A strategically located serine residue is critical for the mutator activity of DNA polymerase IV from Escherichia coli

Amit Sharma, Jithesh Kottur, Naveen Narayanan and Deepak T. Nair*

National Centre for Biological Sciences (NCBS-TIFR), UAS-GKVK Campus, Bellary Road, Bangalore 560065, India

Received August 23, 2012; Accepted February 6, 2013

ABSTRACT

The Y-family DNA polymerase IV or PolIV (Escherichia coli) is the founding member of the DinB family and is known to play an important role in stress-induced mutagenesis. We have determined four crystal structures of this enzyme in its pre-catalytic state in complex with substrate DNA presenting the four possible template nucleotides that are paired with the corresponding incoming nucleotide triphosphates. In all four structures, the Ser42 residue in the active site forms interactions with the base moieties of the incipient Watson–Crick base pair. This residue is located close to the centre of the nascent base pair towards the minor groove. In vitro and in vivo assays show that the fidelity of the PolIV enzyme increases drastically when this Ser residue was mutated to Ala. In addition, the structure of PolIV with the mismatch A:C in the active site shows that the Ser42 residue plays an important role in stabilizing dCTP in a conformation compatible with catalysis. Overall, the structural, biochemical and functional data presented here show that the Ser42 residue is present at a strategic location to stabilize mismatches in the PolIV active site, and thus facilitate the appearance of transition and transversion mutations.

INTRODUCTION

Many prokaryotic organisms respond to environmental stress through activation of the SOS response, which involves upregulation of a number of genes. One major aspect of this response involves enhancement in the frequency at which mutations appear in the genome (1). As evolution operates through the selection of mutations that confer fitness for a specific environment, it is believed that this strategy will ultimately aid in relieving selection pressure imposed by an adverse environment. This phenomenon is sometimes termed as adaptive mutagenesis (2,3).

Error-prone DNA polymerases (dPols) belonging to the Y-family have been implicated in stress-induced mutagenesis (4–10). Escherichia coli possesses two members belonging to the Y-family of DNA polymerases (dPols)—PolIV (gene: dinB) and PolV or UmuCD (genes: umuC and umuD) (4,11,12). The expression of both these enzymes is upregulated during the SOS response. Mutant strains of E. coli lacking these polymerases exhibit substantial reduction in fitness during starvation conditions as compared with wild-type strains (13). In addition to stress-induced mutagenesis, it has been shown that the PolIV is responsible for enhancing the frequency of mutations during the stationary phase. The expression of this enzyme is upregulated during the SOS response (LexA-dependent) and stationary phase (RpoS-dependent); therefore, this enzyme is involved in both stress-induced and spontaneous mutagenesis (1,7,14–17). During the stationary phase, PolIV has been shown to be responsible for as much as 85% of adaptive point mutations (14). Additionally, upregulation of PolIV is sufficient for stress-induced mutagenesis in E. coli and in conjunction with other molecules can lead to the transition of a subpopulation of cells to a hypermutable state (18). Also, transcription of the dinB gene can be induced independent of the SOS response because of β-lactam antibiotics like ceftazidime even in RpoS-deficient strains (19).

A number of groups have shown that PolIV can promote base substitutions and frameshift mutations (20–23). As predicted from its observed role in adaptive mutagenesis, PolIV and homologous molecules do play a role in the appearance of resistance towards antibiotics in bacteria (24–27). It is expected that the ability of PolIV to create mutations has to be regulated to exist within a certain range, as too many mutations can be deleterious and too few will not provide enough templates for natural selection. The structural attributes of the PolIV active site...
that allow calibration of mutagenesis within the correct limits are not known.

We present four crystal structures of PolIV in complex with DNA substrates that present dA, dT, dG and dC at the templating position paired with dTTP, dATP, dCTP and dGTP, respectively. An unusual feature of PolIV evident from the structures is the presence of a serine residue (Ser42) in the active site that forms interactions with the base moieties of all four incipient base pairs. When this Ser residue was mutated to Ala, there was a substantial enhancement in the fidelity of PolIV. In addition, the structure of a ternary complex of PolIV with the A:C mismatch in the active site showed that Ser42 forms direct and water-mediated interactions with the template and incoming nucleotides, respectively. Overall, structural, biochemical and functional studies presented here show that the Ser42 is present at a strategic location to stabilize mismatches in the PolIV active site; thus, this residue plays a critical role in the mutator activity of DNA polymerase IV.

