Published in final edited form as:

J Mol Biol. 2010 July 16; 400(3): 295–308. doi:10.1016/j.jmb.2010.05.030.

Identification of a New Motif in Family B DNA Polymerases by Mutational Analyses of the Bacteriophage T4 DNA Polymerase

Vincent Li¹, Matthew Hogg², and Linda J. Reha-Krantz^{1,*}

¹Department of Biological Sciences, University of Alberta, Edmonton, Alberta T6G 2E9, Canada

²Department of Microbiology and Molecular Genetics, University of Vermont, Burlington, Vermont 05405, USA

Abstract

Structure-based protein sequence alignments of family B DNA polymerases revealed a conserved motif that is formed from interacting residues between loops from the N-terminal and palm domains and between the N-terminal loop and a conserved proline residue. The importance of the motif for function of the bacteriophage T4 DNA polymerase was revealed by suppressor analysis. T4 DNA polymerases that form weak replicating complexes cannot replicate DNA when the dGTP pool is reduced. The conditional lethality provides the means to identify amino acid substitutions that restore replication activity under low dGTP conditions by either correcting the defect produced by the first amino acid substitution or by generally increasing the stability of polymerase complexes; the second type are global suppressors that can effectively counter the reduced stability caused by a variety of amino acid substitutions. Some amino acid substitutions that increase the stability of polymerase complexes produce a new phenotype - sensitivity to the antiviral drug phosphonoacetic acid. Amino acid substitutions that confer decreased ability to replicate DNA under low dGTP conditions or drug sensitivity were identified in the new motif, which suggests that the motif functions in regulating the stability of polymerase complexes. Additional suppressor analyses revealed an apparent network of interactions that link the new motif to the fingers domain and to two patches of conserved residues that bind DNA. The collection of mutant T4 DNA polymerases provides a foundation for future biochemical studies to determine how DNA polymerases remain stably associated with DNA while waiting for the next available dNTP, how DNA polymerases translocate, and the biochemical basis for sensitivity to antiviral drugs.

Keywords

NPL motif in family B DNA polymerases; stability of DNA polymerase complexes; DNA replication fidelity; sensitivity to phosphonoacetic acid; DNA polymerase translocation

© 2010 Elsevier Ltd. All rights reserved.

*Corresponding author: Department of Biological Sciences, CW405 Biosciences Bldg., University of Alberta, Edmonton, Alberta, T6G 2E9, Canada. Linda.Reha-Krantz@ualberta.ca, Phone: +1 780 492 5383, Fax: +1 780 492 9234 .

Publisher's Disclaimer: This is a PDF file of an unedited manuscript that has been accepted for publication. As a service to our customers we are providing this early version of the manuscript. The manuscript will undergo copyediting, typesetting, and review of the resulting proof before it is published in its final citable form. Please note that during the production process errors may be discovered which could affect the content, and all legal disclaimers that apply to the journal pertain.

Introduction

DNA polymerases interact with DNA in the polymerase active site and at several locations along the DNA binding groove (Fig. 1a). While the DNA polymerase holds the DNA primer-terminal region in the correct alignment for nucleotide incorporation, these interactions must relax to allow the DNA polymerase to reposition on the DNA template in order to bind the next incoming nucleotide. DNA polymerases must also be able to dissociate when DNA damage is encountered and to proofread, which requires strand separation and repositioning of the primer terminus in the exonuclease active site to form exonuclease complexes.¹ The processivity of many DNA polymerases is enhanced by association with “clamp” proteins such as PCNA and related doughnut-shaped structures that tether the DNA polymerase to DNA.² DNA polymerases, however, control the stability of replicating complexes, since single amino acid substitutions in the DNA polymerase can significantly affect stability. As explained below, the I417V and A737V substitutions in the bacteriophage T4 DNA polymerase decrease stability while the L412M substitution increases stability.³⁻⁶

DNA polymerase-DNA complexes formed with the wild type T4 DNA polymerase holoenzyme remain bound to DNA with a half-life of ~2.5 minutes when replication pauses because of reduced nucleotide pools,⁷ which is in sharp contrast to the almost immediate release (half-life of ~1 sec) when the holoenzyme encounters a blocking DNA hairpin structure in the template strand.⁸ Biochemical experiments reveal that the I417V- and A737V-DNA polymerases form less stable polymerase complexes with or without the clamp, and instability is exacerbated when dNTP pools are reduced.³⁻⁶ *In vivo*, the instability of polymerase complexes formed with the I417V- and A737V-DNA polymerases results in lethality when the mutant phage are propagated under low dGTP conditions on the *optA1* bacterial host,⁹ which expresses an elevated level of a dGTPase that degrades dGTP to the deoxynucleoside (dG) and tri-poly phosphate (PPP).^{10,11} Wild type phage T4 DNA replication is only slightly affected by reduced dGTP, but a reduced dGTP pool severely inhibits replication by the I417V- and A737V-DNA polymerases.^{3,9} The reduced ability of the I417V- and A737V-DNA polymerases to form stable polymerase complexes produces another characteristic called the antimutator phenotype. If polymerase complexes cannot be formed, this increases the opportunity to form exonuclease complexes with DNAs with mismatched and even matched primer termini.^{1,3-5,12-17}

The conditional lethality of the *optA1*-sensitive I417V- and A737V-DNA polymerases provides the means to select for *optA1*-resistant phage that have acquired a second mutation in the polymerase gene that restores polymerizing activity under low dGTP conditions. The doubly mutant DNA polymerases phenotypically resemble wild type because the second amino acid substitution provides a balancing change that counters the reduced stability of polymerase complexes conferred by the I417V and A737V amino acid substitutions. Most of the *optA1*-resistant phage recovered retain the original mutations in the T4 DNA polymerase gene that encode the I417V or A737V substitutions, but a second mutation was acquired which in many cases encoded the L412M substitution in the polymerase active site.¹⁸ T4 DNA polymerase residue L412 corresponds to L415 in the phage RB69 DNA polymerase (Fig. 1a). The structure of the closely related RB69 DNA polymerase is shown because a structure of the T4 DNA polymerase is not available, but the high amount of sequence and functional similarities allow structural assumptions for the T4 DNA polymerase based on RB69 DNA polymerase structures.¹⁹

The L412M substitution suppresses *optA1*-sensitivity, but only partially rescues the antimutator phenotype observed for the singly mutant I417V- and A737V-DNA polymerases.³ The biochemical basis for suppression of *optA1*-sensitivity is due to the

ability of the L412M substitution to increase the stability of polymerase complexes, even when dNTP pools are low.^{3,4,12} The increased stability of polymerase complexes has the potential to increase mismatch extension; hence, a mutator phenotype is observed for the L412M-DNA polymerase.^{1,3} The L412M-DNA polymerase is also sensitive to phosphonoacetic acid (PAA),^{3,20} a pyrophosphate (PPi) analog that inhibits replication by the herpes viral DNA polymerase by inhibiting the PPi exchange and pyrophosphorolysis reactions.²¹ PAA sensitivity requires formation of stable polymerizing complexes, specifically complexes that have not translocated to be in position to bind the next incoming nucleotide.^{22,23} PAA-sensitivity can be used to identify mutations that suppress drug-sensitivity and many of these encode amino acid substitutions that decrease the stability of polymerase complexes and confer *optA1*-sensitivity.^{3,22} Thus, identification of suppressors of *optA1*- and PAA-sensitivity can be used to gain insights into how the T4 DNA polymerase produces polymerase complexes that maintain stability under low dNTP pools, but still discriminate against the extension of mismatches that is required for high fidelity DNA replication.

