# 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*[†]

Francesca M. Pisani,*[,‡] Mariarita De Felice,[‡] and Mosè Rossi[‡,§]

*Istituto di Biochimica delle Proteine ed Enzimologia, Consiglio Nazionale delle Ricerche, Via G. Marconi 10, 80125 Napoli, Italy, and Dipartimento di Chimica Organica e Biologica, Università degli studi di Napoli "Federico II", Via Mezzocannone 16, 80134 Napoli, Italy*

*Received May 14, 1998; Revised Manuscript Received August 3, 1998*

ABSTRACT: Herein, we report on the mutational analysis of a 70-amino acid segment (region 1, residues 438−508) of family B DNA polymerase from the thermoacidophilic archaeon *Sulfolobus solfataricus* (Sso DNA pol). Region 1, which lies between the Exo III sequence and the similarity motif D- -SLYP, connects the exonuclease and polymerase domains of Sso DNA pol. Two C-terminally deleted forms of the enzyme, proteins N438 (residues 1−438) and N508 (residues 1−508), were overproduced in the recombinant form and biochemically characterized. They contain the three evolutionarily conserved Exo motifs, but differ in the extent of the C-terminal deletion, since only N508 includes region 1. Both have been found to retain a $Mn^{2+}$-dependent 3′−5′ exonuclease activity, whose thermal stability appears to be increased in comparison to that of the full-sized enzyme. Assays for processive 3′−5′ exonuclease activity, carried out with the heparin trap method on a 24-base oligonucleotide, have revealed that protein N508, as well as the full-length Sso DNA pol, retains a level of processivity of the degradative function substantially higher than that for protein N438. In addition, six site-specific mutations have been introduced at the highly conserved Y-GG/A motif, which has been found within Sso DNA pol region 1. All mutant proteins (Lys491Ile, Tyr495Ser, Lys496Ile, Gly497Ala, and Ala498Val) display increased processivity of their 3′−5′ exonuclease activity, with the exception of protein Tyr495Phe. By a steady-state kinetic analysis of the exonucleolytic reaction on a 24-base oligonucleotide, the above site-specific mutations have been found to affect $K_m$ values consistently with the observed differences in the processivity values, whereas the effect on the $k_{cat}$ values seems to be less important. The results from this mutational analysis indicate that region 1 is involved in determining the processivity of the proofreading function, directly interacting with the nucleic acid substrate.

One of the most interesting questions in studies of the DNA polymerase structure−function relationships concerns the molecular mechanisms responsible for the processive mode of synthesis. Processivity of the polymerase function (the ability of the enzyme to synthesize long stretches of DNA without dissociation from the template) is an intrinsic property of the enzyme molecule by itself or, more often, the result of the interaction with accessory proteins. These latter proteins, known as processivity factors, form multimeric complexes that "clamp" the DNA polymerase onto the primer−template substrate, thus permitting efficient elongation of nucleotide chains (*1*). Although the processive mode of synthesis has been extensively investigated, the interpretation of the experimental data at the molecular level has appeared to be rather difficult. Indeed, a general model for the interaction of nucleotide-polymerizing enzymes with their nucleic acid substrates has not yet been devised (*2*, *3*). The most recent structural analysis of Klenow fragment cocrystallyzed with DNA indicates that the nucleic acid substrate is located in a deep cleft between the polymerase and exonuclease domains (*4*). This cleft is partly formed by a significant movement of the "thumb" subdomain from the extended position occupied in the apoenzyme to encircling the bound DNA molecule. The flexibility of the thumb subdomain has been postulated to be important for primer−template translocation during the synthesis and, therefore, for processivity of the polymerase function. On the other hand, the molecular bases for processive 3′−5′ exonuclease activity have been poorly investigated, although the exonucleolytic reaction mechanism is known in great detail, due to the considerable number of site-specific mutagenesis studies (for a review, see ref *5*) and to the crystallographic analysis of the Klenow fragment (*6*, *7*) and, more recently, of family B DNA polymerase from *Escherichia coli* bacteriophage RB69 (*8*).

[†] This work received financial support in the form of grants from Consiglio Nazionale delle Ricerche (Progetto Finalizzato Biotecnologie) and from the European Union (Contract B104-CT96-0488).

* To whom all correspondence should be addressed: IBPE-CNR, Via G. Marconi 10, 80125 Napoli, Italy. Telephone: ++39-081-7257316. Fax: ++39-081-7257240 (or ++39-081-2396525). E-mail: pisani@dafne.ibpe.na.cnr.it.

[‡] Consiglio Nazionale delle Ricerche.

[§] Università degli studi di Napoli "Federico II".

10.1021/bi981127s CCC: $15.00 
Published on Web 10/01/1998

In this study, this issue has been approached using as a model system a thermostable family B DNA polymerase from the archaeon *Sulfolobus solfataricus* (Sso DNA pol).[1] This enzyme, a monomer of about 100 kDa, possesses 3′−5′ exonuclease and polymerase activities that were demonstrated to reside on structurally independent protein domains (*9*, *10*). On the basis of the results of a partial proteolysis study, we demonstrated that a 70-amino acid segment of the Sso DNA pol polypeptide chain (region 1, residues 438−508), which lies at the connection between the exonuclease and polymerase domains, is structurally flexible, exposed to the solvent on the protein molecule surface, and subjected to conformational changes upon DNA substrate binding, and we also proposed that it could be directly involved in nucleic acid interaction (*11*). Herein we have analyzed the function of region 1 by kinetically characterizing two C-terminal deletion derivatives of Sso DNA pol (proteins N438 and N508), both retaining highly thermostable 3′−5′ exonuclease activity. They differ in the polypeptide chain length, since only protein N508 includes at its carboxyl terminus region 1. In addition, for a more detailed molecular analysis, six site-specific mutant proteins have been produced at the highly conserved "Y-GG/A" sequence motif, which has been found to lie within region 1 of Sso DNA pol. The kinetic parameters ($k_{cat}$ and $K_m$) and the processivity level of the exonucleolytic activity on a single-stranded DNA substrate have been determined for all these mutants and compared with those of the wild type enzyme. This biochemical analysis suggests a role for Sso DNA pol region 1 in nucleic acid binding and in processivity of the degradative function. The results are discussed in light of the recently determined three-dimensional structure of family B DNA polymerase from *E. coli* bacteriophage RB69 (*8*).

