Protein family review

Replicative DNA polymerases

M Mar Albà

Address: Wohl Virion Centre, Department of Immunology and Molecular Pathology, University College London, Cleveland Street, London, W1T 4JF, UK. E-mail: m.alba@ucl.ac.uk

Published: 12 January 2001

Genome ***Biology*** 2001, **2(1)**:reviews3002.1–3002.4

The electronic version of this article is the complete one and can be found online at http://genomebiology.com/2001/2/1/reviews/3002

© BioMed Central Ltd (Print ISSN 1465-6906; Online ISSN 1465-6914)

Summary

Replicative DNA polymerases are essential for the replication of the genomes of all living organisms. On the basis of sequence similarities they can be classified into three types. Type A polymerases are homologous to bacterial polymerases I, Type B comprises archaebacterial DNA polymerases and eukaryotic DNA polymerase α, and the bacterial polymerase III class make up type C. Structures have been solved for several type A and B polymerases, which share a similar architecture. The structure of type C is not yet known. The catalytic mechanism of all three types involves two metal-ion-binding acidic residues in the active site. Replicative polymerases are constitutively expressed, but their activity is regulated through the cell cycle and in response to different growth conditions.

Gene organization and evolutionary history

Classification

On the basis of sequence similarities, DNA polymerases can fall into three groups: type A, type B and type C, which have homology to *polA* (pol I), *polB* (pol II) and *polC* (pol III) from *Escherichia coli*, respectively [1,2]. Type C polymerases are not known to share structural similarity with types A and B, so they will not be covered in much detail in this review. In addition to replicative DNA polymerases, these groups also include polymerases involved in other types of DNA synthesis and in DNA repair. For example in bacteria the main replicative DNA polymerase is pol III (type C), while DNA polymerase I (type A) is not essential for replication and DNA polymerase II (type B) is only present in a few bacteria. Pol I has a role in nucleotide excision repair and in the processing of Okazaki fragments that are generated on the lagging strand during DNA replication, while pol II is known to be involved in repair of DNA cross-links. Polymerase delta from eukaryotes belongs to the same polymerase group as the replicative DNA polymerase α (type B). The role of polymerase delta in replication however is not clear although there is evidence that it participates in post-replicational DNA repair.

Replicative DNA polymerases from some bacteriophages (T3, T5 and T7) and eukaryotic mitochondrial DNA polymerases have homology to bacterial polymerases I and are therefore type A polymerases. Eukaryote replicative polymerase α, archaebacterial DNA polymerases, viral DNA polymerases, DNA polymerases encoded in mitochondrial plasmids of various fungi and plants and some bacteriophage polymerases (T4 and RB69) all belong to type B. The bacterial DNA polymerase III class, members of which are responsible for most of the replicative DNA synthesis in bacteria, are type C DNA polymerases [1]. The three types share no obvious sequence similarity, but types A and B are structurally similar to each other (see below). In general, a single gene for each type is found in the different organisms, but there are exceptions; for example, some bacteria have several genes for type C DNA polymerases. Little is known about the structure of the DNA polymerase α gene in eukaryotes, although the mouse gene is known to contain four exons.

Characteristic structural features

Sequence features

Type A polymerases contain three conserved motifs: A, B (Prosite signature PS00447 [3]) and C. Motifs A and C (Figure 1) are part of the catalytic site, whereas motif B is involved in the binding of dNTPs. In type B polymerases, up to six regions of sequence homology have been identified. Regions I (Prosite signature PS00116) and II form part of the active site and are considered to be equivalent to polymerase

type A motifs C and A, respectively (Figure 1). In these regions, metal-ion-binding aspartic acid residues are in equivalent structural positions [4,5].

In addition to the catalytic domain, DNA polymerases often have additional domains required for editing activity, excision of Okazaki primers during replication (structure-specific 5' nuclease activity), or for interactions with other proteins. Bacterial and archaebacterial replicative DNA polymerases can contain a 3'-to-5' proofreading exonuclease domain, and the eukaryotic α polymerase catalytic subunit contains a zinc finger domain required for interactions with other subunits of the polymerase-primase complex [6]. A 35-residue fragment with the potential to form a Zn-finger has also been identified in type C DNA polymerase from bacteria. This region appears essential for the proper formation and/or function of the enzyme's polymerase site [7].

Structural features

The *E. coli* DNA polymerase I structure was detemined in 1985 [8], and since then, several other type A DNA polymerase structures, such as T7 DNA polymerase [9] (Figure 2a), have been solved. For type B DNA polymerases three crystal structures are available, from phage RB69 [10] and from two *Thermococcus* species of archaebacteria (Figure 2b) [11,12].