Here we present structure-based protein sequence alignments and mutational studies of the T4 DNA polymerase that revealed a previously unidentified DNA polymerase motif that is conserved in many family B DNA polymerases. The new motif, which we call the NPL core, is formed at sites of interactions between residues in loops from the N-terminal and Palm domains, and interactions between the N-terminal domain loop and a conserved proline residue in the Linker region, which connects the N-terminal and palm domains (Fig. 1). Amino acid substitutions in the NPL core confer *optA1*- and PAA-sensitivities and, thus, we propose that the NPL core functions in regulating the stability of polymerase complexes. Although the NPL core does not contact DNA directly, further genetic studies revealed an apparent network of interactions that connect the NPL core to the fingers domain and to two conserved patches of residues that contact DNA. Since the NPL core is conserved in family B DNA polymerases, we propose that the NPL core is a major determinant of DNA polymerase stability during chromosome replication.

Results

Identification of NPL core residues by structure-based protein sequence alignments

Family B DNA polymerases are composed of palm, fingers, thumb and exonuclease domains and many have an N-terminal domain as shown for structural studies of the phage RB69 DNA polymerase (Fig. 1a),²⁴ for archaeal family B DNA polymerases,²⁵⁻²⁸ for the herpes viral DNA polymerase,²⁹ for the yeast DNA polymerase δ ,³⁰ and for *E. coli* DNA pol II.³¹ We observed that the N-terminal domain of the phage RB69 DNA polymerase interacts with the palm domain via close interactions between loops in the N-terminal and palm domains (Fig. 1b), as noted previously for the archaeal Sso DNA polymerase.²⁷ But additional interactions were observed between residues in the N-terminal loop with residues in a region of the polypeptide chain that joins the N-terminal to the palm domain – the linker region (Fig. 1b). These interactions were observed in all structures of family B DNA polymerases with N-terminal domains.²⁴⁻³⁴

Structure-based protein sequence alignments were performed as described in Materials and Methods to determine if NPL interactions involve protein sequence motifs that are conserved in family B DNA polymerases (Fig. 2). Note that a subset of family B DNA polymerases, which includes the adeno and ϕ 29 DNA polymerases, do not have a domain that corresponds to the N-terminal domain observed for other family B DNA polymerases³⁵ and this subset of family B DNA polymerases is not discussed here. Protein sequence alignments of family B DNA polymerases based on structure are indicated by an asterisk (*) and were extended to DNA polymerases from other organisms based on sequence

similarities. Previously undetected short motifs were revealed. Most striking was the presence of a highly conserved proline residue in the linker region, which is present in all of the family B DNA polymerases examined. The highly conserved proline residue, which is P378 in the T4 DNA polymerase and P381 in the RB69 DNA polymerase (Fig. 1b), is flanked in general by hydrophobic residues on the amino terminal side and by polar residues on the carboxyl terminal side (Fig. 2).

The palm loop sequence (Fig. 2) follows the highly conserved Motif A sequence (residues 411-420 in the RB69 DNA polymerase and 408-417 in the T4 DNA polymerase) in the primary structure, which assisted alignment. Motif A forms part of the polymerase active site (note residue L415 in the RB69 DNA polymerase in Fig. 1a, which corresponds to L412 in the T4 DNA polymerase). A conserved proline residue is present in the palm loop of phage and many archaeal family B DNA polymerases, most frequently adjacent to a serine residue (the SP motif), but the herpes viral DNA polymerase and many eukaryotic family B DNA polymerases have an alternative CF/CY motif at the analogous position. The archaeal Sso DNA polymerase has an apparent hybrid SY motif. Eukaryotic DNA polymerase ϵ has a P residue, as observed for the phage and archaeal DNA polymerases, but in a QP motif; *E. coli* DNA pol II has a DP motif. The palm motif also has a partially conserved N residue on the amino terminal side of the loop. On the carboxyl terminal side, the ET/DT sequence is observed for DNA polymerases with the SP palm loop motif; an ST or TT is present for the DNA polymerases with the CF/CY motif; DS and SA are observed for the yeast and human ϵ DNA polymerases, respectively. The consensus motif for the palm loop is Nh(SP or CF/CY)(D/ET or S/TT)h, where h indicates a hydrophobic residue.

The N-terminal loop interacts with the palm loop and with the linker region (Fig. 1b), but there is considerable variation in the size of the loop in family B DNA polymerases and little protein sequence conservation was observed (Fig. 2). For these reasons, N-terminal loop alignments are presented only for phage DNA polymerases and for family B DNA polymerases for which a structure has been determined.

Identification of residues in the NPL core by suppressor analyses

Suppressor analysis is a powerful genetic method to probe function and amino acid/domain interactions within a protein (intragenic suppression) or between proteins (intergenic suppression). Most mutations that suppress *optA1*-sensitivity are within the T4 DNA polymerase gene. The L412M substitution is a global suppressor of *optA1*-sensitivity since the increased stability of polymerase complexes conferred by the L412M substitution suppresses the *optA1*-sensitivity produced by amino acid substitutions that decrease stability.¹⁸ The increased stability of polymerase complexes produces PAA-sensitivity and the mutator phenotype.³ In the same genetic experiment that identified L412M,¹⁸ the I50L substitution in the NPL core was also selected as a global suppressor of *optA1*-sensitivity (Fig. 1), and like the L412M-DNA polymerase, the I50L-DNA polymerase is sensitive to PAA and also displays a mutator phenotype (Table 1, Fig. 3). Thus, the I50L substitution in the NPL core is expected to increase the stability of polymerase complexes. [Note that I50 in the T4 DNA polymerase corresponds to I52 in the RB69 DNA polymerase.] While the L412M substitution in the polymerase active center is in position to affect PPi and PAA binding directly, the I50L substitution in the NPL core must act indirectly to produce drug-sensitive polymerase complexes.

In a new experiment, two amino acid substitutions were identified in the NPL core that suppress *optA1*-sensitivity; these amino acid substitutions – P378L and P424L (P381 and P427 in the RB69 DNA polymerase) suppress the *optA1*-sensitivity of the R335C-DNA polymerase (R338 in the RB69 DNA polymerase) (Table 1, Fig. 1). The R335C amino acid substitution was identified as a suppressor of the PAA-sensitivity of the *tsP36*-DNA

polymerase, which has a duplication of residue D863 in the carboxyl terminal region of the T4 DNA polymerase.²² Since amino acid substitutions that suppress PAA-sensitivity often confer *optA1*-sensitivity, it was not surprising that the R335C-DNA polymerase is *optA1*-sensitive and displays an antimutator phenotype; both phenotypes are correlated with mutant DNA polymerases that form weak polymerase complexes. Although this was a small study, it is telling that the only amino acid substitutions identified that suppressed the *optA1*-sensitivity conferred by the R335C substitution were in the NPL core.

T4 strains expressing the singly mutant P378L- and P424L-DNA polymerases were constructed and characterized *in vivo* (Table 1). Amino acid substitutions that suppress *optA1*-sensitivity may be global suppressors that generally increase the stability of polymerase complexes and confer PAA-sensitivity and the mutator phenotype as observed for the L412M substitution, or the substitutions may specifically correct the defect produced by the first amino acid substitution. The P378L and P424L substitutions do not appear to be global suppressors. Only a 5-fold increase in the *rII* revertant frequency was observed for the P378L-DNA polymerase (Table 1) and PAA-sensitivity was not detected (Fig. 3). Furthermore, the P424L-DNA polymerase was phenotypically opposite to the PAA-sensitive/mutator L412M-DNA polymerase and displayed *optA1*-sensitivity and a strong antimutator phenotype as the *rII*⁺ revertant frequency was reduced 50-fold (Table 1). Thus, the P378L and P424L substitutions likely correct the defect in function produced by the R335C substitution. A series of additional genetic experiments are described below that provide insights into how the P378L and P424L substitutions suppress the *optA1*-sensitivity produced by R335C substitution. These experiments exploit the *optA1*-sensitivity of the P424L-DNA polymerase.