## EXPERIMENTAL PROCEDURES

*Materials.* All chemicals were reagent grade. DNA restriction and modification enzymes were from Promega, unless otherwise stated. Oligonucleotides were synthesized by Primm srl (Milan, Italy). All the radioactive reagents were purchased from Amersham Corp.

*Plasmid Construction*. Overexpression of Sso DNA pol in *E. coli* was achieved by subcloning its encoding gene (*polS*) into a pT7 vector derivative under the control of the T7 RNA polymerase Φ10 promoter (*12*). The 882-codon ORF encoding Sso DNA pol was amplified using the plasmid pFC*polS* (*9*) as the DNA template, recombinant *Taq* DNA polymerase, and oligonucleotides EcoATG-for (5′-GGATCTGAATTCTTATGACTAAGCAACTTACCTTA-3′) and HindEnd-rev (5′-CATACCGAAGTAAAGGAAATAGAAGCTTATCATCCAT-3′) as forward and reverse primers, respectively, in a 30-cycle polymerase chain reaction (1 min at 94 °C, 1 min at 55 °C, and 2 min at 72 °C). The above amplification primers were designed to introduce *Eco*RI and *Hin*dIII restriction sites upstream from the initiation codon and downstream from the stop codon of *polS* ORF, respectively. The PCR product, eluted from gel and digested with *Eco*RI and *Hin*dIII, was ligated into *Eco*RI−*Hin*dIII-linearized pT7-7 vector to create the pT7*polS* construct. The cloned fragment was completely sequenced using the T7 DNA polymerase sequencing kit (Pharmacia) to verify that no undesired mutations were introduced during amplification.

The plasmid expressing protein N508 (spanning $Met^{1}$−$Phe^{508}$ of Sso DNA pol) as a fusion with *Schistosoma japonicum* GST (referred to as pGEX-2T/N508) was constructed as follows. Sso DNA pol ORF was PCR-amplified from nucleotide 1 to 1524 with oligonucleotides EcoATG-for and EcoNter-rev (5′-ACGTGAATTCGAATATTCCAGCAGGTGGG-3′) as forward and reverse primers, respectively. These primers introduced an *Eco*RI restriction site at both ends of the amplified fragment. The PCR product was digested with *Eco*RI and subcloned into *Eco*RI-linearized pGEX-2T plasmid (Pharmacia). The nucleotide sequence of the insert was confirmed by sequencing. To construct the recombinant plasmid expressing the fusion protein GST−N438 (pGEX-2T/N438), the *Eco*RI−*Xba*I fragment, which includes Sso DNA pol ORF from nucleotide 1 to 1314 and encodes the polypeptide chain from $Met^{1}$ to $Arg^{438}$, was cut out from the previous construct, blunt-ended, and subcloned into *Sma*I-linearized pGEX-2T vector.

*Site-Directed Mutagenesis.* The mutants presented in this study were prepared by the overlap extension method (*13*) using the polymerase chain reaction. The plasmid pT7*polS* was used as the template for amplification reactions carried out with the Expand high-fidelity PCR system (Boehringer Mannheim), using two end primers (Stu-for and HindEnd-rev) and two complementary mutagenic primers for each site-specific mutation. The sequences of the oligonucleotides used were as follows (mismatch sites for site-directed mutagenesis are underlined): Stu-for, 5′-TGCAAAGGCCTTATTAGGGACATC-3′; K491I-for, 5′-CTAATATAAGAACTTCTGCTCTAATATCGGGAAAAGGATATAAAG-3′; K491I-rev, 5′-CTTTATATCCTTTTCCCGATATTAGAGCAGAAGTTCTTATATTAG-3′; Y495F-for, 5′-GCTCTAATAAAGGGAAAAGGATTCAAAGGCGCAGTAGTTATAGAC-3′; Y495F-rev, 5′-GTCTATAACTACTGCGCCTTTGAATCCTTTTCCCTTTATTAGAGC-3′; Y495S-for, 5′-GCTCTAATAAAGGGAAAAGGATCCAAAGGCGCAGTAGTTATAGAC-3′; Y495S-rev, 5′-GTCTATAACTACTGCGCCTTTGGATCCTTTTCCCTTTATTAGAGC-3′; K496I-for, 5′-CTAATAAAGGGAAAAGGATATATCGGCGCAGTAGTTATAGACCC-3′; K496I-rev, 5′-GGGTCTATAACTACTGCGCCGATATATCCTTTTCCCTTTATTAG-3′; G497A-for, 5′-ATAAAGGGAAAAGGATATAAAGCCGCAGTAGTTATAGACCCACCT-3′; G497A-rev, 5′-AGGTGGGTCTATAACTACTGCGGCTTTATATCCTTTTCCCTTTAT-3′; A498V-for, 5′-GGGAAAAGGATATAAAGGCGTCGTAGTTATAGACCCACCTGCTGG-3′; and A498V-rev, 5′-CCAGCAGGTGGGTCTATAACTACGACGCCTTTATATCCTTTTCCC-3′. Each amplification reaction was performed using 30 cycles of 94 °C for 1 min, 55 °C for 1 min, and 72 °C for 2 min. The final amplified products were cut with *Stu*I and *Hin*dIII and then ligated into pT7*polS* that had been linearized with the same restriction enzymes. The cloned DNA segment was entirely sequenced to confirm the presence of the mutation and to rule out the possibility that replication errors were made during amplification.

*Expression and Purification of Proteins N508 and N438.* Fusion protein GST−N508 (or GST−N438) was purified

[1] Abbreviations: DNA pol, DNA polymerase; Sso, *Sulfolobus solfataricus*; PCR, polymerase chain reaction; GST, glutathione *S*-transferase; ORF, open reading frame; IPTG, isopropyl β-D-1-thiogalactopyranoside; PMSF, phenylmethanesulfonyl fluoride; PAGE, polyacrylamide gel electrophopresis; PVDF, poly(vinylidene difluoride).

from extracts of IPTG-induced *E. coli* TOP 10 cells (Invitrogen), harboring the plasmid pGEX-2T/N508 (or pGEX-2T/N438), by affinity chromatography on a glutathione Sepharose-4B matrix (Pharmacia) according to the manufacturer's instructions. The buffer utilized [10 mM sodium phosphate (pH 7.2), 150 mM NaCl, 10% glycerol, and 0.1% Triton X-100] was supplemented with some protease inhibitors (50 $\mu$g/mL PMSF, 0.2 $\mu$g/mL benzamidine, 1 $\mu$g/mL aprotinin, and 10 $\mu$g/mL soybean trypsin inhibitor). After this chromatographic step, the chimeric protein was cleaved with the protease thrombin and GST was then removed by binding it in batch to the affinity resin. Protein N508 (or N438), collected in the flow-through fraction, was concentrated to about 1 mg/mL using the Centriprep 30 system (Amicon) and stored at −20 °C, after the addition of an equal volume of glycerol. The final yield of the recombinant proteins purified with this procedure was of about 1 mg per liter of bacterial culture.