The catalytic domains of type A and B DNA polymerases have a common overall architecture, which resembles a right hand and consists of 'thumb', 'palm' and 'fingers' domains [4] (Figure 2a,b). A similar structure and catalytic mechanism is shared by other families of polymerases, such as eukaryotic DNA polymerase β, reverse transcriptases and RNA-dependent RNA polymerases. The most conserved region is the palm domain, which contains the catalytic site. The fingers and thumb have somewhat different arrangements in the two families, although the thumb always contains parallel or anti-parallel α helices that appear to interact with the minor groove of the primer-template complex, and the fingers have an α helix with conserved sidechains positioned at the blunt end of the primer-template complex.

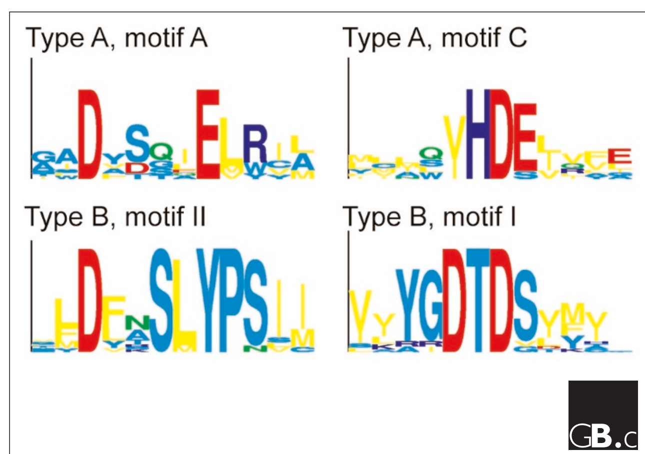

Figure 1
Logo representation of DNA polymerase domains from representative polymerase sequences. The sequences were extracted from SwissProt links in Prosite [3] to the DNA polymerase type A (PS00447) and DNA polymerase type B (PS00116) families. The families were initially found using 'DNA polymerase' as keyword. Sequence logos show the relative representation of the different amino acids at each sequence position in terms of the size of the appropriate letter in the single-letter amino acid code; the largest letters are the most conserved. They were constructed using the online server WebLogo [17]. Type A DNA polymerases (10 sequences) were from *E. coli* (polA), *Bacillus subtilis*, *Thermus aquaticus*, *Aquifex aeolicus*, *Mycobacterium tuberculosis*, phage T3, phage T5, phage T7, mitochondrial *Saccharomyces cerevisiae* DNA polymerase and mitochondrial *Homo sapiens* DNA polymerases. Type B DNA polymerases (15 sequences) were from *H. sapiens* (α and δ), *S. cerevisiae* (α and δ), *Drosophila melanogaster* (α), phage T4, phage RB69, *Chlorella* virus NY-2A, *Zea mays* plasmid S-1, herpes simplex virus 1, vaccinia virus, African swine fever virus, *Archaeoglobus fulgidus*, *Pyrococcus horikoshii*, *E. coli* (polB).

Localization and function

The replicative polymerases are required for the faithful replication of the genetic material and they perform this role by attaching the appropiate nucleotide to the nascent strand to match the template strand. The replication machinery, which includes a number of other well-conserved enzymes such as helicases and primases, assembles at the replication origin, where the replicative DNA polymerase initiates DNA synthesis using short DNA or RNA primers. All known DNA polymerases synthesize DNA in a 5'-to-3' direction.

The replication of DNA occurs before cell division. Replicative DNA polymerase genes are housekeeping genes and in eukaryotes the enzyme is found in all nuclei. The polymerase activity can be controlled in a cell-cycle-dependent manner or in response to different environmental conditions, however. For example, in mammals the activity of replicative DNA polymerase α is regulated by phosphorylation during the cell cycle [13] and its expression is stimulated by specific transcription factors during growth [14].

Enzyme mechanism

DNA synthesis is mediated by transfer of a phosphoryl group from the incoming dNTP to the DNA 3' OH, liberating a pyrophosphate and forming a new DNA phosphodiester bond. This reaction is catalysed by a mechanism that involves two metal ions, normally Mg^{2+}, with the participation of two aspartic acid residues that are structurally conserved among the different enzymes [5,6]. These are the aspartic acids conserved in motifs A and II, and C and I, from type A and type B polymerases, respectively (Figure 1). The first metal ion activates the 3'-OH for attack of the α phosphate of the incoming dNTP and the second metal ion