The L412M substitution suppresses the *optA1*-sensitivity of the P424L-DNA polymerase

If the *optA1*-sensitivity and antimutator phenotypes of the P424L-DNA polymerase are caused by formation of polymerase complexes with low stability, then the L412M substitution is expected to be identified as a suppressor. Individual high-titer cultures of T4 phage expressing the P424L-DNA polymerase were prepared under permissive conditions and then the lysates were plated on the restrictive *optA1*-host to identify *optA1*-resistant phage, which appear at a frequency of about 1 in 10 to 100 million. *optA1*-resistant phage were selected at 42 and 30 °C; the higher temperature is predicted to select for the most robust suppressor mutations. Wild type phage were isolated at 42 °C in which the mutation coding the P424L substitution reverted back to the wild type proline codon. The L412M substitution was also identified at 42 as well as 30 °C (Table 2), which indicates that the general ability of this amino acid substitution to suppress *optA1*-sensitivity extends to the P424L substitution in the NPL core. The doubly mutant L412M/P424L-DNA polymerase is not *optA1*-sensitive and replication fidelity as measured by reversion of the *rIIUV199oc* allele is between the strong antimutator phenotype observed for the P424L-DNA polymerase and the modest mutator phenotype observed for the L412M-DNA polymerase (Table 3).

Identification of amino acids within the NPL core that suppress the *optA1*-sensitivity of the P424L-DNA polymerase

A pseudorevertant was also selected at 42 °C in which the leucine 424 codon was converted to the codon for isoleucine (ATT) (Table 2). The P424I-DNA polymerase strain is informative. Since this strain produces plaques on *optA1* bacteria at 42 °C and displays only a 10-fold antimutator phenotype (Table 1), a proline residue is not essential for function at this position.

If the *optA1*-sensitivity of the P424L-DNA polymerase is caused by altered interactions within the NPL core, then compensating amino acid substitutions in the N-terminal loop and

linker region are expected to restore function. This proposal was confirmed by the identification of four *optA1*-resistant strains at 30 °C that retained the mutation that encoded the P424L substitution, but had acquired additional mutations that encoded amino acid substitutions within the NPL core. Three substitutions were located in the N-terminal loop: D49Y, I50T, and a duplication of residue I50, which was accompanied by the A580S substitution which resides outside of the NPL core. The P378H substitution was identified for the conserved proline residue in the linker region (Table 2). The identification of suppressor amino acid substitutions at the closest sites for direct physical interaction within the NPL core (Figs. 1 and 4a) provides evidence that residues in the NPL core interact and that these interactions are important for DNA polymerase function.

The compensatory amino acid substitutions within the NPL core are predicted to correct the defect in function caused by the P424L substitution rather than to be global suppressors like L412M since these amino acid substitutions were not identified in previous experiments and the close proximity of NPL core residues provides the means for physical interaction. This proposal was tested by constructing the I50L/P424L-DNA polymerase since I50L was not identified as a suppressor of the *optA1*-sensitivity conferred by the P424L substitution even though I50L was identified as a global suppressor in previous experiments.¹⁸ The reason for this is clear. While I50T significantly suppressed the *optA1*-sensitivity conferred by the P424L substitution and partially suppressed the antimutator phenotype, the I50L/P424L-DNA polymerase was still *optA1*-sensitive; the plating efficiency (0.1) is too low to be detected by genetic selection (Table 1). Thus, while the I50L substitution is a global suppressor of *optA1*-sensitivity conferred by amino acid substitutions *outside* of the NPL core, the I50T substitution is required to correct the defect produced by P424L substitution *within* the NPL core.

The use of suppressor analysis to identify NPL networks; connecting the NPL core to the fingers domain and to residues that bind DNA

Y460F/H (RB69 Y463), G466S (RB69 G469), L567F (RB69 L570), N579L (RB69 N582), A580S (RB69 A583)—Several amino acid substitutions were identified outside of the NPL core that suppress the *optA1*-sensitivity of the P424L-DNA polymerase. The DNA polymerase view in Fig. 4a shows the NPL core residues - I50, P378 and P424 in the T4 DNA polymerase (I52, P381, and P427 in the RB69 DNA polymerase) at the hub of interactions between several protein domains. Amino acid substitutions were identified in the base regions of the N (light green) and P (yellow) helices of the fingers domain - G466S and L567F, respectively (residues G469 and L570 in the RB69 DNA polymerase) (Fig. 4a and e). Note that residue L567 resides within the conserved Motif B sequence in family B DNA polymerases (Fig. 5a).

Amino acid substitutions were also identified for residues Y460 and N579 (residues Y463 and N582 in the RB69 DNA polymerase), which are in position to form H-bonds (Fig. 4a). The H-bond distance for RB69 DNA polymerase residues Y463 and N582 is ~ 2.6 Å in the closed structure, which increases to ~2.9 Å in the open structure. The Y460F/H and N579L substitutions in the T4 DNA polymerase are expected to disrupt H-bonding and other interactions that may extend to L567, which is discussed above and to A580 (A583 in the RB69 DNA polymerase) (Fig. 4a). The A580S substitution was identified in combination with a duplication of residue I50 in the NPL core (Table 2). While one of the substitutions alone may be responsible for suppression, it is likely that the combination of both substitutions is required, which will be tested in future experiments. Note, however, that A580 is a highly conserved residue in helix Q (Fig. 5a). Helix Q (dark green in Fig. 4a) interacts with residues in the conserved Y/FxGG/AxV motif (red residues in Fig. 4a), a sequence that is important for interactions with DNA in the primer-terminal region.^{36,37}

Thus, the amino acid substitutions for residues at base positions of helices N and P (G466S, L567F) and the substitutions for residues Y460, N582, and A580 appear to link P424 in the NPL core to the fingers domain and to conserved residues that are important for DNA interactions.

E98K (RB69 E100), S369A (RB69 S372), Q478R/H (RB69 Q481), F487L (RB69 L490)—Several amino acid substitutions that counter the *optA1*-sensitivity conferred by the P424L substitution are in position to affect fingers domain movements in addition to residues in the base regions of helices N and P of the fingers domain, G466 and L567, which were discussed above. The S369A substitution (S372 in the RB69 DNA polymerase) is in position to affect interactions with E471 (E474 in the RB69 DNA polymerase) in helix N (Figs. 4b and c). Q478 in helix N (Q481 in the RB69 DNA polymerase) and E98 (E100 in the RB69 DNA polymerase) appear to form an interacting pair of amino acids; three amino acid substitutions (E98K, Q478R, and Q478H) were identified that could affect such interactions.

The F487L substitution in helix N (L490 in the RB69 DNA polymerase) is at the edge of the only region of significant protein sequence divergence between the T4 and RB69 DNA polymerases, but the closest interacting residues, E716-G717 in the RB69 DNA polymerase and E713-D714 in the T4 DNA polymerase, are flanked by conserved residues (Fig. 4 d and e). Residues E716 and G717 in the RB69 DNA polymerase reside in the loop region of a β hairpin structure (Fig. 4e), which connects with a patch of highly conserved basic residues, K705, K706, and R707, that interact with the primer template region. Movement of the fingers domain brings residue L490 in the RB69 DNA polymerase to about 6.7 Å from residue G717 in the closed structure (Fig. 4e) and to about 13 Å in the open structure (Fig. 4d). We speculate that the F487L substitution in the T4 DNA polymerase suppresses the *optA1*-sensitivity conferred by the P424L substitution by a mechanism that links movements of the fingers domain to interactions that affect DNA binding by the patch of conserved basic residues.

Mechanism for suppressing the *optA1*-sensitivity conferred by the R335C amino acid substitution; regulating movement of the fingers domain

The *optA1*-sensitivity of the R335C-DNA polymerase is suppressed by the P424L substitution, which also confers *optA1*-sensitivity and antimutator phenotypes. The ability of the R335C and P424L substitutions individually to confer *optA1*-sensitivity and antimutator phenotypes but together to suppress both phenotypes suggests that suppression is achieved by a physical link between the two amino acid substitutions. The fingers domain is the most reasonable physical connection between residues R335 and P424 based on structure (Fig. 4b and c). Residue R338 in the RB69 DNA polymerase (R335 in the T4 DNA polymerase) forms part of a platform that cradles helix P in the RB69 DNA polymerase closed complex (Fig. 4c), but is separated from helix P in the open complex (Fig. 4b). Thus, the fingers domain moves with respect to R338 during cycles of nucleotide incorporation. Residue P424 in the NPL core is located near the base of the fingers domain (Fig. 4d and e). We propose that the R335C and P424L substitutions individually disrupt the normal positioning and movement of the fingers domain to produce DNA polymerase complexes with reduced stability, but together both substitutions correct positioning to restore function. These observations along with the identification of several amino acid substitutions in the fingers domain that suppress the *optA1*-sensitivity conferred by the P424L substitution indicate a functional relationship between NPL core residues and the fingers domain.