MATERIALS AND METHODS

Cloning, purification and crystallization

The dinB gene was amplified from genomic DNA of Escherichia coli and cloned into pGEX_6P1 vector and expressed using the C41DE3 strain. The GST–PolIV fusion polypeptide with an N-terminal GST-tag was purified by affinity chromatography using GST–sepharose resin (GE Healthcare). The GST-tag was removed by PreScission protease, and the PolIV protein was further purified by gel filtration using a Superdex 200 column. For phasing, selenomethionine-labelled PolIV was prepared using B834 cells, and the labelled protein was purified and crystallized using identical protocols as the native protein. Purified oligonucleotides were purchased from Sigma Genosys and annealed to provide duplex DNA substrates. In addition, oligonucleotide with dA at the templating position and a dideoxy nucleotide at the 3'-end were purchased from Keck Centre (Yale University), purified using ion exchange chromatography, desalted, lyophilized, solubilized and annealed. The ternary complexes were reconstituted by mixing PolIV (0.3 mM) with non-terminated DNA at a 1:1:2 molar ratio followed by addition of 5 mM dNMPnPP (Jena Biosciences) and 5 mM MgCl2. In the case of chain termination DNA, dCTP (GE Amersham) was added to a final concentration of 5 mM. After screening and optimization, the best crystals were obtained in 0.1 M acetate (pH 4.8) and 5–12% (w/v) MPD.

Structure determination and crystallographic refinement

X-ray diffraction data were collected at the BM14 beamline of ESRF (PolIV<sub>IG4:CTP</sub>, SeMet–PolIV<sub>IG4:CTP</sub> and PolIV<sub>GT:AMPnPP</sub>) and the PXIII beamline of SLS (PolIV<sub>IG4:GMPnPP</sub>, PolIV<sub>IG4:ATMPnPP</sub> and PolIV<sub>DA:ACTP</sub>). The structure of the SeMet–PolIV<sub>IG4:CTP</sub> complex was determined by multiwavelength anomalous dispersion method using the SHELEX suite of programs (28). The initial model was built manually using O, and the available structures of MsPolIV (4DEZ) and that of the Polymerase associated domain (PAD) region of PolIV (1UNN) were used to guide the model building (29–31). There are two complexes in the asymmetric unit based on the chain ID of the protein, template strand, primer strand, incoming nucleotide and ion, the two complexes were named ABCDE and FGHJ. The other structures were determined by molecular replacement using the structures of SeMet–PolIV<sub>IG4:CTP</sub> as a search model (with template and incoming nucleotide removed). All the structures were refined using CNS until convergence after which TLS refinement was carried out using Refmac in CCP4 (32,33). The presence and conformation of the incoming nucleotide in all the complexes was confirmed through calculation of simulated annealed omit maps. Majority of the residues are in the favourable regions of the Ramachandran plot with only 1% residues in the disallowed regions. The data and refinement statistics are provided in Tables 1 and 2. In all the structures, the electron density map of the FGHJ complex was better than that of the ABCDE complex, and this complex was selected for analysis.