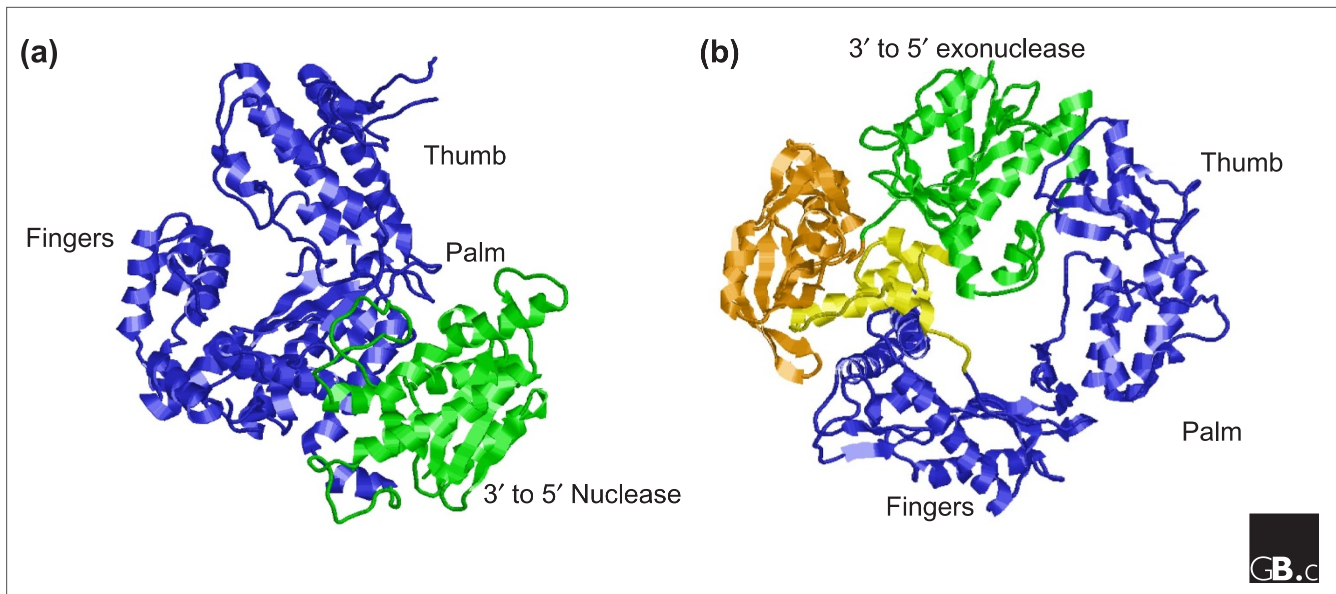

Figure 2
DNA polymerase protein structures. The figures were generated with the free software Rasmol (RasWin Molecular Graphics, Windows Version 2.4) and files extracted from the Protein Data Bank (PDB) [18]. **(a)** Bacteriophage T7 DNA polymerase (type A; PDB code: 1t7p) [9]. Blue, catalytic domain; green, 3'-to-5' exonuclease domain. **(b)** *Thermococcus gorgorianus* type B DNA polymerase [11] (PDB code: 1tgo). Blue, catalytic domain; green, 3'-to-5' exonuclease domain; orange, amino-terminal domain; yellow, linker region.

stabilizes the negative charge that builds up on the leaving oxygen and chelates the β- and γ-phosphates.

The catalytic mechanism of the type C DNA polymerase catalytic subunit may be similar to type A and B polymerases. Three conserved aspartic acid residues have been identified that strongly affect the polymerase activity and that may therefore be involved in the coordination of metal ions in the active site [15].

Frontiers

Replication in prokaryotes and eukaryotes is a relatively well studied pathway and a number of solved type A and type B DNA polymerase structures have provided evidence of a common architecture and active-site mechanism. Similar features have been observed in other non-replicative DNA polymerases. The determination of structures of type C DNA polymerases will be interesting, as it may show whether all polymerases can be unified as a single structural class.

An interesting aspect of replicative DNA polymerases is the diversity of enzymes, covering at least three different sequence families, that perform the same function. DNA replication is an ancestral function and as such is expected to be well conserved. The presence of different types of functionally homologous polymerases in different organisms suggests a complex evolution, however. One hypothesis, which involves transfer of DNA polymerase genes from viruses to eukaryotic genomes, has recently been put forward to attempt to address this question [16]. There is clearly much still to be learnt about these important enzymes.