Discussion

Using suppressor analysis to determine how the bacteriophage T4 DNA polymerase forms polymerase complexes that remain stable when dNTP pools are low; identification of the NPL core

Suppressor analysis or “forced evolution” is a powerful genetic method to uncover interactions between amino acid residues and structures that are important for function; this method is especially useful in revealing interactions that are well removed from the active site. We propose that suppressor analysis can be used to uncover the mechanism used by the phage T4 DNA polymerase to form stable polymerase complexes that can replicate long stretches of DNA even when dNTP pools are low. Two T4 DNA polymerase phenotypes were used: *optAI*- and PAA-sensitivity, which are correlated with mutant DNA polymerases that form polymerase complexes with decreased or increased stability, respectively. Since DNA replication by *optAI*-sensitive T4 DNA polymerases is severely restricted under conditions where the dGTP pool is reduced, complexes formed with these mutant DNA polymerases dissociate prematurely while waiting for the next available dGTP. Selection for *optAI*-resistant phage reveals second-site mutations that encode amino acid substitutions that suppress *optAI*-sensitivity by either correcting the defect caused by the first amino acid substitution or by generally increasing the stability of polymerase complexes, as observed for the L412M substitution in the polymerase active site. The L412M substitution also confers PAA-sensitivity (Fig. 3). Amino acid substitutions that confer *optAI*- and PAA-sensitivity were identified in a cluster of residues that we have named the NPL core (Fig. 1); we propose that the NPL core functions in regulating the stability of polymerase complexes. Residues in the tripartite NPL core are partially conserved in family B DNA polymerases, especially the proline residue in the linker region (Fig. 2).

We were first drawn to the NPL core because of identification of the I50L substitution as a global suppressor of the *optAI*-sensitivity of several DNA polymerases.¹⁸ The I50L substitution resembles the L412M substitution by conferring sensitivity to the antiviral drug phosphonoacetic acid (PAA) (Fig. 3) and a mutator phenotype (Table 1); thus, we predict that the I50L substitution increases the stability of polymerase complexes as observed for the L412M substitution.^{3,4,12} Additional amino acid substitutions in the NPL core were identified as suppressors of the *optAI*-sensitivity of the R335C-DNA polymerase – P378L and P424L (Table 1). While a leucine substitution for P378 had only a small effect on function, the P424L substitution conferred *optAI*-sensitivity and an antimutator phenotype (Table 1). Since *optAI*-sensitivity and antimutator phenotypes are observed for amino acid substitutions that reduce the stability of polymerase complexes, as observed for the I417V and A737V amino substitutions,³⁻⁵ we predict that the P424L substitution reduces stability. This proposal is supported by the ability of the L412M substitution to suppress the *optAI*-sensitivity conferred by the P424L substitution. Thus, amino acid substitutions that apparently increase (I50L) and decrease (P424L) the stability of polymerase complexes were identified in the NPL core.

If residues within the NPL core interact during DNA replication, then substitutions for NPL core residues should be identified that suppress the *optAI*-sensitivity conferred by the P424L substitution. This proposal was confirmed. Three “compensating” amino acid substitutions were identified in the N-terminal loop – D49Y, I50T, a duplication of residue I50, which was accompanied by the A580S substitution, and P378H was identified in the linker region (Tables 1 and 2).

Since amino acid substitutions within the NPL core confer *optAI*-sensitivity or PAA-sensitivity, as observed for amino acid substitutions within the polymerase active site, the NPL core may be acting as a second control center in regulating the stability of polymerase

complexes. As already discussed, the L412M substitution was selected as a strong suppressor of the *optAI*-sensitivity conferred by the P424L substitution (Table 2). But what happens if the L412M substitution in the polymerase active site is combined with the I50L substitution in the NPL core? Since individually both amino acid substitutions confer PAA-sensitivity and mutator phenotypes, if residues in the polymerase active center and the NPL core function independently then increased PAA-sensitivity and more replication errors are expected for the double mutant compared to the single mutants. We attempted to construct the double mutant, but synthetic lethality was observed (Table 3). Thus, while amino acid substitutions that apparently increase (L412M) and decrease (P424L) the stability of polymerase complexes can “balance” each other, the combination of two amino acid substitutions that confer the same phenotypes produces a DNA polymerase that cannot function *in vivo*. Lethality may be the consequence of too many mutations (error catastrophe) or, if both amino acid substitutions act to increase the stability of polymerase complexes, then translocation may be impeded. Both possibilities can be tested in future *in vitro* experiments.

Networks of amino acid interactions that connect the NPL core to DNA

If the NPL core contributes to determining the stability of replicating complexes, how can this be done without contacting DNA? Suppressor analysis revealed apparent networks of interactions that connect NPL core residues with DNA in the primer-terminal region via the conserved Y/FxGG/AxV motif (colored red in Fig. 4a) and possibly to the conserved patch of basic residues (Fig. 4d and e). Since many of the amino acid substitutions that suppress the *optAI*-sensitivity conferred by the P424L substitution are predicted to affect movement of the fingers domain (E98K, R335C, S379A, G466S, Q478H/R, F487L, and L567F), we propose that the NPL core is in position to relay movements of the fingers domain to residues that interact with DNA in the primer-terminal region.

Strong evidence for the link between the NPL core and fingers domain is provided by the ability of the P424L substitution in the NPL core to suppress the *optAI*-sensitivity conferred by the R335C substitution even though the P424L substitution also confers *optAI*-sensitivity. Thus, these amino acid substitutions are not acting independently, but appear to be connected by the fingers domain. Since helix P of the fingers domain makes close contact with R338 in the RB69 DNA polymerase (corresponds to R335 in the T4 DNA polymerase) in the closed structure (Fig. 4c), the R335C substitution in the T4 DNA polymerase may disrupt positioning of the fingers domain, which appears to be restored by the P424L substitution. The L340P substitution (L343 in the RB69 DNA polymerase) provides additional evidence that positioning of the fingers domain in the closed structure affects function (Fig. 4c). The L340P substitution confers PAA-sensitivity and a strong mutator phenotype,^{20,38} which are opposite to the *optAI*-sensitivity and antimutator phenotypes conferred by the R335C substitution. Thus, amino acid substitutions that perturb interactions with helix P in the fingers domain can either decrease (R335C) or increase (L340P) the ability of the DNA polymerase to form stable polymerase complexes when dGTP is in short supply.

Model

We propose the following model to explain the *optAI*-sensitivity/antimutator and PAA-sensitivity/mutator phenotypes and to provide insights into how the T4 DNA polymerase forms stable polymerase complexes that can withstand low dNTP pools but still retain the ability to proofread mismatches.

One of the critical steps in determining the fidelity of DNA replication occurs just following nucleotide incorporation. If an incorrect nucleotide is incorporated at the primer terminus,

then proofreading ensues which requires separation of the end of the primer strand from the template and repositioning of the primer end within the exonuclease active site; these steps may require enzyme dissociation.¹ Incorporation of correct nucleotides, however, is the norm. DNA polymerases are proposed to form polymerase complexes that are in rapid equilibrium between the untranslocated and translocated positions until the correct dNTP is bound in the nucleotide binding pocket that is formed in the translocated complex.^{22,23} Initial binding of the correct nucleotide is rapid⁴ and is followed by a slower conformational change to form a closed complex which provides increased nucleotide discrimination.³⁹ But what happens if the correct nucleotide is not available? One possibility is proofreading even though the primer terminus is correct. *optAI*-sensitive/antimutator T4 DNA polymerases degrade correctly replicated DNA even in the presence of dNTPs and degradation is exacerbated by low dNTP pools.^{1,5} In contrast, PAA-sensitive/mutator T4 DNA polymerases proofread fewer mismatches and do not degrade correctly replicated DNA as readily when dNTP pools are low.³ Since these mutant DNA polymerases are PAA-sensitive and PAA is a PPi (pyrophosphate) analog that binds to the untranslocated complex, PAA-sensitive DNA polymerases are predicted to favor formation of untranslocated complexes. Thus, enhanced stability of polymerase complexes is correlated with formation of PAA-sensitive, untranslocated polymerase complexes.

