGGQICCD	DS	NGKLNNSKSVIDLEGDRHSSGAQIPVDSLSELKEYSLFFGQ	300
tr AOA2R9BL46 AOA2R9BL46_PANPA	TPLLAPAPAQEPTVTLGGQICCD	DS	NGKLNNSKSVIDLEGDRHSSGAQIPVDSLSELKEYSLFFGQ	299
<b>sp Q14181 DPOA2_HUMAN</b>	<b>TPLLAPAPAQEPTVTLGGQICCD</b>	<b>DS</b>	<b>NGKLNNSKSVIDLEGDRHSSGAQIPVDSLSELKEYSLFFGQ</b>	<b>300</b>
tr K7B9Q6 K7B9Q6_PANTR	TPLLAPAPAQEPTVTLGGQICCD	DS	NGKLNNSKSVIDLEGDRHSSGAQIPVDSLSELKEYSLFFGQ	300
tr A0A2K6NHA0 AOA2K6NHA0_RHIRO	TPLLAPAPAQEPTVTLGGQICCD	DS	NGKLNNSKSVIDLEGDRHSSGAQIPVDSLSELKEYSLFFGQ	273
tr A0A2K5ZN81 AOA2K5ZN81_MANLE	TPLLAPAPAQEPTVTLGGQICCD	DS	NGKLNNSKSVIDLEGDRHSSGAQIPVDSLSELKEYSLFFGQ	250
tr A0A2K6EC81 AOA2K6EC81_MACNE	TPLLAPAPAQEPTVTLGGQICCD	DS	NGKLNNSKSVIDLEGDRHSSGAQIPVDSLSELKEYSLFFGQ	280
tr A0A2K5KER4 AOA2K5KER4_COLAP	TPLLAPAPAQEPTVTLGGQICCD	DS	NGKLNNSKSVIDLEGDRHSSGAQIPVDSLSELKEYSLFFGQ	300
tr A0A2K6NHC8 AOA2K6NHC8_RHIRO	TPLLAPAPAQEPTVTLGGQICCD	DS	NGKLNNSKSVIDLEGDRHSSGAQIPVDSLSELKEYSLFFGQ	300
tr AOA096MMDO AOA096MMDO_PAPAN	TPLLAPAPAQEPTVTLGGQICCD	DS	NGKLNNSKSVIDLEGDRHSSGAQIPVDSLSELKEYSLFFGQ	300
*****	*****	*****	*****	*****
tr A0A1D5QNG5 AOA1D5QNG5_MACMU	VVIMEGINNTGRKLVATKLYEGVPLPFYQPTTEEDGDFEQSMVLVACGHYTTSDSITYDPL	360		
tr A0A2K5KX68 AOA2K5KX68_CERAT	VVIMEGINNTGRKLVATKLYEGVPLPFYQPTTEEDGDFEQSMVLVACGHYTTSDSITYDPL	360		
tr A0A2I3SDP8 AOA2I3SDP8_PANTR	VVIMEGINNTGRKLVATKLYEGVPLPFYQPTTEADAEFQSMVLVACGHYTTSDSITYDPL	360		
tr A0A2J8TZ97 AOA2J8TZ97_PONAB	VVIMEGINNTGRKLVATKLYEGVPLPFYQPTTEADAEFQSMVLVACGHYTTSDSITYDPL	360		
tr G3QKNO G3QKNO_GORGO	VVIMEGINNTGRKLVATKLYEGVPLPFYQPTTEADAEFQSMVLVACGHYTTSDSITYDPL	360		
tr AOA2R9BL46 AOA2R9BL46_PANPA	VVIMEGINNTGRKLVATKLYEGVPLPFYQPTTEADAEFQSMVLVACGHYTTSDSITYDPL	359		
<b>sp Q14181 DPOA2_HUMAN</b>	<b>VVIMEGINNTGRKLVATKLYEGVPLPFYQPTTEADAEFQSMVLVACGHYTTSDSITYDPL</b>	<b>360</b>		
tr K7B9Q6 K7B9Q6_PANTR	VVIMEGINNTGRKLVATKLYEGVPLPFYQPTTEADAEFQSMVLVACGHYTTSDSITYDPL	360		
tr A0A2K6NHA0 AOA2K6NHA0_RHIRO	VVIMEGINNTGRKLVATKLYEGVPLPFYQPTTEADAEFQSMVLVACGHYTTSDSITYDPL	333		
tr A0A2K5ZN81 AOA2K5ZN81_MANLE	VVIMEGINNTGRKLVATKLYEGVPLPFYQPTTEADAEFQSMVLVACGHYTTSDSITYDPL	310		
tr A0A2K6EC81 AOA2K6EC81_MACNE	VVIMEGINNTGRKLVATKLYEGVPLPFYQPTTEADAEFQSMVLVACGHYTTSDSITYDPL	340		
tr A0A2K5KER4 AOA2K5KER4_COLAP	VVIMEGINNTGRKLVATKLYEGVPLPFYQPTTEADAEFQSMVLVACGHYTTSDSITYDPL	360		
tr A0A2K6NHC8 AOA2K6NHC8_RHIRO	VVIMEGINNTGRKLVATKLYEGVPLPFYQPTTEADAEFQSMVLVACGHYTTSDSITYDPL	360		
tr AOA096MMDO AOA096MMDO_PAPAN	VVIMEGINNTGRKLVATKLYEGVPLPFYQPTTEADAEFQSMVLVACGHYTTSDSITYDPL	360		
*****	*****	*****	*****	*****

