Biochemical analysis of point mutations in the 5’-3’ exonuclease of DNA polymerase I of *Streptococcus pneumoniae*: functional and structural implications

Mónica Amblar, Mario García de Lacoba, Maria A. Corrales and Paloma López

Centro de Investigaciones Biológicas, Consejo Superior de Investigaciones Científicas, Velázquez 144, 28006 Madrid, Spain.

Running title: active site of the DNA polymerase I 5’-3’ exonuclease

Send correspondence to: Mónica Amblar
Centro de Investigaciones Biológicas
Velázquez 144
28006 Madrid, Spain
Phone: 34-91-5611800
Fax: 34-91-5627518
E-mail: mamblar@cib.csic.es
SUMMARY

To define the active site of the 5'-3' exonucleolytic domain of the *Streptococcus pneumoniae* DNA polymerase I (Spn PolI), we have constructed His-tagged-Spn PolI fusion protein and introduced mutations at residues Asp\(^{10}\), Glu\(^{88}\) and Glu\(^{114}\), that are conserved among all prokaryotic and eukaryotic 5' nucleases. The mutations, but not the fusion to the C-terminal end of the wild-type, reduced the exonuclease activity. The residual exonuclease activity of the mutant proteins has been kinetically studied, together with potential alterations in metal binding at the active site. Comparison of the catalytic rate and dissociation constant of the D10G, E114G and E88K mutants and the control fusion protein support: (i) a critical function of Asp\(^{10}\) in the catalytic event, (ii) a role of Glu\(^{114}\) in the exonucleolytic reaction, being secondarily involved in both catalysis and DNA binding, and (iii) a nonessential function of Glu\(^{88}\) for the exonuclease activity of Spn PolI. Moreover, the pattern of metal activation of the mutant proteins indicates that none of the three residues is a metal-ligand at the active site.

These findings and those previously obtained with D190A mutant of Spn PolII are discussed in relation to structural and mutational data for related 5’ nucleases.

* This research was under the auspices of the Consejo Superior de Investigaciones Científicas, and supported by the European Union grant QLK2-CT-2000-00543 and the Programme of Strategic Groups of the Comunidad de Madrid.

\(^1\) The abbreviations used are: Spn PolI, *S. pneumoniae* DNA polymerase; Spn PolID10A, Spn PolI with the amino acid substitution D10A; Spn PolID190A, Spn PolI with the amino acid substitution D190A; FEN, flap endonuclease: Taq Pol, *T. aquaticus* DNA polymerase I; Pfu Pol, *P. furiosus* DNA polymerase
II; IPTG, thio-β-D-galactoside; Spn PolII-(his), Spn PolII fused to an His-Tag at the C-terminal end; Eco PolII, *E. coli* DNA polymerase I; Mtb PolII, *M. tuberculosis* DNA polymerase I; MjFEN-1, *M. jannaschii* FEN-1; hFEN-1, human FEN-1
INTRODUCTION

The polymerase function of type-I-like DNA polymerases has been studied in considerable detail, with biochemical and mutagenesis analysis proceeding in parallel with the structural studies (1-6). Such detailed study have provided a prototypical molecular model of DNA-dependent DNA polymerization and important insights into the architecture of the primer and nucleotide binding sites. By contrast, in spite of the mutational and structural analysis of several 5’-3’ exonucleolytic domains of eubacterial polymerases (7-9) and related bacteriophage 5’ nucleases (10-12), the molecular mechanism of the exonucleolytic reaction still remains obscure.

Sequence comparisons and enzymatic studies indicate that the eubacterial PolI-associated 5’ nucleases share significant sequence homology with the polymerase-independent 5’ nucleases from several bacteriophages (13). The prokaryotic 5’ nucleases are also related to mammalian FEN-1 proteins and several yeast proteins of the RAD2 family, having two large blocks of sequence similarity that bear some resemblance to the bacterial and bacteriophage nuclease sequences (14, 15). Some clues to identify important residues in the bacterial 5’-3’ exonuclease family derive from the multiple sequence alignment of ten bacterial and bacteriophage nucleases (13). Six conserved sequence motifs containing 14 invariant amino acids were identified, nine of which were carboxylate residues. The presence of highly conserved carboxylate residues, led to the proposal that some of these amino acids could be involved in metal binding at the active site of the 5’-3’ exonucleases, as occurs in other enzymes catalyzing phosphoryl transfer reactions (1). This hypothesis has been further supported by structural data from the 5’ nucleases from bacteriophage T5 (10) and Taq Pol (9), and from T4 RNase H (11). In these three proteins, the 5’ nuclease active site consists of a set of carboxylate
residues which coordinate metal ligands that are essential for the nuclease activity. However, several intriguing differences among the three active sites exist, that pose some important questions.