Previous studies have focused on interactions within the polymerase active site, but studies presented here indicate that residues in the NPL core play an important additional role that is linked to movement of the fingers domain and to conserved residues that bind DNA. DNA polymerase dissociation likely requires release of DNA contacts in the polymerase active center as well as release of contacts that are mediated via the NPL core. The additive interactions observed for amino acid substitutions in the NPL core and in the polymerase active site suggests that both sites function in DNA binding, but independently of each other. If the NPL core mediates DNA contacts between the Y/FxGG/AxV motif and possibly the patch of conserved basic residues (Fig. 4), then DNA association during translocation may be maintained by alternating DNA binding with NPL-mediated contacts and contacts within the polymerase active site.

We have made a movie that illustrates the relationship between the open and closed forms of the RB69 DNA polymerase as it cycles between rounds of nucleotide incorporation (Supplemental Materials). While the movie does not show the entire reaction cycle, it clearly demonstrates the spatial arrangements of the amino acid substitutions described in this manuscript as the fingers domain swings back and forth between both conformations. The relationship between the NPL core and the NPL interaction network is shown. The movements in the NPL hinge are subtle, but rapid and subtle conformation changes are predicted while the polymerase translocates along the DNA template.

Implications for other family B DNA polymerases and future directions

Two types of amino acid substitutions in the phage T4 DNA polymerase are presented: substitutions that confer *optAI*-sensitivity and antimutator phenotypes or PAA-sensitivity and mutator phenotypes. These opposite phenotypes, however, represent two sides of the same coin. While enhanced ability to form stable polymerase complexes is an advantage when dNTP pools fluctuate, the down side is less proofreading and more replication errors. Conversely, low stability means premature dissociation when dNTPs are in short supply, which increases proofreading but also reduces replication efficiency. We propose that DNA polymerases in different organisms have optimized the balance between forming stable polymerase complexes and proofreading to meet specific environmental conditions. For example, the PAA-sensitivity of the wild type herpes viral DNA polymerase reflects increased formation of the untranslocated complex, which may be necessary if dNTP pools fluctuate. Archaeal DNA polymerases that replicate DNA at high temperature may also

favor formation of stable polymerase complexes. In contrast, the T4 and RB69 DNA polymerases replicate DNA under moderate temperatures and high dNTP concentrations.⁴⁰ Since dNTPs are consistently available, the T4 and RB69 DNA polymerases may form less stable polymerase complexes in order to reduce mismatch extension and increase proofreading.

Suppressor analysis identified several amino acid substitutions for conserved residues in family B DNA polymerase, for example P358 and P424 in the NPL core (Fig. 2), G466 in helix N (Fig. 4a) and A580 in helix Q (Fig. 5a), which demonstrates similarities among family B DNA polymerases. Several amino acid substitutions were also identified for non-conserved amino acids and these may provide insights into differences in function. For example, three amino acid substitutions that suppressed the *optAI*-sensitivity conferred by the P424L substitution were identified that affect interaction between residues Y460 and N579 in the T4 DNA polymerase (Fig. 5b). The phenylalanine and histidine substitutions for Y460 and the leucine substitution for N579 are predicted to disrupt H-bonding interactions. Since the Y460F/H and N579L substitutions counter the *optAI*-sensitivity and apparent decreased stability conferred by the P424L substitution, these amino acid substitutions are predicted to increase the stability of polymerase complexes. Residues corresponding to Y460 and N579 in other family B DNA polymerases are replaced by hydrophobic residues, frequently phenylalanine for Y460 and a variety of amino acids for N579, which may indicate that polymerase complexes formed with these DNA polymerases are intrinsically more stable. Interestingly, *E. coli* DNA polymerase II, which has increased mismatch extension ability,³¹ has a phenylalanine and a leucine for residues that correspond to Y460 and N579 in the T4 DNA polymerase; Y460F and N579L were identified as *optAI*-suppressors in the T4 DNA polymerase (Fig. 5b). However, since high fidelity DNA polymerases also have a phenylalanine instead of the Y460 observed in the T4 DNA polymerase and a hydrophobic amino acid instead of N579, it is unlikely that the phenylalanine and leucine substitutions in *E. coli* DNA polymerase II alone are responsible for translesion synthesis activity. But these substitutions may combine with other subtle differences to allow mismatch extension ability.³¹ Similarly, subtle differences are predicted to be the basis for DNA polymerase drug-sensitivity or - resistance.

These studies provide evidence that the NPL core is conserved in family B DNA polymerases and that the NPL core plays a role in regulating the stability of polymerase complexes via interactions that involve the fingers domain and conserved residues that bind DNA. The networks will have to be expanded in future experiments to include the thumb domain since the *optAI*-sensitivity and antimutator phenotypes of the A737V-DNA polymerase are suppressed by the I50L substitution in the NPL core and by the L412M substitution in the polymerase active site,¹⁸ but it is not clear how the A737V substitution produces *optAI*-sensitivity. The collection of amino acid substitutions reported here can also be used to test and to further develop computational methods that have been used to probe the dynamics of DNA and RNA polymerases.⁴¹

Materials and Methods

Databases, structure-based sequence alignments, and DNA polymerase dynamics

The protein sequence database at the National Center for Biotechnology Information (NIH) was the source for the family B DNA polymerase sequences discussed here. Family B DNA polymerase structures were visualized using the applications Spdb-viewer⁴² and PyMOL.⁴³ CLUSTALX⁴⁴ was used to align protein sequences; structurally equivalent regions were manually put into the CLUSTAL program as distinct alignment regions.

A movie describing the relationships between open and closed forms of the RB69 DNA polymerase is provided. The open, binary complex of the RB69 DNA polymerase with DNA containing a templating abasic site, PDB ID:1Q9X³³ and the closed, ternary complex with DNA containing a templating A and an incoming dTTP, PDB ID: 1IG9²⁴ were aligned via their palm domains (residues 383-468 and 573-729). The morph between open and closed forms of the DNA polymerase was calculated with the program LSQMAN.⁴⁵ The abasic site in the open complex was replaced with an adenine, the dTTP from the ternary complex was removed, and residues 505-534 of the fingers domain of both complexes were removed to aid in clarity. Individual movie frames were rendered in PyMOL⁴³ and assembled in QuickTime Pro (Apple Inc, Cupertino, CA).

Bacterial and bacteriophage T4 strains

The *Escherichia coli* strain CR63 (K strain, *supD*) was used to prepare T4 phage cultures using standard procedures. Determinations of *rII*⁺ mutant frequencies were made using the CR63 lambda lysogenic strain (CR63λ). The *optA*⁺(2395) and *optA1* strains (2396)⁹ were used to detect *optA1*-sensitivity. The bacteriophage T4 strains are derived from strain T4D.