*Purification of Recombinant Sso DNA Pol and Its Site-Specific Mutant Derivatives.* The plasmid pT7*polS* and its derivatives expressing Sso DNA pol mutant proteins were transformed into *E. coli* BL21(DE3) cells (*12*). Transformed bacteria were grown at 37 °C in a 2 L culture of Luria-Bertani broth (supplemented with 100 $\mu$g/mL ampicillin) to late-logarithmic phase, before adding IPTG to a final concetration of 0.5 mM to induce recombinant protein expression. After 2 h at 37 °C, cells were harvested by centrifugation and the bacterial pellets stored at −20 °C until they were used. The cell paste was defrozen and suspended in 100 mL of buffer A [10 mM Tris/HCl (pH 8.5), 2.5 mM $MgCl_2$, and 100 $\mu$M EDTA] supplemented with some protease inhibitors (50 $\mu$g/mL PMSF, 0.2 $\mu$g/mL benzamidine, 1 $\mu$g/mL aprotinin, and 10 $\mu$g/mL soybean trypsin inhibitor). Cells were broken by two consecutive passages through a French pressure cell apparatus (Aminco Co., Silver Spring, MD) at 2000 psi. Bacterial lysate was clarified by ultracentrifugation for 15 min at 30 000 rpm (Sorvall 50-2Ti rotor) at 10 °C. The supernatant was subjected to thermoprecipitation steps at 55, 60, and 65 °C for 5 min. After each heat treatment, the sample was incubated in ice for 10 min and the precipitated protein was removed by ultracentrifugation for 20 min at 30 000 rpm (Sorvall 50-2Ti rotor). The sample was then applied at a flow rate of 0.4 mL/min to a heparin−Sepharose column (8 cm × 1.5 cm, Pharmacia) equilibrated in buffer A. After an extensive washing step, a 120 mL linear gradient from 0 to 1 M NaCl in buffer A was applied to the column at a flow rate of 1 mL/min. Fractions containing the recombinant protein, centered at 0.75 M NaCl, were pooled and concentrated to a final volume of about 15 mL using the Centriplus 30 system (Amicon). The sample was then dialyzed overnight against 4 L of buffer B [25 mM Hepes/NaOH (pH 7.5) and 50 mM NaCl]. The following day, the sample was chromatographed at a flow rate of 1 mL/min through a Mono S HR 5/5 column using an FPLC system (Pharmacia). The column was washed with buffer B and the elution carried out with a 20 mL linear gradient from 50 to 500 mM NaCl in buffer B. One milliliter fractions were collected, and the recombinant protein was detected by enzymatic assays and denaturing PAGE. The recombinant DNA polymerase was found to elute from 0.2 to 0.3 M NaCl. The corresponding fractions were pooled, concentrated using the Centricon 30 system (Amicon) in storage buffer [10 mM Tris/HCl (pH 8.0), 2.5 mM $MgCl_2$, 2.5 mM 2-mercaptoethanol, and 50% glycerol], and stored at −20 °C.

*Bidimensional Activity Gel Electrophoresis.* Bidimensional activity gel analyses were performed with the procedure of Longley and Mosbaugh (*14*), with the modifications previously described (*11*).

*$NH_2$-Terminal Sequence Analysis.* $NH_2$-terminal sequence analysis of protein N438 and N508 was carried out after electrotransfer onto PVDF membrane ProBlott (Applied Biosystems) from denaturing polyacrylamide gels, as previously described (*11*).

*DNA Polymerase Assay.* Calf thymus activated DNA was used as the substrate for polymerase activity assays, which were carried out at 75 °C, as previously described (*11*).

*3′−5′ Exonuclease Assays.* In assay 1, 3′−5′ exonuclease activity was monitored by measuring the release of [$^{32}$P]-dNMP from a 17-mer oligonucleotide 3′-end labeled with terminal deoxynucleotydil transferase and [$\alpha$-$^{32}$P]dATP, as described previously (*5*). The composition of the reaction mixture (final volume of 10 $\mu$L) was as follows: 50 mM Tris/HCl (pH 8.0), 2.5 mM 2-mercaptoethanol, 2.5 mM $MnCl_2$, 3 $\mu$M labeled oligonucleotide (specific activity of $0.5-1.0 \times 10^4$ cpm/pmol), and the indicated amounts of protein. Assays were carried out at 75 °C in 0.5 mL tubes using a heating block equipped with a thermostated lid to prevent evaporation.

In assay 2, the 24-mer oligonucleotide, 5′-end labeled with [$\gamma$-$^{32}$P]ATP by T4 polynucleotide kinase, was utilized as the substrate (specific activity of $1-5 \times 10^6$ cpm/pmol). The assay temperature and buffer were the same as those for assay 1. Reactions (final volume of 10 $\mu$L) were stopped by the addition of 2.5 $\mu$L of 95% formamide, 20 mM EDTA, 0.05% bromophenol blue, and 0.05% xylene cyanol (stop solution). Samples were then heated at 75 °C for 4 min and subjected to electrophoresis on a 20% polyacrylamide/8 M urea gel. Degradation products were quantitated by phosphorimagery using an Imagequant phosphorimager 445 SI (Molecular Dynamics). The rates for excision of 3′-terminal nucleotides were obtained by calculating the number of catalytic events giving rise to each degradation product, according to the equation reported by Cheng and Kuchta (*15*). The amount of each protein and the incubation times were adjusted to obtain linear conditions. To determine the kinetic parameters ($K_m$ and $k_{cat}$) for the 3′−5′ exonucleolytic reaction, assays at various concentrations of single-strand DNA substrate were carried out. The initial rate value for nucleotide excision was calculated at each substrate concentration; kinetic data were analyzed with Lineweaver−Burk double-reciprocal plots by using the computer program GraFit (*16*).

3′−5′ exonuclease assays on double-stranded DNA were performed using as the substrate 3′-end, $^3$H-labeled activated calf thymus DNA, as previously described (*11*).