Fig. 14 MSA of the regulatory subunit B of the DNA polymerase  $\alpha$

*A0A1D5QNG5\_MACMU, Macaca mulatta  
 A0A2I3SDP8\_PANTR, Pan troglodytes  
 G3QKNO\_GORGO, Gorilla gorilla gorilla  
 Q14181/DPOA2\_HUMAN, Homo sapiens  
 A0A2K6NHA0\_RHIRO, Rhinopithecus roxellana  
 A0A2K6EC81\_MACNE, Macaca nemestrina  
 A0A2K6NHC8\_RHIRO, Rhinopithecus roxellana*

*A0A2K5KX68\_CERAT, Cercocebus atys  
 A0A2J8TZ97\_PONAB, Pongo abelii  
 A0A2R9BL46\_PANPA, Pan paniscus  
 K7B9Q6\_PANTR, Pan troglodytes  
 A0A2K5ZN81\_MANLE, Mandrillus leucophaeus  
 A0A2K5KER4\_COLAP, Colobus angolensis palliates  
 A0A096MMDO\_PAPAN, Papio Anubis*

### PR function DNA polymerase $\alpha$ of eukaryotic primosomes

Error-free transcription is the hallmark of all DNA and RNA polymerases. Nonetheless, these enzymes do make mistakes, but at a rate of about  $10^{-4}$  to  $10^{-5}$ . The RNA primases from prokaryotes and eukaryotes which make only very short oligonucleotide primers ( $\sim 10$  nts) to initiate replication process do not possess a PR activity. [32] (These mistakes may not be deleterious, as they are removed in the final step of the DNA replication process and are replaced with dNTPs

by a PR-proficient enzyme). In eukaryotes, the RNA primer synthesized by the RNA primase is further extended by the DNA pol  $\alpha$ , which also again makes only short DNA primers ( $\sim 25$  nts). Eukaryotic primase also generates large amount of di- and tri-nucleotides, even though the associated DNA pol  $\alpha$  only uses primers at least 7 nucleotides long. [33] Pavlov et al. found that the DNA fragments synthesized by the DNA pol  $\alpha$  contained errors, as it does not have PR exonuclease activity. [28] In contrast to other DNA polymerases, DNA pol  $\alpha$ , purified from higher eukaryotes, showed no detectable exonuclease activity. [34 & references therein]

However, a potent exonuclease was detected after separating the 182-kDa polymerase subunit from the other three subunits of the enzyme in *Drosophila melanogaster*. [34] They found that the exonuclease activity co-sedimented with the polymerase subunit, suggesting that the two activities reside in the same polypeptide. MSA analysis of DNA pol  $\alpha$  from *D. melanogaster* and related organisms by sequence similarity showed a typical PR active site ( $\text{DxE} \rightarrow \text{CD} \rightarrow \text{YxxxD}$ -). However, one of the invariant amino acid –D- was not found in the animal sources, including the human enzyme. It is interesting to note, by sequence similarity, a PR exonuclease active-site amino acids are found from the animal sources, in the regulatory subunit B of DNA pol  $\alpha$