Bacteriophage T4 methods

Replication fidelity was determined for wild type and mutant phage T4 strains by measuring revertant frequencies for the *rII* mutant *rIIUV199oc* to *rII*⁺. The total number of phage per culture was measured by plating on the permissive host CR63 and the number of *rII*⁺ revertants was determined by plating on the host CR63λ, which restricts growth of *rII* mutant T4 phage. The revertant frequency is the ratio of *rII*⁺ revertants to the total number of phage. At least 8 parallel cultures were tested for each experiment; the median revertant frequency is reported.

optA1-sensitivity was determined by measuring the plating efficiency of phage T4 strains on the *optA1* host (2396) and the isogenic *optA*⁺ host (2395). At least 5 cultures were tested; the average plating efficiency is reported.

optA1-resistant phage were selected by plating phage from independent, high-titer cultures (>10¹¹ phage/ml) on the nonpermissive *optA1* bacterial host 2396. The plates were incubated overnight at 42 °C (most stringent condition) or 30 °C (less stringent condition). Phage plaques were produced only if the infecting phage could replicate genomic DNA and produce phage progeny under the low dGTP condition. *optA1*-resistant phage were detected at a frequency of about one in 10 to 100 million. Only one (for the most part) or two *optA1*-resistant phage isolates were selected from each culture for further study.

The T4 DNA polymerase gene (*g43*) of each isolate was sequenced to identify the mutation that conferred *optA1*-resistance. Phage T4 genomic DNA was prepared from 0.75 ml of a frozen/thawed phage culture (~ 10¹¹ phage) using the standard phenol/chloroform extraction method. The T4 DNA polymerase gene was amplified using the following primers: F 5'-GCCTAATAACTCGGGCTATAAACTAAGG and R 5'-GGGACCTGGAGGTCCTAG. The PCR product was purified by gel electrophoresis and sequenced using the F and R primers plus 3 additional primers that extended in the reverse direction: 5'-GCGATTTGCTGAACCGCATC, 5'-TTCTGGTGTAAATGAGTCCTAT, and 5'-GGCGCAGGGTCTTGTTCAAC.

Mutant T4 DNA polymerase strains were constructed by recombination between co-infecting mutant phage strains or between the infecting phage with cloned copies of mutant DNA polymerase genes carried on plasmids within the host bacteria.⁴⁶ The T4 DNA polymerase expression vector⁴⁷ was mutated using a variety of site-directed mutagenesis procedures. The single mutagenic-primer method⁴⁸ was used to construct the P424L-DNA

polymerase using the primer: 5'-CGCCAGGTTAACATTAGTCTTGAACTATTCGTGGTCAG. The mutagenic base is underlined. The DNA polymerase gene was sequenced to confirm the presence of all mutations.

Sensitivity to the antiviral drug phosphonoacetic acid (PAA) was determined by spotting phage (10 µl) onto a PAA gradient that was overlaid with soft agar containing host bacteria (CR63). The PAA gradient plates were formed by first pouring 35 ml of an agar solution containing 2 mg/ml PAA into a 100 × 100-mm plate and then elevating one edge with a pencil. After the agar hardened, the plate was placed flat on the bench and 35 ml of drug-free agar solution was added. The plates were used the following day. The agar solution contained in 100 ml 1.3 g tryptone, 0.8 g NaCl, 0.1 g Na citrate, 1.2 g Noble agar, and 200 mg PAA for PAA-agar and no PAA for drug-free agar.

Supplementary Material

Refer to Web version on PubMed Central for supplementary material.

Acknowledgments

We thank M. Bryman for the initial selection of the R335C/P378L- and R335C/P424L-DNA polymerase strains. This work was supported from an operating grant from the Natural Sciences and Engineering Research Council of Canada to L.R.-K. M.H. was supported by PHS grant CA-52040 from the National Cancer Institute. L.R.-K. is a Scientist of the Alberta Heritage Foundation for Medical Research.

Abbreviations used

dGTP	deoxyguanosine triphosphate
dNTP	deoxynucleoside triphosphate
PAA	phosphonoacetic acid
PCNA	proliferating cell nuclear antigen
<i>optA1</i> allele	increased dGTPase reduces the pool of dGTP available for DNA replication

References

1. Reha-Krantz LJ. DNA polymerase proofreading: multiple roles maintain genome stability. *BBA Proteins and Proteomics*. 2009 doi:10.1016/j.bbapap.2009.06.012.
2. Bloom LB. Loading clamps for DNA replication and repair. *DNA Repair*. 2009; 8:570–578. [PubMed: 19213612]
3. Reha-Krantz LJ, Nonay RL. Motif A of bacteriophage T4 DNA polymerase: role in primer extension and DNA replication fidelity. *J. Biol. Chem*. 1994; 269:5635–5643. [PubMed: 8119900]
4. Fidalgo da Silva E, Mandal SS, Reha-Krantz LJ. Using 2-aminopurine fluorescence to measure incorporation of incorrect nucleotides by wild type and mutant bacteriophage T4 DNA polymerases. *J. Biol. Chem*. 2002; 277:40640–40649. [PubMed: 12189135]
5. Spacciapoli P, Nossal NG. A single mutation in bacteriophage T4 DNA polymerase (A737V, *tsL141*) decreases the processivity as a polymerase and increases its processivity as a 3'→5' exonuclease. *J. Biol. Chem*. 1994; 269:438–446. [PubMed: 8276833]
6. Spacciapoli P, Nossal NG. Interactions of DNA polymerase and DNA helicases within the bacteriophage T4 DNA replication complex. *J. Biol. Chem*. 1994; 269:447–455. [PubMed: 8276834]
7. Hacker KJ, Alberts BM. The slow dissociation of the T4 DNA polymerase holoenzyme when stalled by nucleotide omission. *J. Biol. Chem*. 1994; 269:24209–24220. [PubMed: 7929077]

8. Hacker KJ, Alberts BM. The rapid dissociation of the T4 DNA polymerase holoenzyme when stopped by a hairpin helix. *J. Biol. Chem.* 1994; 269:24221–24228. [PubMed: 7929078]
9. Gauss P, Doherty DH, Gold L. Bacterial and phage mutations that reveal helix-unwinding activities required for bacteriophage T4 replication. *Proc. Natl. Acad. Sci. USA.* 1983; 80:1669–1673. [PubMed: 6300866]
10. Meyers JA, Beauchamp BH, Richardson CC. Gene 1.2 of bacteriophage T7. Effect on deoxynucleotide pools. *J. Biol. Chem.* 1987; 262:5288–5292. [PubMed: 3549718]
11. Seto D, Bhatnager SK, Bessman MJ. The purification and properties of deoxyguanosine triphosphate triphosphorhydrolase from *Escherichia coli*. *J. Biol. Chem.* 1988; 263:1494–1499. [PubMed: 2826481]
12. Hariharan C, Reha-Krantz LJ. Using 2-aminopurine fluorescence to detect bacteriophage T4 DNA polymerase-DNA complexes that are important for primer extension and proofreading reactions. *Biochemistry.* 2005; 44:15674–15684. [PubMed: 16313170]
13. Drake J, Allen EF, Forsberg SA, Preparata R-M, Greening EO. Spontaneous mutation. *Nature.* 1969; 221:1128–1132. [PubMed: 4975273]
14. Wang FJ, Ripley LS. The spectrum of acridine resistant mutants of bacteriophage T4 reveals cryptic effects of the *tsL141* DNA polymerase allele on spontaneous mutagenesis. *Genetics.* 1998; 148:1655–1665. [PubMed: 9560385]
15. Reha-Krantz LJ. Learning about DNA polymerase function by studying antimutator DNA polymerases. *Trends Biochem. Sci.* 1995; 20:136–140. [PubMed: 7770910]
16. Drake JW. General mutators are improbable. *J. Mol. Biol.* 1993; 229:8–13. [PubMed: 8421317]
17. Schaaper RM. Antimutator mutants in bacteriophage T4 and *Escherichia coli*. *Genetics.* 1998; 148:1579–1585. [PubMed: 9560377]
18. Stocki SA, Nonay RL, Reha-Krantz LJ. Dynamics of bacteriophage T4 DNA polymerase function: identification of amino acid residues that affect switching between polymerase and 3'→5' exonuclease activities. *J. Mol. Biol.* 1995; 254:15–28. [PubMed: 7473755]
19. Hogg M, Cooper W, Reha-Krantz L, Wallace S. Kinetics of error generation in homologous B-family DNA polymerases. *Nucl. Acids Res.* 2006; 34:2528–2535. [PubMed: 16687658]
20. Reha-Krantz LJ, Nonay RL, Stocki S. Bacteriophage T4 DNA polymerase mutations that confer sensitivity to the PPi analog phosphonoacetic acid. *J. Virol.* 1993; 67:60–66. [PubMed: 8380094]
21. Leinbach SS, Reno JM, Lee LF, Isbell AF, Boezi JA. Mechanisms of phosphonoacetate inhibition of herpes virus-induced DNA polymerase. *Biochemistry.* 1976; 15:426–430. [PubMed: 55273]
22. Reha-Krantz LJ, Wong C. Selection of bacteriophage T4 antimutator DNA polymerases: a link between proofreading and sensitivity to phosphonoacetic acid. *Mut. Res.* 1996; 350:9–16. [PubMed: 8657202]
23. Marchand B, Tchesnokov EP, Götte M. The pyrophosphate analogue foscarnet traps the pre-translocated state of the HIV-1 reverse transcriptase in a Brownian ratchet model of polymerase translocation. *J. Biol. Chem.* 2007; 282:3337–3346. [PubMed: 17145704]
24. Franklin MC, Wang J, Steitz TA. Structure of the replicating complex of a pol α family DNA polymerase. *Cell.* 2001; 105:657–667. [PubMed: 11389835]
25. Hopfner K-P, 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. [PubMed: 10097083]
26. Rodriguez AC, Park HW, Mao C, Beese LS. Crystal structure of a pol α family DNA polymerase from the hyperthermophilic archaeon *Therococcus* sp 9 ° N-7. *J. Mol. Biol.* 2000; 299:447–462. [PubMed: 10860752]
27. Savino C, Federici L, Johnson KA, Vallone B, Nastapoulos V, Rossi M, Pisani FM, Tsernoglou D. Insights into DNA replication: the crystal structure of DNA polymerase B1 from the archaeon *Sulfolobus solfataricus*. *Structure.* 2004; 12:2001–2008. [PubMed: 15530364]
28. Kim SW, Kim DU, Kim JK, Kang LW, Cho HS. Crystal structure of Pfu, the high fidelity DNA polymerase from *Pyrococcus furiosus*. *Internat. J. of Biol. Molecules.* 2008; 42:356–361.
29. Liu S, Knafels JD, Chang JS, Waszak GA, Baldwin ET, Deibel MR Jr, Thomsen DR, Homa FL, Wells PA, Tory MC. Crystal structure of the herpes simplex virus 1 DNA polymerase. *J. Biol. Chem.* 2006; 281:18193–18200. [PubMed: 16638752]