*Assay for Processive 3′−5′ Exonuclease Activity.* Processivity of the 3′−5′ exonuclease activity was assayed using the heparin trap method, essentially as described by Reddy et al. (*17*). The indicated amount of each protein was preincubated for 2 min at 75 °C with 5′-end labeled 24-mer oligonucleotide (0.1 pmols) in buffer composed of 50 mM Tris/HCl (pH 8.0) and 2.5 mM 2-mercaptoethanol (reaction volume of 10 $\mu$L). Then, $MnCl_2$ (to start substrate degradation) and heparin (to ensure single-turnover conditions) were

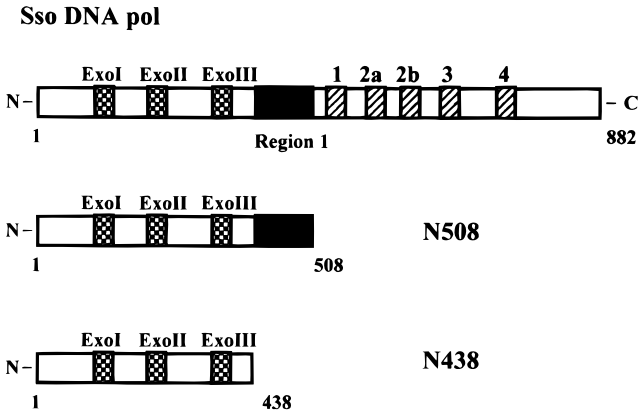


FIGURE 1: Diagramatic representation of the polypeptide chain of Sso DNA pol and its truncated derivatives. Sequence similarity motifs are described by Blanco and Salas (*23*).

added at final concentrations of 2.5 mM and 1 mg/mL, respectively. Incubation was continued for additional 0.5 min and reaction stopped by addition of stop solution (2.5 $\mu$L). In parallel control experiments designed to test the effectiveness of the trap, heparin was added to the assay mixture before the enzyme. Samples were then denatured by incubation at 75 °C for 4 min and subjected to denaturing electrophoresis on a 20% polyacrylamide/8 M urea gel. Gels were autoradiographed on X-ray films.

## RESULTS AND DISCUSSION

*Functional Dissection of Sso DNA Pol Region 1 by Mutagenesis Studies.* The aim of this biochemical analysis was to learn more about DNA polymerase−nucleic acid substrate interaction, in particular concerning the molecular mechanisms responsible for the processivity of the 3′−5′ exonuclease function. In a previous report, the structure of the recombinant Sso DNA pol in free and DNA-bound states was investigated by a limited proteolysis study (*11*). The enzyme polypeptide chain was found to be sensitive to the proteolytic attack within two main locations: a central segment (around residue $Lys^{491}$), located between the exonuclease and polymerase domains (region 1); and the extreme C terminus (region 2). In addition, comparison of the trypsin cleavage maps in the absence and presence of DNA ligands revealed that both these segments of the Sso DNA pol amino acid chain are affected by conformational changes upon nucleic acid binding. Furthermore, only when partial proteolysis experiments were carried out in the presence of DNA, was a proteolytic product of about 8 kDa, which includes region 1, recovered. The appearance of this fragment only after partial tryptic digestion in the presence of DNA suggested its involvement in the nucleic acid binding function so that site occupancy by ligand would exclude access by proteases. All that considered, to thoroughly investigate the role of region 1 in DNA substrate interaction, we initially decided to analyze the biochemical properties of two carboxyl-deleted forms of Sso DNA pol which spanned residues 1−438 (protein N438) and 1−508 (protein N508) of the enzyme polypeptide chain. Both these proteins contain the evolutionarily conserved sequence motifs Exo I−III, which were shown to be critical for the degradative activity in various systems (see Figure 1). Only protein N508 includes at its C-terminal end the 70-amino acid sequence

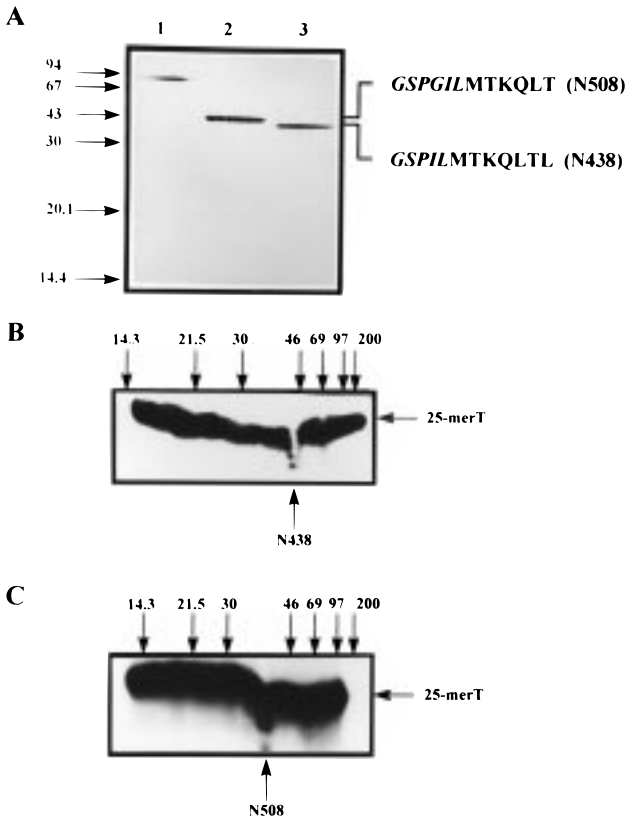


FIGURE 2: Electrophoretic analysis of recombinant Sso DNA pol and its truncated derivatives. (A) Aliquots (about 2 $\mu$g) of purified recombinant proteins Sso DNA pol (lane 1), N508 (lane 2), and N438 (lane 3) were run through a denaturing 10% polyacrylamide gel, which was stained with Coomassie Brilliant Blue. The partial N-terminal sequence of purified proteins N438 and N508 is reported on the right side. The amino acid residues derived from pGEX-2T vector are italic. (B and C) In situ detection of the 3′−5′ exonuclease of proteins N438 and N508, respectively. An aliquot (1 $\mu$g) of purified protein N438 (or N508) was subjected to the bidimensional activity gel analysis, as described in the text. In each panel, the position of protein markers is indicated by arrows.

referred to as region 1, which is located between the Exo III motif and the conserved sequence D- -SLYP (motif 1 in Figure 1). Furthermore, the functional role of region 1 was investigated by a mutational analysis of the Y-GG/A motif, highly conserved among family B DNA polymerases (*18*), that has been found to lie within the above amino acid segment of Sso DNA pol (see Figure 4A).