(Fig. 14). Table 4 shows the 5 invariant amino acids (-DxE-----D-----Y---D-) that make the PR active site in DNA/RNA polymerases in bacteria and viruses. MSA analysis have shown that out of the 5 amino acids that are making the DEDD family (DxE---D---H/Y---D) of PR exonucleases is also found in DNA pol  $\alpha$  sequences from yeasts and animals. These data suggest that the DNA pol  $\alpha$  might have lost its PR function during evolution because of mutation(s) in the active site acidic amino acid(s). These data show that the active site amino acids in the PR 3'→5' exonucleases are conserved evolutionarily. Out of 5 catalytic amino acids (-DxE→D→Y/H→D-), the third invariant D could possibly be replaced by an E in  $\alpha$  polymerases but with no activity (Table 4). This is further supported by with *Drosophila* primase-polymerase. [32]

Table 4 PR active site amino acids in different DdDps and RdRps

PR enzyme	First triad	D	Y/H*	→Gap←	D
Pol $\alpha$ -Yeasts ( <i>S. c.</i> ) <b>DdDps</b>	DV^E	--	Y	3	D
Pol $\alpha$ - Animals ( <i>H. s.</i> ) „	DVE	--	Y	3	D
Pol I- Bacteria ( <i>E. coli</i> ) „	DTE	YD	Y	3	D
Pol II- Bacteria ( <i>E. coli</i> ) „	DIE	FD	Y	3	D
RNase D- Bacteria ( <i>E. coli</i> )	DTE	FD	Y	3	D
Pol III (e-subunit)-Bacteria ( <i>E. coli</i> ) <b>DdDps</b>	DTE	FD	H	4	D
RNase T- Bacteria ( <i>E. coli</i> )	DVE	FD	H	4	D
SARS-CoV-1 (NSP14)- <b>RdRps</b>	DVE	FD	H	4	D
MERS-CoV (NSP14)- „	DVE	FD	H	4	D
SARS-CoV-2 (NSP14)- „	DVE	FD	H	4	D
SARS-related & Human-CoVs (Bat-RaTG3, Pangolin, Civet & 229E, HKU, OC43) NSP14-„	DVE	I/VD	H	4	D
Human-CoV (NL63) (NSP14) -„	DVE	FD	H	4	D
Adenovirus 5 (Human) <b>DdDps</b>	DVE	FD	Y	3	D
Vaccinia virus „	DIE	FD	Y	3	D
Small pox virus „	DIE	FD	Y	3	D
HSV-1 „	DIE	FD	Y	3	D
T2, T4, T6,T3,T7 phages „	DIE	FD	Y	3	D
Phage RB69-gp43 ( <i>E. coli</i> ) „	DIE	FD	Y	3	D
Klebsiella phage E1 „	DIE	FD	Y	3	D
Yersinia phage (JC221) „	DIE	FD	Y	3	D

Adapted from [35, 36]

CoVs, Coronaviruses; *S. c*, *Saccharomyces cerevisiae*; *H. s*, *Homo sapiens*; HSV, *Herpes simplex virus*.

The non-structural protein (NSP14) codes for the PR function in RdRps of coronaviruses

\*The invariant Y/H acts as the proton acceptor in the catalytic reaction, and the last invariant D is placed 3 to 4 amino acids downstream from the proton acceptor.

<sup>^</sup>The middle amino acid is not highly conserved as in animals.

## Conclusions

Replication of genomes is an indispensable activity in the life-cycle of all living cells. This ensures making an exact copy of the genome that is to be transferred to daughter cells during cell division. The primases and DNA polymerases play essential roles in this process. MSA analysis showed that the eubacterial primases are structurally different with no significant identity to the human primase, suggesting that it could act as potential drug targets. However, as the active sites are found to be similar, they may follow a similar catalytic mechanism to initiate the replication process. Even though, the PR function is reported to be muted in the catalytic subunit of eukaryotic DNA

primase, DNA pol  $\alpha$ , the PR active-site amino acids, found in the regulatory subunit may perform the PR function.

## Acknowledgments

The author wishes to thank Dr. N. Srinivasan, Former Professor, Department of Endocrinology, Post Graduate Institute of Basic Medical Sciences, University of Madras, Chennai, for his useful suggestions on the manuscript.

## Competing interests

The author has declared that no competing interests exist.