30. Swan MK, Johnson RE, Prakash L, Prakash S, Aggarwal AK. Structural basis of high-fidelity DNA synthesis by yeast DNA polymerase δ . *Nature Struct. Mol. Biol.* 2009; 16:979–986. [PubMed: 19718023]
31. Wang F, Yang W. Structural insight into translesion synthesis by DNA pol II. *Cell.* 2009; 139:1279–1289. [PubMed: 20064374]
32. Shamoo Y, Steitz TA. Building a replisome from interacting pieces: sliding clamp complexed to a peptide from DNA polymerase and a polymerase editing complex. *Cell.* 1999; 99:155–166. [PubMed: 10535734]
33. Freisinger E, Grollman AP, Miller H, Kisker C. Lesion (in)tolerance reveals insights into DNA replication fidelity. *EMBO J.* 2004; 23:1494–1505. [PubMed: 15057282]
34. Hogg M, Wallace SS, Doublié S. Crystallographic snapshots of a replicative DNA polymerase encountering an abasic site. *EMBO J.* 2004; 23:1483–1493. [PubMed: 15057283]
35. Braithwaite D, Ito J. Compilation, alignment, and phylogenetic relationships of DNA polymerases. *Nucl. Acids Res.* 1993; 21:787–802. [PubMed: 8451181]
36. Pisani FM, de Felice M, Rossi M. Amino acid residues involved in determining the processivity of the 3'-5' exonuclease activity in a family B DNA polymerase from the thermoacidophilic archaeon *Sulfolobus solfataricus*. *Biochemistry.* 1998; 37:15005–15012. [PubMed: 9778379]
37. Truniger V, Bonnin A, Lázaro JM, de Vega M, Salas M. Involvement of the “linker” region between the exonuclease and polymerase domains of the ϕ 29 DNA polymerase in DNA and TP binding. *Gene.* 2005; 348:89–99. [PubMed: 15777661]
38. Reha-Krantz LJ. Amino acid changes coded by bacteriophage T4 DNA polymerase mutator mutants. Relating structure to function. *J. Mol. Biol.* 1988; 202:711–724. [PubMed: 3172235]
39. Hariharan C, Bloom LB, Helquist SA, Kool KT, Reha-Krantz LJ. Dynamics of nucleotide incorporation: snapshots revealed by 2-aminopurine studies. *Biochemistry.* 2006; 45:2836–2844. [PubMed: 16503638]
40. Mathews CK, Sinha NK. Are precursors concentrated at replication sites? *Proc. Natl. Acad. Sci. USA.* 1984; 79:302–306. [PubMed: 7043458]
41. Zheng W, Brooks BR, Doniach S, Thirumalai D. Network of dynamically important residues in the open/closed transition in polymerases is strongly conserved. *Structure.* 2005; 13:565–577. [PubMed: 15837195]
42. Kaplan W, Littlejohn TG. Swiss-PDB viewer (deep view). *Briefings in Bioinformatics.* 2001; 2:195–197. [PubMed: 11465736]
43. DeLano, WL. The PyMol Molecular Graphics System. DeLano Scientific; San Carlos, CA, USA: 2002. <http://www.pymol.org>
44. Larken MA, Blackshields G, Brown NP, Chenna R, McGleddigan PA, McWilliam H, Valentin F, Wallace IM, Wilm A, Lopez R. Clustal W and Clustal X version 2.0. *Bioinformatics.* 2007; 23:2947–29948. [PubMed: 17846036]
45. Kleywegt GJ. Use of non-crystallographic symmetry in protein structure refinement. *Acta Cryst.* 1996; D52:842–859.
46. Li, V. M.Sc. Thesis. University of Alberta; Edmonton, Alberta, Canada: 2009. Domain interactions in DNA polymerases.
47. Lin TC, Rush J, Spicer EK, Konigsberg WH. Cloning and expression of T4 DNA polymerase. *Proc. Natl. Acad. Sci. USA.* 1987; 84:7000–7005. [PubMed: 3478676]
48. Salerno JC, Jones r. J. Erdogan E, Smith SME. A single-stage polymerase-based protocol for the introduction of deletions and insertions without subcloning. *Molec. Biotech.* 2005; 29:225–232.

Fig. 1.

Location of the NPL core in the structure of the ternary complex of the phage RB69 DNA polymerase; the figure is adapted from FRANKLIN *et al.*²⁴ (PDB ID: 1IG9).

(a) The five domains of the RB69 DNA polymerase are colored as follows: the N-terminal domain is brown, the exonuclease domain is light blue, the palm domain is green, the N helix of the fingers domain is lime green and the P helix is yellow, and the thumb domain is magenta. The linker peptide is illustrated as a red thread. The DNA substrate is in orange and the deoxynucleosides are depicted as blue sticks. The polymerase and exonuclease active sites are enclosed by black circles. Amino acid residues are indicated for the RB69 DNA polymerase and the corresponding amino acids in the phage T4 DNA polymerase are in parentheses. (b) Enlargement to show the NPL core residues. Residue I52, which resides in a loop structure in the N-terminal domain of the RB69 DNA polymerase interacts with P427, which resides in a loop structure in the palm domain, and with a conserved proline residue, P381 in the linker peptide. The corresponding residues in the phage T4 DNA polymerases are I50, P424, and P378, respectively.

Fig. 2.