Primer extension assays for wild-type and mutant

The Ser42Ala mutant was generated using the Quickchange kit (Stratagene), and the presence of the mutation was confirmed by sequencing. The mutant enzyme was purified using a protocol similar to that for wild-type (wt) PolIV. The presence of the mutation was also confirmed using mass spectrometry. Primer extension assays to generate fidelity profiles for wild-type and mutant enzymes were carried out as described elsewhere (34). Four template oligonucleotides (TA<sub>1</sub>, T<sub>T</sub>, T<sub>G</sub> and T<sub>C</sub>) were purchased from Sigma Genosys. T<sub>1</sub>, T<sub>T</sub>, T<sub>G</sub> and T<sub>C</sub> were designed to present four different nucleotides A, T, G and C, respectively, at the templating position in the active site. DNA substrates TA<sub>P</sub>, T<sub>T</sub>P<sub>+</sub>, T<sub>G</sub>P<sub>+</sub> and T<sub>C</sub>P<sub>+</sub> prepared by annealing the fluorescently labelled primer P<sub>+</sub> (5'-6-FAM-CTGACTCGAGCAT-3') with each of the four 50mer templates TA<sub>1</sub> (TCTTACCGTGCCCTACC TGAACCTGGTCTCTGCTAAATGCCATACGAGAGATAC G), T<sub>T</sub> (TCTTACCGTGCCCTACGTGAACCTGGTCTC TACGAGAGTACG), T<sub>G</sub> (TCTTACCGTGCCCTACGTGAACCTGGTCTC TACGAGAGTACG) and T<sub>C</sub> (TCTTACCGTGCCCTACGTGAACCTGGTCTC TACGAGAGTACG), respectively. For primer extension assay, the reaction mixture (20 μl) consisted of 5 μM of dNTPs, 20 nM of DNA substrate, 0.1 M ammonium sulphate, 2.5 mM MgCl2, 4 μl of 5X assay buffer (125 mM Tris–Cl, pH 8.0, and 5 mM DTT) and 10 nM of PolIV (wt or mutant). After incubation for 2 h at 37°C, the reaction was terminated by adding 10 μl of stop solution (80% formamide, 1 mg/ml of xylene cyanol, 1 mg/ml bromophenol blue and 20 mM ethylenediaminetetraacetic acid) followed by 2 min of incubation at 95°C. This mixture was immediately transferred to ice for 10 min. In all, 15 μl of this sample was loaded onto a 20% polyacrylamide gel containing 8 M urea, and 1X TBE was used to resolve the reaction products. The gel was pre-run for at
least 60 min before loading the samples. Resolved products on the gel were observed by excitation at 488 nm, and the bands were visualized and recorded using Biorad Pharos FX Plus Molecular Imager. The intensity of the observed bands was quantified using Quantity one, 1D analysis software. The level of incorporation was calculated using the following equation: Percentage incorporation = \(\frac{I_s}{I_s + I_p}\) \times 100 where \(I_s\) = intensity of band that has shifted upwards because of incorporation, and \(I_p\) = intensity of the non-shifted primer band in the same lane. All assays were repeated thrice, and the results were reproducible. The standard error in the measured values of the intensities of the bands was <15%.

**Table 1.** Data collection and refinement statistics (matched base pairs)

<table>
<thead>
<tr>
<th>Data Collection</th>
<th>PolIV(_{dG}:dCMPnPP)</th>
<th>PolIV(_{dG}:dGMPnPP)</th>
<th>PolIV(_{dA}:dTMPnPP)</th>
<th>PolIV(_{dT}:dAMPnPP)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wavelength ((\AA))</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
<td>1.0</td>
</tr>
<tr>
<td>Space group</td>
<td>(P_2)</td>
<td>(P_2)</td>
<td>(P_2)</td>
<td>(P_2)</td>
</tr>
<tr>
<td>Cell ((\AA)) dimensions</td>
<td>86.14, 56.95, 110.81</td>
<td>86.76, 57.04, 110.7</td>
<td>85.75, 57.25, 110.3</td>
<td>86.69, 56.88, 110.93</td>
</tr>
<tr>
<td>Resolution ((\AA))</td>
<td>2.67 (2.81–2.67)</td>
<td>2.33 (2.46–2.33)</td>
<td>2.38 (2.38–2.51)</td>
<td>2.48 (2.61–2.48)</td>
</tr>
<tr>
<td>(R_{sym}) or (R_{merge})</td>
<td>7.4 (52.9)</td>
<td>6.1 (51.4)</td>
<td>5.4 (51.8)</td>
<td>5.9 (52.2)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>99.7 (100)</td>
<td>99.4 (99.5)</td>
<td>99.9 (100)</td>
<td>100 (99.9)</td>
</tr>
<tr>
<td>Redundancy</td>
<td>3.5 (3.5)</td>
<td>3.7 (3.4)</td>
<td>3.6 (3.6)</td>
<td>3.8 (3.8)</td>
</tr>
</tbody>
</table>