References

1. Delarue M, Poch O, Tordo N, Moras D, Argos P: **An attempt to unify the structure of polymerases.** *Prot Engineering* 1990, **3:**461-467.
Pioneer work in the identification of conserved regions in various polymerase types and comparison with the *E. coli* polymerase I fold.
2. Braithwaite DK, Ito J: **Compilation, alignment, and phylogenetic relationships of DNA polymerases.** *Nucleic Acids Res* 1993, **21:**787-802.
A collection of aligned type A, B and C DNA polymerase sequences. Although somewhat outdated, this article is still a good reference for DNA polymerase families.
3. **Prosite** [http://www.expasy.ch/prosite/]
Database of protein domains.
4. Joyce CM: **Choosing the right sugar: how polymerases select a nucleotide substrate.** *Proc Natl Acad Sci USA* 1997, **94:**1619-1622.
Comparison of residues in the active site of different types of DNA and RNA polymerases. Emphasis is put on the possible substrate discriminative mechanisms.
5. Steitz TA: **DNA polymerases: structural diversity and common mechanisms.** *J Biol Chem* 1999, **274:**17395-17398.
An excellent review on the structure of DNA polymerases.
6. Mizuno T, Yamagishi K, Miyazawa H, Hanaoka F: **Molecular architecture of the mouse DNA polymerase alpha-primase complex.** *Mol Cell Biol*, 1999 **19:**7886-7896.
A detailed study of interactions between the four different subunits of the polymerase complex by selectively expressing combinations of them in cultured mammalian cells.
7. Barnes MH, Leo CJ, Brown NC: **DNA polymerase III of Gram-positive eubacteria is a zinc metalloprotein conserving an essential finger-like domain.** *Biochemistry* 1998, **37:**15254-15260.
Analysis of the functionality of the finger-like domain fragment in pol III.

8. Ollis DL, Brick P, Hamlin R, Xuong NG, Steitz TA: **Structure of large fragment of *Escherichia coli* DNA polymerase I complexed with dTMP.** *Nature* 1985, **313:**762-766.
Pioneer work in the determination of a DNA polymerase structure.
9. Doublié S, Tabor S, Long AM, Richardson CC, Ellenberger T: **Crystal structure of a bacteriophage T7 DNA replication complex at 2.2 Å resolution.** *Nature* 1998, **391:**251-258.
Includes the complete structure of the bacteriophage T7 DNA polymerase (type A), with polymerase and 3' to 5' exonuclease domains.
10. Wang J, Sattar AKMA, Wang CC, Karam JD, Konigsberg WH, Steitz TA: **Crystal structure of a pol α replication DNA polymerase from bacteriophage RB69.** *Cell* 1997, **89:**1087-1089.
First structure of a type B DNA polymerase to be determined.
11. Hopfner KP, Eichinger A, Engh RA, Laue F, Ankenbauer W, Huber R, Angerer B: **Crystal structure of a thermostable type B DNA polymerase from *Thermococcus gorgonarius*.** *Proc Natl Acad Sci USA* 1999, **96:**3600-3605.
Structure of an archaebacterial DNA polymerase and comparison to the RB69 DNA polymerase structure for the identification of features that confer increased thermostability to Taq polymerase.
12. Rodriguez AC, Park HW, Mao C, Beese LS: **Crystal structure of a pol alpha family DNA polymerase from the hyperthermophilic archaeon thermococcus sp. 9 degrees N-7.** *J Mol Biol* 2000, **299:**471-487.
Crystal structure of another thermostable archaeal DNA polymerase. A putative RNA-binding domain is identified in the amino-terminal region.
13. Voitenleitner C, Rehfuess C, Hilmes M, O'Rear L, Liao PC, Gage DA, Ott R, Nasheuer HP, Fanning E: **Cell cycle-dependent regulation of human DNA polymerase alpha-primase activity by phosphorylation.** *Mol Cell Biol* 1999, **19:**646-656.
An analysis of the phosphorylation of the polymerase α catalytic and p68 subunits by cyclin A/cdk2 and cyclin E/cdk2 kinases and its effect on initiation of replication.
14. Izumi M, Yokoi M, Nishikawa NS, Miyazawa H, Sugino A, Yamagishi M, Yamaguchi M, Matsukage A, Yatahaia F, Hanaoka F: **Transcription of the catalytic 180-kDa subunit gene of mouse DNA polymerase alpha is controlled by E2F, an Ets-related transcription factor, and Sp1.** *Biochim Biophys Acta* 2000, **1492:**341-352.
Identification of an E2F-binding site near the transcription start site that is essential for growth-dependent stimulation of gene expression.
15. Pritchard AE, McHenry CS: **Identification of the acidic residues in the active site of DNA polymerase III.** *J Mol Biol* 1999, **285:**1067-1080.
Comparison of sequences and mutagenesis analysis were used to identify residues in the active site of the catalytic subunit of DNA polymerase III.
16. Villarreal LP, DeFilippis VR: **A hypothesis for DNA viruses as the origin of eukaryotic replication proteins.** *J Virol* 2000, **74:**7079-7084.
Sequence comparison and phylogenetic trees of type A DNA polymerases from a variety of organisms, including DNA viruses. The hypothesis is formulated that replication genes, including those encoding DNA polymerases, were transferred from viruses to eukaryotes during evolution.
17. **WebLogo** [http://www.bio.cam.ac.uk/cgi-bin/seqlogo/logo.cgi]
A service for constructing logos (as in Figure 1) from sequence alignments.
18. **Protein Data Bank (PDB)** [http://pdb.ccdc.cam.ac.uk/pdb/]
Database of protein structures.