*Overexpression and Purification of Sso DNA Pol Truncated Forms N508 and N438 Which Retain Highly Thermostable 3′−5′ Exonuclease Activity.* Proteins N438 and N508 were overexpressed in *E. coli* as a fusion with *S. japonicum* GST and purified by affinity chromatography on a glutathione−Sepharose matrix. Figure 2A shows the homogeneous recombinant Sso DNA pol together with its purified derivatives N508 and N438. The identity of these truncated proteins was probed by Western blot experiments using a rabbit polyclonal antiserum raised against the full-sized Sso DNA pol[2] and $NH_2$-terminal sequence analysis after electroblotting from a denaturing gel to a PVDF membrane (see Figure 2). Initially, to verify whether proteins N508 and N438 indeed retained 3′−5′ exonuclease activity, as predicted on the basis of the modular organization model proposed

[2] F. M. Pisani et al., unpublished results.

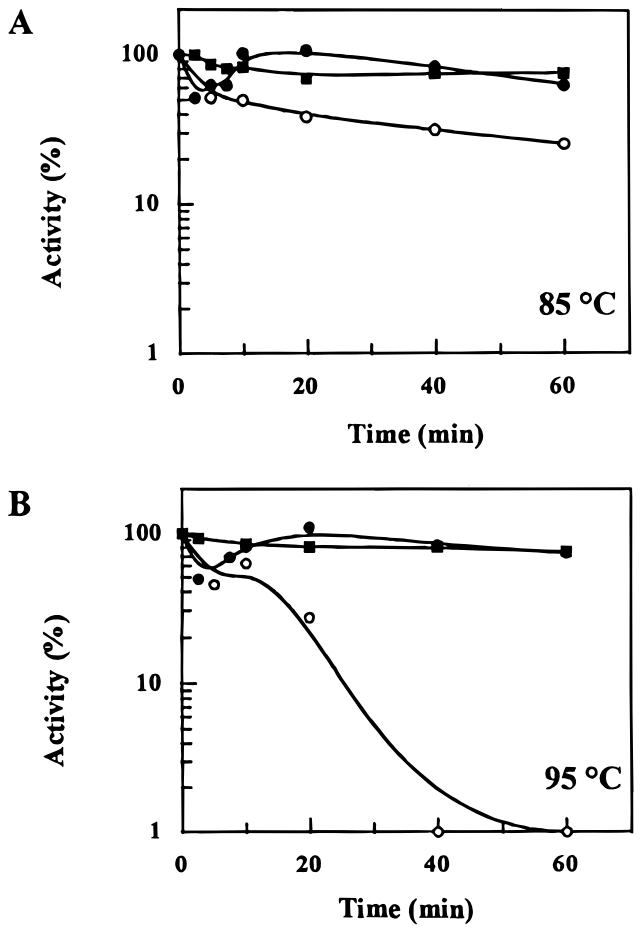


FIGURE 3: Thermal stability of 3′−5′ exonuclease activity of Sso DNA pol and its truncated derivatives. The heat stability was determined by incubating a sample of each protein (0.25 mg/mL) at 85 (A) and at 95 °C (B) in assay buffer without substrate. At each indicated time, aliquots (0.3 pmol of Sso DNA pol, 1.7 pmol of N508, and 3.4 pmol of N438) were withdrawn and tested for residual 3′−5′ exonuclease activity in assay 1, as described in Experimental Procedures. Data reported are mean values of at least three independent experiments within a 20% experimental error. Curves were obtained by interpolating data points with the computer program GraFit (*16*). The following symbols were used: Sso DNA pol (○), N508 (●), and N438 (■).

for Sso DNA pol (*10*), we carried out a bidimensional in situ activity gel analysis (*14*). In this assay, the substrate for the exonucleolytic reaction consists of a $^{32}$P-5′-end-labeled 25-mer oligonucleotide annealed to M13 single-stranded DNA but forming with this template a 3′-terminal (T·C) mispair. Both proteins N438 and N508 (panels B and C of Figure 2, respectively) were found to retain a 3′−5′ exonuclease activity which was strongly dependent on $Mn^{2+}$ ions, in line with previously reported results (*10*).

The thermal stability of Sso DNA pol and its truncated derivatives was monitored at 75, 85, and 95 °C. The 3′−5′ exonuclease activity of each protein was substantially unaffected by incubation at 75 °C for at least 1 h.[2] On the other hand, as shown in Figure 3, the rate of inactivation at 85 °C was larger for Sso DNA pol than for proteins N508 and N438 (panel A); this difference was even more pronounced when the thermal stability was tested at 95 °C (panel B). In this case, the 3′−5′ exonuclease activity of Sso DNA pol dropped to an undetectable level after an incubation of about 40 min, whereas that of proteins N438 and N508 showed only a 20% reduction after 60 min with respect to the initial value. The increased thermal resistance of proteins N508 and N438 in comparison to that of the intact Sso DNA pol was in some way an unexpected result which is likely to depend on a lower stability of the enzyme carboxyl-terminal domain with respect to the amino-terminal one. However, the high thermal resistance features of proteins N508 and N438, taken together with the results of the in situ activity gel assay (see Figure 2), left no doubt that the activity we had monitored was an intrinsic property of these proteins more than the result of some *E. coli* contaminating exonuclease.

*Construction and Purification of Site-Specific Mutants in Region 1 of Sso DNA Pol.* Analysis of the primary structure of Sso DNA pol region 1 has revealed the presence of the Y-GG/A consensus motif, which is highly conserved among family B DNA polymerases (*18*). Multiple sequence alignments indicate that in each family B enzyme the above motif is located between Exo III and D- -SLYP conserved sequences. In Sso DNA pol, the Y-GG/A motif is very close to Lys$^{491}$, which is the site within region 1 of high protease sensitivity in the experiments of partial proteolysis with trypsin, as previously discussed (see Figure 4A; *11*). Therefore, to better analyze the role of region 1 in nucleic acid substrate interaction, amino acid residues at the above consensus sequence were targeted for mutagenesis. As shown in Figure 4, Lys$^{491}$ and Lys$^{496}$ were changed to Ile and Tyr$^{495}$ was changed to Phe and Ser, Gly$^{497}$ to Ala, and Ala$^{498}$ to Val. These conservative changes were introduced to either remove or structurally alter the functional group of the side chains. Site-directed mutagenesis, overproduction, and purification of these mutant proteins were carried out as described in Experimental Procedures. The purification procedure, which includes heat precipitation steps and heparin−Sepharose and ionic exchange chromatographies, yielded highly homogeneous recombinant proteins of about 100 kDa, whose identity was verified by immunoblot experiments with a polyclonal antiserum raised against the native Sso DNA pol.[2] Since all the mutant proteins showed solubility and chromatographic behavior identical to those of the wild type enzyme, we inferred that the conserved mutations did not cause global structural alterations. The initial analysis of the 3′−5′ exonuclease activity of the above mutants on a single-stranded DNA substrate (using assay method 1) indicated that none of these amino acid substitutions had drastic effects on the exonucleolytic catalysis, since all proteins displayed a level of degradative activity similar to that of the wild type enzyme.