## References

1. Griep MA. Primase structure and function. Indian J Biochem Biophys, 1995; 32:171-178.
2. Burgers PMJ, Kunkel TA. Eukaryotic DNA replication fork. Annu Rev Biochem, 2017; 86, 417–438.
3. Palanivelu P. DNA polymerases – An insight into their active sites and mechanism of action, In: Recent Advances in Biological Research, SCIENCE DOMAIN International Book Publishers, UK, ISBN: 9788193422441, DOI: 10.9734/bpi/rabr/v1; Vol 1, pp 1-39: 2019.
4. Ilic S, Cohen S, Singh M, Tam B, Dayan A, Akabayov B. DnaG Primase - a target for the development of novel antibacterial agents. Antibiotics (Basel), 2018; 7: 72 1-19.
5. Johnson SK, Bhattacharyya S, Griep MA. DnaB Helicase Stimulates Primer Synthesis Activity on Short Oligonucleotide Templates. Biochemistry, 2000; 39:736–744.
6. Marians KJ, Prokaryotic DNA replication. Annu Rev Biochem, 1992; 61, 673.
7. Ogawa T, Okazaki T. Function of RNase H in DNA replication revealed by RNase H defective mutants of *Escherichia coli*. Mol Genet, 1984;193, 231–237.
8. Sun W, Tormo J, Steitz TA, Godson GN. Domains of *Escherichia coli* primase: functional activity of a 47-kDa N-terminal proteolytic fragment. Proc Natl Acad Sci (USA), 1994; 91, 11462-11466.
9. Keck JL, Roche DD, Lynch AS, Berger JM. "Structure of the RNA polymerase domain of *E. coli* primase". Science, 2000; 287: 2482–2486.
10. Ziegelin G, Linderoth NA, Calendar R, Lanka E. Domain Structure of Phage P4 a Protein Deduced by Mutational Analysis. J Bacteriol, 1995; 177: 4333-4341.
11. Godson GN, Schoenich J, Sun W, Mustaev AA. Identification of the magnesium ion binding site in the catalytic center of *Escherichia coli* primase by iron cleavage. Biochemistry, 2000; 39: 332-339.
12. Doublié S, Zahn KE. Structural insights into eukaryotic DNA replication. Front Microbiol, 2014; 5:1-8.
13. Perera RL, Torella R, Klinge S, Kilkenny ML, Maman JD, Pellegrini L. Mechanism for priming DNA synthesis by yeast DNA Polymerase  $\alpha$ . eLife, 2013; 2:e00482 DOI: [10.7554/eLife.00482](https://doi.org/10.7554/eLife.00482).