Structure-based protein sequence alignments of the NPL core for several family B DNA polymerases. Structural information was used to align protein sequences for the DNA polymerases identified by an asterisk (*). Abbreviations used are as follows: T4, enterbacteriophage T4; RB69, enteriobacteriophage RB69; Phi1, enteriobacteriophage phi1; Aeh1, *Aeromonas* phage Aeh1; P-SSM4, *Prochlorococcus* phage P-SSM4; Syn9, *Synechococcus cyanophage* sny9; Pfu, *Pyrococcus furiosus*; 9N-7, *Thermococcus* sp. 9°N-7; Vent, *Thermococcus litoralis*; DTok, *Desulfurococcus Tok*; KOD, *Pyrococcus kodakaraensis*; Sso, *Sulfolobus solfataricus*; HSV, herpes simplex virus 1; EcoII, *Escherichia coli* DNA pol II; Vac, vaccinia virus; Yα, *Saccharomyces cerevisiae* DNA pol α; Yδ, *Saccharomyces cerevisiae* DNA pol δ; Yε, *Saccharomyces cerevisiae* DNA pol ε; Yζ, *Saccharomyces cerevisiae* DNA pol ζ; Hα, human DNA pol α; Hδ, human DNA pol δ; and Hε, human DNA pol ε. The standard Clustal coloring for amino acid residues is used: GPST – orange; HKR – red; FWY – blue; ILMV – green. A lower case “h” indicates a hydrophobic residue.

Fig. 3.

The T4 I50L- and L412M-DNA polymerases are sensitive to the antiviral drug phosphonoacetic acid (PAA). Soft agar containing host bacteria was overlaid on a PAA gradient from 0 to 2 mg/ml. Phage were spotted across the gradient. Plaques, which are the circular clearings on the bacterial lawn, were observed from 0 to 2 mg/ml PAA for the wild type, P424I- and P378L-DNA polymerases. Smaller and fewer plaques were observed for phage expressing the L412M- and I50L-DNA polymerases; these phage are PAA-sensitive.

Fig. 4.

Locations of amino acid substitutions that suppress the inability of the phage T4 P424L-DNA polymerase to replicate DNA when the dGTP pool is reduced (*optA1*-sensitivity). The open (PDB ID: 1Q9X) and closed (PDB ID: 1IG9) RB69 DNA polymerase structures were adapted from Franklin *et al.*²⁴ and Freisinger *et al.*³³, respectively. Amino acids of interest are depicted in space fill; T4 DNA polymerase residues are indicated in parentheses. The protein domains are colored as in Figure 1. (a) NPL core residues in the closed structure of the RB69 DNA polymerase - residues I52, P424, and P378 (T4 residues I50, P424, and P378) are near a cluster of amino acids identified by suppressor analysis that form an apparent network of interactions that connect to the Y/FxGG/AxV motif in the linker (residues 391-396 in the RB69 DNA polymerase are depicted in red space fill) that interacts with the template strand. Open (b) and closed (c) conformations, respectively, showing movement of the fingers domain – helix N (lime green) and helix P (yellow). Open (d) and closed (e) conformations, respectively, showing the conserved basic residues that interact with DNA and the relative positions of G717 and L490 in the RB69 DNA polymerase (D714 and F487 in the T4 DNA polymerase).

(a)	Motif B					Helix Q				
	K	NS	YG	G	Y	A	T	G		
T4	557	KILINSLY	GALG	NIHFRYYDLRN	NATAITIFG					
RB69	560	KILINSLY	GALG	NVWFRYYDLRN	NATAITIFG					
Pfu	488	KLLANSFY	GYGYAKARWYCKE	CAESVTAWG						
Sso	603	KVFINATY	GVFGAETFPLYAPAV	AEVSTALG						
HSV1	811	KVVCNSVY	GFTGVQHGLLPCIHVA	AATVTTIG						
Yδ	700	KISANSVY	GFTGATV	GKLPCLAISSSV	TAYG					
Ye	825	KVILNSFY	GYVMRKGSRWYSMEM	AGITCLTG						

(b) Interactions in the extended NPL

T4	Y460 (F/H)	N579 (L)
RB69	Y463	N582
Pfu	F442	C510
Sso	V555	V625
HSV1	F763	V833
Yδ	F653	I723
EcII	F459	L515

Fig. 5.

Amino acids identified by suppressor analysis of the T4 DNA polymerase compared to other family B DNA polymerases. (a) Protein sequence alignments in Motif B and helix Q. Sites of amino acid substitutions that suppress the *optA1*-sensitivity of the T4 P424L-DNA polymerase are indicated by an asterisk. The alanine residue in helix Q is conserved, but not the adjacent asparagine. The leucine residue in Motif B is also not conserved. (b) The Y460F/H and N579L substitutions were identified as suppressors of the *optA1*-sensitivity conferred by the P424L substitution in the T4 DNA polymerase. Y460 and N579 are in position to form H-bonds (Fig. 4a). Residues corresponding to Y460 and N579 are not conserved in other family B DNA polymerases.

Table 1
Effects of amino acid substitutions in the NPL core

DNA polymerase	<i>optAI</i> -sensitivity ^a (relative) ^a	Replication fidelity (relative) ^b
Wild type	1	1
I50L	0.7	100
R335C	0	0.01
R335C/P378L	0.7	0.13
R335C/P424L	0.8	0.03
P378L	0.8	5
P424L	0	0.02
P424I	0.8	0.1
I50T/P424L	0.8	0.13
I50L/P424L	0.1	0.2

^a *optAI*-sensitivity is the plating efficiency of T4 strains on the mutant *optAI* bacterial host that expresses elevated levels of a dGTPase compared to the plating efficiency on the isogenic wild type host. The plating efficiency of the wild type T4 strain is 0.85; plating efficiencies relative to the wild type T4 phage strain are reported. A “0” indicates that no plaques were detected on the *optAI* host except for *optAI*-resistant phage that appeared at a frequency of ~1 in 10⁶ to 10⁸.

^b Replication fidelity was determined by measuring the frequency of *rIIUV199oc* revertants as described in Materials and Methods. The revertant frequencies reported are relative to the *rIIUV199oc*⁺ revertant frequency of wild T4 phage, ~1 × 10⁻⁶. A revertant frequency >1 indicates a mutator phenotype and a frequency <1 indicates an antimutator phenotype.

Table 2
Amino acid substitutions that suppress the *optAI*-sensitivity of the T4 P424L-DNA polymerase

Amino acid substitutions in the T4 P424L-DNA polymerase	Corresponding amino acids in the RB69 DNA polymerase	Nucleotide sequence of the suppressor mutations (T4)
D49Y	D51	¹⁴⁵ GAC → TAC
I50 duplication + A580S	I52 duplication + A583	¹⁴⁸ ATC → ATCATC
		¹⁷³⁸ GCT → TCT
I50T	I52	¹⁴⁸ ATC → ACC
E98K	E100	²⁹² GAA → AAA
S369A	S372	¹¹⁰⁵ TCA → GCA
P378H	P381	¹¹³² CCT → CAT
L412M	L415	¹²³⁴ CTG → ATG
P424I	P427	¹²⁷⁰ CTT → ATT
Y460F	Y463	¹³⁵¹ TAT → TTT
Y460H	Y463	¹³⁵¹ TAT → CAT
G466S	G469	¹³⁹⁶ GGT → AGT
Q478H	Q481	¹⁴³² CAG → CAT
Q478R	Q481	¹⁴³² CAG → CGG
F487L	L490	¹⁴⁵⁹ TTC → TTA
L567F	L570	¹⁶⁹⁹ CTT → TTT
N579L	N582	¹⁷³⁵ AAT → CTT

Table 3
Effects of the L412M substitution on the P424L and I50L substitutions in the NPL core

Amino acid substitutions in the T4 DNA polymerase	<i>optA1</i> -sensitivity	Replication fidelity
Wild type	1	1
P424L	0	0.02
L412M	0.8	9
L412M/P424L	0.8	0.1
I50L/L412M	Lethal	Lethal

The same conditions as for Table 1.