*Kinetic Analysis of the 3′−5′ Exonuclease Activity of Sso DNA Pol and Its Mutant Derivatives.* The kinetic characterization of Sso DNA pol and its mutant derivatives was approached with measurement of the 3′−5′ exonuclease processivity. This latter can be defined as the number of nucleotides excised by the enzyme for each binding event to the substrate. Processivity of the 3′−5′ exonuclease is a net result of both the $k_{cat}$ for excision of the next 3′-terminal nucleotide and the dissociation of the enzyme from DNA ($k_{off}$). The processivity assay was carried out at 75 °C using the heparin trap method (*17*). This procedure relies on the strategic use of heparin to trap the enzyme molecules that dissociate from the substrate, thereby preventing rebinding and ensuring that DNA degradation products are derived

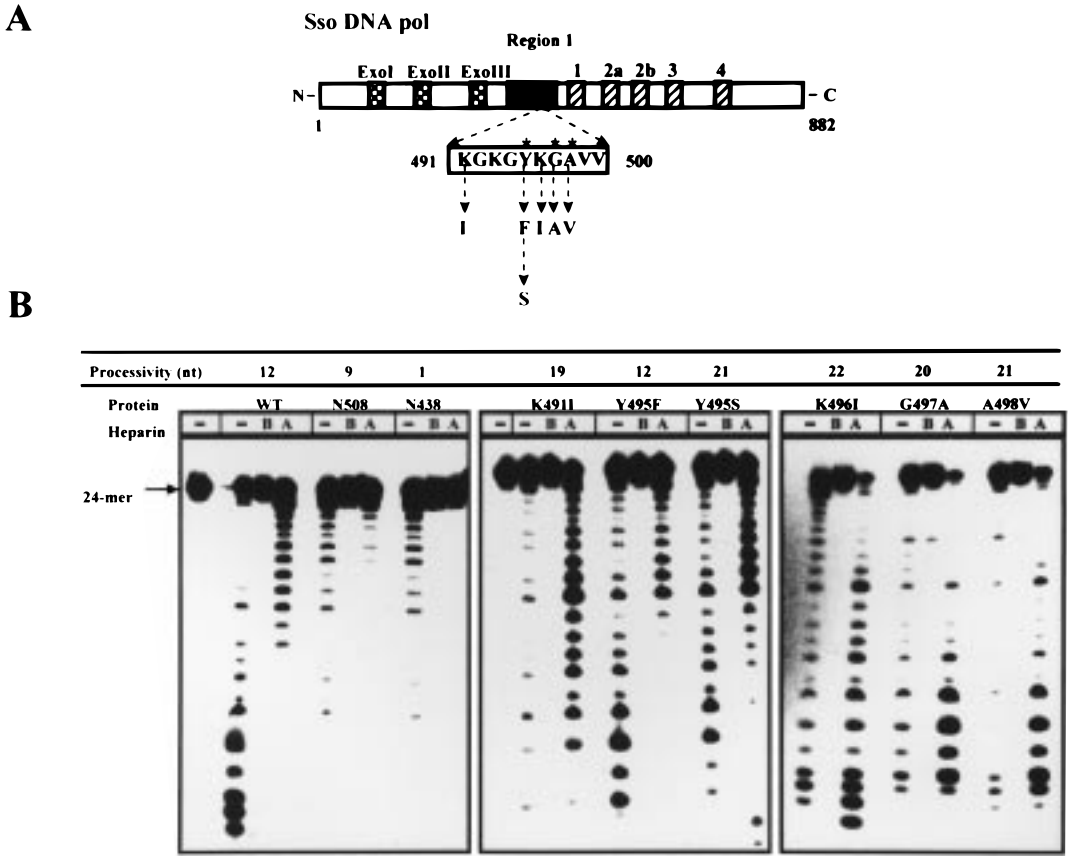


FIGURE 4: Site-specific mutations of Sso DNA pol at the highly conserved Y-GG/A consensus motif and their effect on the processivity of 3′−5′ exonuclease activity. (A) The amino acid substitutions produced for this study are shown. A diagramatic representation of the Sso DNA pol polypeptide chain is reported. Sequence similarity motifs are described by Blanco and Salas (*23*). The relative location of the Y-GG/A motif is indicated, and its amino acid sequence is reported within the box in the single-letter notation (invariant residues are indicated by asterisks). (B) Sso DNA pol (56 nM), N508 (320 nM), N438 (640 nM), and each site-specific mutant protein (56 nM) were incubated with the substrate (10 nM) in assay buffer without $MnCl_2$ for 2 min at 75 °C, and the reaction was begun by addition of $MnCl_2$ (final concentration of 2.5 mM). Where indicated, heparin (final concentration of 2 mg/mL) was added either before (lanes B) or after each protein at the same time as the $MnCl_2$ (lanes A), and the incubation was continued for an additional 0.5 min. Products were then run on a 20% polyacrylamide/8 M urea gel. The position of the 24-mer oligonucleotide is indicated by the arrow. The processivity of the 3′−5′ exonuclease activity expressed as the number of nucleotides excised per binding event is reported on the top for each protein.

from a single productive encounter with the enzyme. As a control experiment for testing the effectiveness of the trap, a mixture containing the radiolabeled substrate, heparin, and $Mn^{2+}$ ions was preincubated at 75 °C and then a sample of enzyme was added. As shown in Figure 4 (lanes B), preaddition of the trap prevents the proteins from binding to the oligonucleotide, and thus, no degradation takes place. On the other hand, single-turnover conditions are achieved by preincubation of the enzyme with the oligonucleotide and subsequent simultaneous addition of heparin and $Mn^{2+}$ ions (lanes A in Figure 4). Under these latter conditions, Sso DNA pol is able to remove about 12 nucleotides from the 24-base oligonucleotide 3′-terminus; protein N508 is able to excise up to eight or nine nucleotides, whereas protein N438 shows a very low level of processivity, being able to hydrolyze only one or two nucleotides. All the site-specific mutants display an increase in the 3′−5′ exonuclease processivity level, with the exception of protein Tyr495Phe, which is able to remove from the oligonucleotide 3′-end about 12 or 13 nucleotides as the wild type enzyme (see Figure 4B). It is interesting to note that mutation of the same Tyr residue to Ser produces a protein with a substantial increase in the processivity level (about 21 nucleotides excised per binding event).