**Table 2.** Data collection and refinement statistics (mismatched base pair)

<table>
<thead>
<tr>
<th>Data collection</th>
<th>PolIV(_{dA}:dTCTP)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wavelength ((\AA))</td>
<td>1.0</td>
</tr>
<tr>
<td>Space group</td>
<td>(P_2)</td>
</tr>
<tr>
<td>Cell ((\AA)) dimensions</td>
<td>86.8, 56.98, 111.34</td>
</tr>
<tr>
<td>Resolution ((\AA))</td>
<td>2.32 (2.45–2.32)</td>
</tr>
<tr>
<td>(R_{sym}) or (R_{merge})</td>
<td>10.2 (53.0)</td>
</tr>
<tr>
<td>Completeness (%)</td>
<td>99.2 (99.2)</td>
</tr>
<tr>
<td>Redundancy</td>
<td>3.4 (3.5)</td>
</tr>
<tr>
<td>Resolution ((\AA))</td>
<td>35.2–2.32</td>
</tr>
<tr>
<td>No. of reflections</td>
<td>85474</td>
</tr>
<tr>
<td>(R_{work}/R_{free})</td>
<td>23.4/27.8</td>
</tr>
<tr>
<td>No. of atoms</td>
<td>Protein 5376</td>
</tr>
<tr>
<td></td>
<td>DNA 1356</td>
</tr>
<tr>
<td></td>
<td>dNMPnPP 58</td>
</tr>
<tr>
<td></td>
<td>Ion 4</td>
</tr>
<tr>
<td></td>
<td>Water 93</td>
</tr>
<tr>
<td>R.m.s deviations</td>
<td>Bond lengths ((\AA))</td>
</tr>
<tr>
<td></td>
<td>Bond angles ((^\circ))</td>
</tr>
</tbody>
</table>

**In vivo assay to assess mutator activity of wild-type and mutant PolIV**

Wild-type and mutated \(dinB\) genes were cloned into a modified pET22b(+) vector to generate the plasmids pDJNNdb and pDJNNdbSER42ALA. The PolIV deleted strain \(dinB749\) (del)::kan was obtained from the Coli Genetic Stock Centre (Yale university). Freshly prepared competent cells of this strain were transformed separately with the plasmids pDJNNdb, pDJNNdbSER42ALA and empty plasmid. A single colony for each of these samples was used to inoculate starter cultures at 37°C. Twenty-five millilitres of sterile LB media was prepared for each sample and inoculated with 100 ml of log-phase cells from each starter culture. These cultures were allowed to grow overnight at 37°C after which the cells were pelleted down by centrifugation. The cells were then resuspended in 500 ml of LB media, and 100 ml each of this suspension was plated on five different plates containing LB-agar media with 100 mg/ml of rifampicin. In addition, dilutions were also plated on LB agar without rifampicin to determine viability of the cells. The plates were incubated overnight at 37°C, and the number of colonies obtained were counted and compared. This assay was repeated three times to ensure reproducibility of the observations.
Data deposition

Atomic coordinates and structure factors have been deposited in the PDB with accession codes 4IRC (PolIVdG:dCMPnPP), 4IR9 (PolIVdC:dGMPnPP), 4IR1 (PolIVdA:dTMPnPP), 4IRD (PolIVdT:dAMPnPP) and 4IRK (PolIVdA:dCTP).