14. Burgers PM. Polymerase dynamics at the eukaryotic DNA replication fork. *J Biol Chem*, 2009; 284: 4041–4045.
15. Roth YF. Eucaryotic primase. *Eur J Biochem*, 1987; 165:473-481.
16. Frick DN, Richardson CC. DNA primases. *Annu Rev Biochem*, 2001; 70: 39–80
17. Baranovskiy AG, Zhang Y, Suwa Y, Babayeva ND, Gu J, Pavlov YI, Tahirov TH. Crystal Structure of the Human Primase. *J Biol Chem*, 2015; 290:5635-5646.
18. Baranovskiy, AG, ZhangY, SuwaY, Gu J, Babayeva ND, Pavlov YI, TahirovTH. Insight into the human DNA primase interaction with template-primer. *J Biol Chem*, 2016; 291, 4793 –4802.
19. Palanivelu, P. Single Subunit RNA Polymerases: An Insight into their Active Sites and Catalytic Mechanism, In: Advances and Trends in Biotechnology and Genetics. SCIENCE DOMAIN International Book Publishers, UK: pp 1-38: 2019
20. Copeland WC, Tan X. (1995) Active site mapping of the catalytic mouse primase subunit by alanine scanning mutagenesis. *J Biol Chem*, 270, 3905–3913.
21. Vaithiyalingam S, Arnett DR, Aggarwal A, Eichman, BF, Fanning E, Chazin WJ. Insights into eukaryotic primer synthesis from structures of the p48 subunit of human DNA primase. *J Mol Biol*, 2014; 426, 558–569.
22. Kilkenny ML, Longo MA, Perera RL, Pellegrini L. Structures of human primase reveal design of nucleotide elongation site and mode of Pol alpha tethering. *Proc Natl Acad Sci (USA)*, 2013; 110:15961–15966.
23. Vaithiyalingam S, Warren EM, Eichman BF, Chazin WJ. Insights into eukaryotic DNA priming from the structure and functional interactions of the 4Fe-4S cluster domain of human DNA primase. *Proc Natl Acad Sci (USA)*, 2010; 107, 13684–13689.
24. Pellegrini L. The Pol alpha-primase complex. *Subcell Biochem*, 2012; 62: 157–169.
25. Martínez-Jiménez MI, Calvo PA, García-Gómez S, Guerra-González S, Blanco L. The Zn-finger domain of human PrimPol is required to stabilize the initiating nucleotide during DNA priming. *Nucleic Acids Res*, 2018; 46: 4138–4151.
26. Suwa Y, Gu J, Baranovskiy AG, Babayeva ND, Pavlov YI, Tahirov TH. [Crystal Structure of the Human Pol α B Subunit in Complex with the C-terminal Domain of the Catalytic Subunit](#). *J Biol Chem*, 2015; 290:14328-14337.
27. Barnes R, Eckert K. Maintenance of Genome Integrity: How Mammalian Cells Orchestrate Genome Duplication by Coordinating Replicative and Specialized DNA Polymerases. *Genes*, 8:19. doi:10.3390/genes8010019: 2017.
28. Pavlov YI, Frahm C, Nick McElhinny SA, Niimi A, Suzuki M, Kunkel TA. Evidence that errors made by DNA polymerase alpha are corrected by DNA polymerase delta. *Curr Biol*, 2006; 16: 202–207.
29. Netz DJ, Stith CM, Stümpfig M, Köpf G, Vogel D, Genau HM, Stodola JL, Lill R, Burgers PM, Pierik AJ. Eukaryotic DNA polymerases require an iron-sulfur cluster for the formation of active complexes. *Nat Chem Biol*, 2012; 8: 125–132.
30. [Baranovskiy AG, Babayeva ND, Zhang Y, Gu J, Suwa Y, Pavlov YI, Tahirov TH](#). Mechanism of Concerted RNA-DNA Primer Synthesis by the Human Primosome. *J Biol Chem*, 2016; 291:10006-10020.
31. DePamphilis ML, 2016. Genome duplication: Genome Duplication: The Heartbeat of Developing Organisms. *Curr Top Dev Biol*, 2016; 116: 201–229.

32. Lehman IR and Kaguni LS. DNA Polymerase  $\alpha$ . J Biol Chem, 1989; 264: 4265:4268.
33. Kuchta RD, Stenge G Mechanism and Evolution of DNA Primases. Biochim Biophys Acta, 2010; 1804: 1180–1189.
34. Cotterill SM, Reyland ME, Loeb LA, Lehman R. A cryptic proofreading 3'-5' exonuclease associated with the polymerase subunit of the DNA polymerase-primase from *Drosophila melanogaster*. Proc Natl Acad Sci (USA), 1987;84: 5635-5639.
35. Ishino Y, Iwasaki H, Kato I, Shinagawa H. Amino Acid Sequence Motifs Essential to 3'→5' Exonuclease Activity of *Escherichia coli* DNA Polymerase II. J Biol Chem, 269; 1994:14655-14660.
36. Palanivelu, P. An overview of the proofreading functions in bacteria and SARS- Coronaviruses. Int J Biochem Res Rev, 2021; 30; 33-62.

**How To Cite This Article:**

**Peramachi Palanivelu** *Analyses of Priming Reactions and Proofreading Functions of Primases during Initiation of Replication of Prokaryotic and Eukaryotic Genomes* Br J Pharm Med Res, Vol.07, Issue 03, Pg. 3790 - 3886, May - June 2022. ISSN:2456-9836 Cross Ref DOI : <https://doi.org/10.24942/bjpmr.2022.988>

**Source of Support:** Nil **Conflict of Interest:** None declared

Your next submission with [British BioMedicine Institute](#) will reach you the below assets

- Quality Editorial service
- Swift Peer Review
- E-prints Service
- Manuscript Podcast for convenient understanding
- Global attainment for your research
- Manuscript accessibility in different formats  
([Pdf](#), [E-pub](#), [Full Text](#))
- Unceasing customer service



Track the below URL for one-step submission

<https://www.bjpmr.org/manuscript-submission/>