To determine whether the variation in the processivity values among these proteins might be due to a change in either the affinity constant for the substrate or the nucleotide excision rate, the kinetic parameters ($K_m$ and $k_{cat}$) for the exonucleolytic degradation reaction were detrmined at 75 °C using a 5′-end-labeled 24-mer oligonucleotide. Reaction products were separated via denaturing polyacrylamide gels and quantitated by phosphorimagery to determine the initial rate of 3′-terminal nucleotide excision, as described in Experimental Procedures. The calculated values are reported in Table 1. Both the truncated derivatives N508 and N438 show reduced affinity and lower turnover number with respect to those of the full-sized Sso DNA pol. However, differences in $K_m$ values among these proteins seem to be more important than variation in $k_{cat}$ values; i.e., the Michaelis constant of N438 for the 24-base oligonucleotide is enhanced by about 3- and 9-fold with respect to those of N508 and Sso DNA pol, respectively. These results indicate that deletion of region 1 greatly affects the strength of the interaction with the oligonucleotide substrate, reducing the processivity level of the degradative function. In addition, a difference in $K_m$ values between protein N508 and the full-sized Sso DNA pol indicates that residues distal to region 1

Table 1: Apparent Kinetic Parameters for 3′−5′ Exonuclease Activity of Sso DNA Pol and Its Mutant Derivatives on a 24-Base Oligonucleotide[a]

| protein | $K_m$ ($\mu$M) | $k_{cat}$ ($s^{-1}$) | $k_{cat}/K_m$ ($\mu M^{-1}$ $s^{-1}$) | processivity[b] (nucleotides) |
|---|---|---|---|---|
| WT | 8 | 200 | 25 | 12 |
| K491I | 7 | 200 | 28 | 19 |
| Y495F | 9 | 173 | 19 | 12 |
| Y495S | ND[c] | ND | ND | 21 |
| K496I | 3 | 153 | 51 | 22 |
| G497A | 4 | 140 | 35 | 20 |
| A498V | 5 | 198 | 40 | 21 |
| N508 | 22 | 180 | 8 | 9 |
| N438 | 71 | 73 | 1 | 1 |

[a] Kinetic data, determined as described in Experimental Procedures, are mean values of assays carried out at least in triplicate with a 20% experimental error. [b] The processivity of the 3′−5′ exonuclease is expressed as the number of nucleotides excised by the enzyme per binding event to the substrate. [c] Not determined.

Table 2: Relative Polymerse and 3′−5′ Exonuclease Activities of Sso DNA Pol and Its Mutant Derivatives on Double-Stranded DNA[a]

| protein | exo activity (%) | pol activity (%) | pol/exo ratio |
|---|---|---|---|
| WT | 100 | 100 | 1 |
| K491I | 55 | 60 | 1.1 |
| Y495F | 68 | 158 | 2.3 |
| Y495S | 5 | 6 | 1.1 |
| K496I | 50 | 55 | 1.1 |
| G497A | 128 | 79 | 0.6 |
| A498V | 9 | 15 | 1.1 |
| N508 | 45 | NA[b] | ND[c] |
| N438 | 10 | NA | ND |

[a] Assays were carried out as described in Experimental Procedures. The percent activity is relative to the specific activity of wild type Sso DNA pol. Reported data are mean values of at least three determinations within a 20% experimental error. [b] Not active. [c] Not determined.

are also required for proper binding of a single-stranded substrate at the exonuclease active site.

As for the site-directed mutants, the measured $k_{cat}$ values are quite similar since the observed differences are within the experimental error found for this kind of enzymatic assay (see Table 1). On the other hand, we have found substantial variations in $K_m$ values, which seem to nicely correlate with the processivity level of each protein; i.e., mutant Lys496Ile with the highest processivity value (up to 22 nucleotides excised) displays the lowest $K_m$ value (about 2.6 $\mu$M). These results indicate that residues of the highly conserved Y-GG/A motif located within Sso DNA pol region 1 play a direct role in determining the single-stranded substrate binding efficiency and are involved in determining processivity of the 3′−5′ exonuclease activity.

*Mutations at the Y-GG/A Motif Affect Polymerase and 3′−5′ Exonuclease Activities on Double-Stranded DNA.* All the mutant derivatives of Sso DNA pol were tested for polymerase and exonuclease activities on double-stranded DNA, using as a substrate activated calf thymus DNA (see Experimental Procedures). Values are reported in Table 2. As expected, the C-terminally deleted proteins N508 and N438 do not retain any detectable DNA polymerase activity. Protein N438 shows a drastic reduction of the 3′−5′ exonuclease activity on double-stranded DNA, whereas protein N508 retains about 50% of the degradative capability on this kind of substrate in comparison to the full-sized Sso DNA pol. This finding suggests that region 1 might also be important for the interaction with double-stranded nucleic acid molecules.

As also shown in Table 2, mutations of $Tyr^{495}$ to Ser and $Ala^{498}$ to Val are found to reduce drastically both polymerase and 3′−5′ exonuclease activities on double-stranded DNA. Less dramatic, and with almost no effect on the polymerase/exonuclease ratio, are the mutations of $Lys^{491}$ and $Lys^{496}$ to Ile, since the corresponding mutant proteins retain almost 50% of both activities with respect to the wild type enzyme. On the other hand, mutant Tyr495Phe displays a substantially higher polymerase activity and reduced degradative function in comparison to the wild type Sso DNA pol, whereas subsitution of $Gly^{497}$ with Ala seems to favor the exonucleolytic activity on double-stranded DNA in comparison to nucleotide incorporation capability. It is interesting to notice that all the amino acid substitutions we have introduced produce effects on both synthetic and degradative activities and some of them (such as Tyr495Phe and Gly497Ala) substantially alter the polymerase/exonuclease ratio. These data strongly support the proposal that region 1 of Sso DNA pol has a direct role in the interaction with the nucleic acid substrate, and that it could be involved in molecular "switching" of the primer strand terminus between polymerase and exonuclease active sites.