RESULTS

Structure of ternary complexes of PolIV

We have determined the structure of four functional complexes of PolIV corresponding to the four possible Watson–Crick pairs. To our knowledge, these are the first structures of a prokaryotic Y-family DNA polymerase captured in the functional state. The four complexes PolIVdG:dCMPnPP, PolIVdC:dGMPnPP, PolIVdA:dTMPnPP and PolIVdT:dAMPnPP crystallized in the space group P2₁ with roughly similar cell constants of \(a = 86.7\ A, b = 56.9\ A, c = 111.3\ A\). The diffraction data could be collected to a maximal resolution of 2.67, 2.33, 2.38 and 2.48 A for the PolIVdG:dCMPnPP, PolIVdC:dGMPnPP, PolIVdA:dTMPnPP and PolIVdT:dAMPnPP complexes, respectively. In all the structures, the asymmetric unit is composed of two complexes. The final refined structure for each complex includes PolIV residues 1–342, nucleotides 1–18 for the template strand, nucleotides 1–14 for the primer strand, incoming dNTP, two Mg\(^{2+}\) ions and water molecules (Figure 1A). The structures show the presence of the four domains characteristic of Y-family dPols, namely, the palm (residues 1–10, 74–165), fingers (11–73), thumb (166–230) and the PAD or little finger (241–341) (Figure 1A). The palm domain houses the catalytic residues—Asp8, Asp103 and Glu104.

In all the complexes, one Mg\(^{2+}\) ion is coordinated in octahedral geometry by the oxygen atoms of \(\beta\)-(2.3 A) and \(\gamma\)-(2.3 A) phosphates of the dNTP molecule and the carboxylates of Asp8 (2.2 A) and Asp103 (2.3 A) in a plane. The coordination sphere is completed by the oxygen of the \(\alpha\)-phosphate (2.2 A) and the main chain carbonyl of Met9 (2.3 A) at the apical positions. This Mg\(^{2+}\) is the equivalent of metal ‘B’ of replicative polymerases and should be involved in the two-metal ion mechanism of DNA synthesis (35). The 3’ oxygen of the primer is present close to the \(\alpha\)-phosphate of the incoming dNTP and is aligned almost linearly with the bond between the \(\alpha\)-phosphate and O3A, suggesting that the enzyme is primed for catalysis and has been captured in the pre-catalytic state. The second Mg\(^{2+}\) equivalent of the metal ‘A’ is located in a position similar to that observed in the functional complexes of human DNA Polβ (36,37).

Active site residues that interact with the incipient base pair

In all four structures, the incipient base pair is present in a cavity and is lined by the residues of three domains—palm, fingers and PAD (Figure 1B). Arg330 of the PAD forms interactions with the phosphodiester backbone of the template nucleotide. The template nucleotide interacts with residues of the fingers domain (Gly33, Arg38, Gly39, Val40, Ser42 and Ala56) and the PAD region (Thr248, Lys291, Pro298 and Arg330). The triad of Arg, Tyr and Lys that is present in all Y-family dPols and serves to stabilize the triphosphate moiety of the incoming nucleotide is also present in PolIV (Tyr46, Arg49 and Lys157) (38). However, unlike other Y-family dPols, the Tyr residue (Tyr46) from this invariant triad is not involved in making contacts with the incoming nucleotide. In addition to Arg49 and Lys157, the residues Met9, Cys11, Phe12 and Thr43 also form polar interactions with the triphosphate moiety of the incoming nucleotide. Overall, the incoming nucleotide is present in the close proximity of residues of the palm (Asp8, Met9, Asp10 and Asp103) and fingers (Cys11, Phe12, Phe13, Ser42, Thr43, Arg49 and Ser55) domains (Figure 1B).

Ser42 forms interactions with the base moieties of the nascent base pair

The side chains of majority of the residues that line the active site form van der Waal interactions with the atoms of the base moieties of the incipient base pair. However, unlike the other residues, Ser42 forms direct and water-mediated polar interactions with the atoms of the bases in the nascent base pair in all four complexes (Figure 2). This residue is located close to the midline of the Watson–Crick base pair towards the minor groove and is, therefore, positioned correctly to interact with the base atoms. In the case of PolIVdG:dCMPnPP complex, the O–H group of the side chain of this residue forms a hydrogen bond with the N2 atom (2.7 A) of the template guanine and the O2 atom (3.3 A) of the incoming nucleotide dCMPnPP (Figure 2A). In the PolIVdC:dGMPnPP complex, this residue forms both direct (3.0 A) and water-mediated hydrogen bonds with the N2 atom of dGMPnPP (Figure 2B). For the PolIVdA:dAMPnPP complex, Ser42 forms direct (2.9 A) and water-mediated interactions with the O2 atom of dAMPnPP (Figure 2C). In addition, Ser42 interacts with the N3 atom of dA through two bridging water molecules. Finally, in the case of the PolIVdT:dAMPnPP complex, Ser42 forms an interaction with the N3 atom (3.2 A) of dAMPnPP (Figure 2D). Overall, it is clear that the Ser42 residue interacts with the bases of the nascent base pair and could influence fidelity of replication.

Sequence analysis of orthologues using the KEGG database shows that a Ser residue at position 42 is conserved in a majority of close prokaryotic homologues of PolIV (Supplementary Figure S1). Also, PolIV structure was superimposed with available structures of other Y-family dPols to ascertain the identity of the residue present at a position equivalent to that of Ser42. A serine residue is present at the equivalent position in the eukaryotic dPol kappa (human; S137). In the only available structure of PolK with undamaged DNA, (PolKdA,dTTTP) S137 is located at a distance of 3.5 Å (Supplementary Figure S2) from the O2 atom of dTTP (39). Serine is not found in an equivalent position in other eukaryotic Y-family dPols, such as 1...
(human; V64) (40), η (human, L48; yeast L60) (36,41) and Rev1 (human, A509; yeast A401) (42,43) and known archaeal dPols Dpo4 (\textit{Sulfolobus solfataricus}, A44) (44) and Dbh (\textit{Sulfolobus acidocaldarius}, A44) (45). Thus, a polar residue is present at this location only in case of orthologues of PolIV and Polk. Additionally, available structures of ternary complexes of dPols belonging to the A, B, C, X and Y families do not show the presence of an active site residue like Ser42 that forms direct and water-mediated polar interactions with the base moieties in all four nascent Watson–Crick base pairs.

The Ser42Ala mutant exhibits higher fidelity than wild-type enzyme

To further probe the role of Ser42 in influencing fidelity, we mutated this residue to Ala. Primer extension assays were carried out to assess the ability of the mutant protein
to incorporate the correct and incorrect nucleotide opposite all four possible template nucleotides in comparison with native PolIV. In the presence of the correct nucleotide, both wt and mutant enzyme showed maximal extension of the primers. For the reaction conditions used, ~80% of the primer was extended on addition of the correct nucleotide in case of all the templates (Figure 3A). However, the wild-type (wt) enzyme exhibited significant misincorporation for a number of different combinations of template and incoming nucleotide (Figure 3A). For template nucleotide dA, the wtPolIV exhibited significant incorporation of dCTP. Also, all three incorrect nucleotides dTTP, dATP and dCTP could be incorporated opposite template dT. Additionally, dATP and dTTP were added opposite template dG. Finally, when the template nucleotide was dC, there was significant misincorporation in the presence of all three incorrect nucleotides (Figure 3A) with maximal extension in case of dTTP (~20%). In comparison, the Ser42Ala mutant exhibited substantially higher fidelity than the native enzyme with no misincorporation opposite the two template nucleotides dA and dG (Figure 3B). Marginal incorporation of dTTP opposite dT was observed and misincorporation opposite dC was considerably reduced and observed only in case of dTTP (~7%). Also, in case of wtPolIV, significant incorporation of dTTP, dATP and dCTP was observed opposite nucleotides 5’ of the correct template nucleotide dA, dT and dG, respectively (Figure 3A). However, in case of the Ser42Ala, these additions were not observed (Figure 3B). Overall, the primer extension assays show that the fidelity of the Ser42Ala mutant is substantially higher than the wtPolIV (Figure 3C).