*Structural and Functional Similarities Shared with Other DNA Polymerases.* On the basis of a mutational analysis carried out on the replicase from *Bacillus subtilis* bacteriophage $\phi$29, a critical role in the coordination or cross-talk between synthesis and degradation was proposed for the Y-GG/A consensus motif (*19*). In addition, Stocki et al. (*20*) identified several T4 DNA polymerase mutants affected in active site switching. A certain number of the mutated residues were mapped within a portion of the polypeptide chain which corresponds to Sso DNA pol region 1, as indicated by multiple sequence alignments among family B DNA polymerases (*18*).

Furthermore, Lin et al. (*21*) reported on the characterization of a T4 DNA polymerase truncated form with 3′−5′ exonuclease activity, referred to as N388. This protein displays a specific exonucleolytic activity with long oligonucleotides and a degree a processivity on oligo$(dT)_{32}$ remarkably lower in comparison to that of the full-sized enzyme. It is interesting to note that protein N388 (like the derivative N438 of Sso DNA pol) does not include at its C terminus the previously discussed Y-GG/A sequence motif.

More recently, the crystal structure of the replicative DNA polymerase from *E. coli* bacteriophage RB69 has been elucidated (*8*). The RB69 replicase is 63% identical to its homologue from T4 phage and is similar in the primary structure to several family B DNA polymerases, including Sso DNA pol. Analysis of this three-dimensional structure has revealed that the highly conserved Y-GG/A motif is located in a long solvent-accessible loop (residues 380−394) between the exonuclease and polymerase domains. Our results, which suggest that region 1 of Sso DNA pol is flexible and exposed to the solvent on the protein molecule surface (*11*), are consistent with this finding. The structure at low resolution of a complex between single-stranded DNA and the RB69 replicase has revealed that the four 3′-terminal nucleotides of oligo $p(dT)_{16}$ are bound within a cleft (editing channel) formed by the exonuclease domain and the tip of the polymerase thumb subdomain. This finding indicates

that residues of the polymerase domain are also required for efficient binding of the nucleic acid substrate at the exonuclease active site. This could explain the increased affinity constant values and the reduced catalytic efficiency of the two C-terminally deleted forms of Sso DNA pol (N438 and N508) described in this study. Although the structure of RB69 DNA pol bound to double-stranded DNA has not yet been published, it has been suggested that duplex DNA, bound to a second cleft (cleft D), occupies the same location whether the primer terminus is situated in either the polymerase or the exonuclease active site of the enzyme and that residue $Tyr^{391}$ of the Y-GG/A consensus motif could make contact with the primer strand 3′-terminus (*8*). Therefore, in RB69 DNA polymerase, the loop segment that is formed by residues 380−394 and includes the Y-GG/A sequence motif could play a critical role in switching the nucleic acid substrate between the polymerase and exonuclease catalytic sites. The above structural data are consistent with the biochemical analysis reported herein if we consider that active site translocation and processivity of the proofreading activity might be correlated functions that are likely to rely on a common mechanism for "sliding" of the enzyme molecule along the nucleic acid substrate (*22*). However, for a greater definition of the amino acid residues involved in nucleic acid interaction, the cocrystal structure of a family B DNA polymerase bound to duplex DNA substrates is needed.

## ACKNOWLEDGMENT

We acknowledge Ottavio Piedimonte and Ferdinando Febbraio for the artwork and Vitale Carratore for the N-terminal sequence analyses.

## REFERENCES

1. Hubscher, U., Maga, G., and Podust, V. N. (1996) in *DNA replication in eukaryotic cells* (DePamphilis, M. L., Ed.) Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
2. Steitz, T. A., Smerdon, S. J., Jager, J., and Joyce, C. M. (1994) *Science 266*, 2022−2025.
3. Pelletier, H. (1994) *Science 266*, 2025−2026.
4. Beese, L. S., Derbyshire, V., and Steitz, T. A. (1993) *Science 260*, 352−355.
5. Derbyshire, V., Pinsonneault, J. K., and Joyce, C. M. (1995) *Methods Enzymol. 262*, 363−385.
6. Derbyshire, V., Grindley, N. D. F., and Joyce, C. M. (1991) *EMBO J. 10*, 17−24.
7. Beese, L. S., and Steitz, T. A. (1991) *EMBO J. 10*, 25−33.
8. Wang, J., Yu, P., Sattar, A. K. M. A., Wang, C. C., Karam, J. D., and Steitz, T. A. (1997) *Cell 89*, 1087−1099.
9. Pisani, F. M., De Martino, C., and Rossi, M. (1992) *Nucleic Acids Res. 20*, 2711−2716.
10. Pisani, F. M., and Rossi, M. (1994) *J. Biol. Chem. 269*, 7887−7892.
11. Pisani, F. M., Manco, G., Carratore, V., and Rossi, M. (1996) *Biochemistry 35*, 9158−9166.
12. Studier, F. W., and Moffatt, B. A. (1986) *J. Mol. Biol. 189*, 113−130.
13. Ho, S. N., Hunt, D. H., Horton, R. M., Pullen, J. K., and Pease, L. R. (1989) *Gene 77*, 51−59.
14. Longley, M. J., and Mosbaugh, D. W. (1991) *Biochemistry 30*, 2655−2664.
15. Cheng, C.-H., and Kuchta, R. D. (1993) *Biochemistry 32*, 8568−8574.
16. Leatherbarrow, R. J. (1992) *GraFit*, version 3.0, Erithacus Software Ltd., Staines, U.K.
17. Reddy, M. K., Weitzel, S. E., and von Hippel, P. H. (1992) *J. Biol. Chem. 267*, 14157−14166.
18. Braithwaite, D. K., and Ito, J. (1993) *Nucleic Acids Res. 21*, 787−802.
19. Truninger, V., Lazaro, J. M., Salas, M., and Blanco, L. (1995) *EMBO J. 15*, 3430−3441.
20. Stocki, S. A., Nonay, R. L., and Reha-Krantz, L. J. (1995) *J. Mol. Biol. 254*, 15−28.
21. Lin, T.-C., Karam, G., and Konigsberg, W. H. (1994) *J. Biol. Chem. 269*, 19286−19294.
22. Johnson, K. A. (1993) *Annu. Rev. Biochem. 62*, 685−713.
23. Blanco, L., and Salas, M. (1996) *J. Biol. Chem. 271*, 8509−8512.

BI981127S