We carried out an in vivo assay to estimate the mutator activity of Ser42Ala mutant in comparison with wtPolIV. The assay probed for the appearance of
rifampicin-resistant colonies when PolIV was expressed ectopically from a plasmid in a strain wherein the \textit{dinB} gene had been deleted. The bactericidal activity of rifampicin has been attributed to its ability to bind close to the active site of the \textit{\beta}-subunit of RNA polymerase (Rpo) and prevent elongation of the nascent RNA transcripts (46). It has been seen earlier that substitution mutations at key locations in the \textit{rpoB} gene render resistance to \textit{E. coli} against this antibiotic. Consequently, the designed assay is an appropriate tool to evaluate the ability of wtPolIV and the Ser42Ala mutant to generate substitution mutations. The results clearly show that a plasmid bearing the wtPolIV gene gave rise to nearly five-fold the number of colonies observed for a plasmid with the Ser42Ala mutant or for empty vector (Figure 4). Also, in this assay, the number of colonies obtained with the plasmid bearing the mutated \textit{dinB} gene and empty vector were nearly equal. These observations show that the Ser42 residue plays a critical role in the ability of PolIV to create substitution mutations.

\textbf{Structure of PolIV}_{dA:dCTP}

The PolIV\textsubscript{dA:dCTP} complex was crystallized in the same conditions as the other complexes, and X-ray diffraction data could be collected to a maximal resolution of 2.32 Å. The final structure refined to a final R\textsubscript{free} and R\textsubscript{cryst} of 27.8 and 23.4%, respectively (Table 2). The real space

\begin{figure}[h]
\centering
\includegraphics[width=\textwidth]{figure3.png}
\caption{Fidelity profile of wt PolIV and the mutant Ser42Ala. The results for primer extension assay wherein wtPolIV (A) or the Ser42Ala mutant (B) was incubated with DNA duplexes (T\textsubscript{A}P\textsuperscript{*}, T\textsubscript{T}P\textsuperscript{*}, T\textsubscript{G}P\textsuperscript{*} and T\textsubscript{C}P\textsuperscript{*}) and different dNTPs are displayed. (B) 3D plot exhibiting the misincorporation profile of wtPolIV and the mutant. Incorporation was quantitated in the form of percentage of primer extended. The level of incorporation of the incorrect nucleotide is plotted on the z-axis for different template nucleotides (dA, dT, dG and dC; y-axis), for all four incoming nucleotides (dTTP, dGTP and dCTP; x-axis) for the two proteins (wtPolIV and Ser42Ala; ζ-axis).}
\end{figure}
that the Ser42 residue plays a critical role in the ability of PolIV to create substitution mutations.

**DISCUSSION**

A number of studies over the years suggest that replicative dPols ensure high fidelity of DNA synthesis primarily through geometric selection of the 3D shape of Watson–Crick base pairs (49–51). To ensure this, the active site residues will have to generate a 3D mould into which the Watson–Crick base would fit perfectly in an orientation that would facilitate catalysis. Any perturbation in the active site should, therefore, lead to a loss of fidelity. In line with this, mutational analysis of different replicative dPols, such as RB69, Polδ and Polγ, has shown that active site residues generally play a role in ensuring fidelity. Substitution of key active site residues with Ala or Gly results in significant decrease in the fidelity of the enzyme (52–56). In case of extensive studies carried out on RB69, some of the designed mutations did not impair the ability of the polymerase to select the correct nucleotide but allowed stabilization of mismatches in the active site to reduce fidelity. Conversely, in case of *E. coli* PolIV, substitution of Ser42 with Ala enhances fidelity and reduces the ability of the enzyme to promote mismatches. The structures presented in this study suggest that the Ser42 residue interacts with incorrect incoming nucleotides and stabilizes them in the PolIV active site to ensure their addition to the growing primer strand. Hence, the Ser42 residue plays a critical role in the ability of PolIV to create substitution mutations.

Although, Ser42 forms interactions with the base atoms of the four possible incoming nucleotides, wtPolIV exhibits variation in its ability to promote different mismatches. It is seen that PolIV is unable to stabilize purine:purine mismatches with dGTP as the incoming nucleotide in its active site. Additionally, only low level of incorporation is seen for purine:purine mismatches with
dATP as the incoming nucleotide. The C1′–C1′ distance in the case of canonical Watson–Crick base pairs is 10.6 Å. An analysis of non–Watson–Crick base pairs by Leontis et al. (57) shows that the C1′–C1′ distance and profile of hydrogen-bonding between bases can exhibit variation for different mismatches. For purine–purine base pairs, this distance is ~12.5 Å, and this increase in the horizontal extent of the base pairs could give rise to steric clashes with the enzyme residues. These steric clashes would adversely influence stabilization of purine–purine mismatches in the PolIV active site. Overall, it is possible that the observed variability in incorporation of different mismatches could be due to the variation in the destabilizing effects of steric interactions that differentially offset the positive contributions arising from interaction of the incoming nucleotide with the Ser42 residue and variable number of hydrogen bonds formed between the mispaired bases.

It should be noted here that the error profiles reported here show marked differences from that observed in earlier studies. Kobayashi et al. (22) assessed fidelity of replication of the lacZa gene, whereas Wagner and Nohmi (21) did the same for the cII gene from λ-phage. Although both studies showed that frameshift mutations dominate the mutation spectrum, the contribution of substitution mutations was significantly higher in case of replication of the cII gene. Additionally, it was seen that although substitution mutations in case of the LacZa gene were dominated by T–G transition (19 of 66 substitutions) and T–C transition (16 of 66 substitutions), these mutations were not observed in case of replication of the cII gene. Thus, the frequency of different substitution mutations exhibited substantial differences in the two studies. The propensity of PolIV to promote different mismatches reported in the present study is distinct from those observed earlier, suggesting that sequence context in substrate DNA influences the ability of PolIV to stabilize different mismatches. Thus, the mutation spectrum generated by PolIV could be complex, and it is possible that this attribute might ensure improved sampling of the mutation space to aid adaptive mutagenesis.

PolIV is part of the machinery responsible for transient enhancement of mutation rates in subpopulations of a culture of *E. coli* (14,18,58,59). The frequency of mutation during adaptive mutagenesis has to be low enough to prevent cell death of the entire pool because of mutational load and high enough to generate enough number of distinct genomic templates for natural selection. As only 5–25% of primer is extended for the different mismatches, the presence of a Ser residue at the correct location in the PolIV active site could allow calibration of mutagenesis within these required limits. The observations made during the course of this study suggest that Ser42 reduces fidelity to a level that will ensure enough genomic variability for natural selection without compromising genetic viability. The strategically placed Ser42 residue could, therefore, play an important role in the appearance of cells that possess sets of mutations that might allow survival in a stressful environment, and thus relieve selection pressure.

--

**ACCESSION NUMBERS**

4IRC, 4IR9, 4IR1, 4IRD and 4IRK.

**SUPPLEMENTARY DATA**

Supplementary Data are available at NAR Online: Supplementary Figures 1–3.

**ACKNOWLEDGEMENTS**

D.T.N. thanks the X-ray diffraction facility located in the Molecular Biophysics Unit of the Indian Institute of Science [funded by Departments of Biotechnology (DBT) and Science and Technology (DST), Government of India] for facilitating screening and data collection. D.T.N. acknowledges the help rendered by Dr Hassan Belrhali (ESRF) and Dr Babu Manjashetty during data collection at the BM-14 beamline of ESRF and by Meitian Wang and Sandro Waltersperger at the PXIII beamline of SLS. Figures were generated using PyMOL (Schroedinger Inc.). The authors would like to thank Prof. Jayant B. Udgaonkar for critically reading the manuscript.

**FUNDING**

NCBS-TIFR (Department of Atomic Energy); Council of Scientific and Industrial Research (CSIR; Sanction Letter No. 37/((1539)/12/EMR-II); Ramanujan fellowship from the Department of Science and Technology (Government of India) (to D.T.N.); Senior and Junior research fellowship from CSIR (Government of India) (to A.S. and J.K., respectively); Junior research fellowship from the Indian Council of Medical Research (ICMR) (to N.N.). Data collection at the BM14 beamline of ESRF (Grenoble, France) was funded by the BM14 project—a collaboration between Department of Biotechnology, EMBL and ESRF. Funding for open access charge: Extramural grant from the council of Scientific and Industrial Research (CSIR).

Conflict of interest statement. None declared.

**REFERENCES**

[Section of the text]
5114  Nucleic Acids Research, 2013, Vol. 41, No. 9