DNA polymerase I of *Streptococcus pneumoniae* (Spn PolI) is a bifunctional protein having two enzymatic activities: DNA polymerase and 5’-3’ exonuclease (16). These activities are located on different domains of the protein that are arranged in the same order in all PolI-like DNA polymerases (17, 18). Like other DNA polymerases of the family, both enzymatic activities of Spn PolI are involved in DNA-repair processes (19, 20). Unlike that of *Escherichia coli* (21), the exonucleolytic domain has proved to be essential for pneumococcal cell viability (20). Previous studies on the exonuclease activity of Spn PolI showed an essential role for Asp\(^{10}\) and Asp\(^{190}\) in the exonucleolytic reaction, since the substitution of these carboxylate residues by Ala led to an almost total inactivation of the nuclease domain in Spn PolID10A protein (22) and to a drastic reduction of the catalytic efficiency as well as to an altered metal binding in the Spn PolID190A mutant protein (23).

In this paper we describe the overproduction and purification of a His-tagged fusion form of Spn PolI and the introduction of mutations at three residues of the exonucleolytic domain of the protein, Asp\(^{10}\) and Glu\(^{114}\) (proposed metal ligands in other nucleases) and Glu\(^{88}\) (highly conserved among prokaryotic 5’nucleases). We also report the use of these enzymes in a kinetic study to explore the roles of the conserved carboxylate residues in the exonucleolytic reaction. Finally, we present a three-dimensional model of the putative 5’-3’ exonucleolytic domain of Spn PolI, built by homology modeling, in order to provide us with a framework to support some structural explanations of the Spn PolI mutant activities.
EXPERIMENTAL PROCEDURES

Materials

Restriction enzymes, T4 DNA ligase and T4 polynucleotide kinase were purchased from New England Biolabs. T7 DNA polymerase and inorganic pyrophosphatase were obtained from Pharmacia, Taq Pol from Boehringer Mannheim and Pfu Pol from Stratagene. Unlabeled oligonucleotide primers were synthesized in a Gene Assembler (Pharmacia) at the Centro de Investigaciones Biológicas, Madrid, Spain.

The *E. coli* strains used were JM109 (*endA1 recA1 gyrA96 thi hsdR17 (rK− mK+)* relA supE44 *Δ(lac-proAB) F* (traD36 proA*B* lacI*ΔZAM15) (24) for cloning experiments and BL21(DE3) (*F* rB− mB− gal ompT (int::PlacUV5 T7 gen1 imm21 nin5) (25) for expression of enzymes to be purified.

Construction of plasmids expressing Spn PolI derivatives

The pMA6 plasmid was constructed by cloning a 5.3 kb *Eco*RI fragment from plasmid pSM29 (26), containing the pneumococcal *polA* gene, into the *Eco*RI site of the phagemid pAlter-1 vector (Promega). By site directed mutagenesis two new restriction sites were introduced at both sides of *polA* gene on pMA6 plasmid: *Eco*RI (between the -10 and -35 boxes) and *Xho*I (at the stop codon of the gene). Then, the mutated *polA* gene was excised from the pMA6 derivative by digestion at the newly created *Eco*RI and *Xho*I restriction sites and ligated to the pET21b (fusion vector derived the *E. coli* pET5 expression vector (27)). The resulting plasmid, named pMA9, contains the *polA* gene fused at its 3’-end to an His-Tag coding sequence under the control of T7 gene φ10 promoter (27). As a consequence, the stop codon of the structural gene was replaced by
a DNA fragment encoding the amino acid sequence LE(H)₆ obtaining the fusion
derivative named Spn PolI-(his), which can be selectively bounded to chelating
sepharose resin through the histidine residues. Plasmid pMA10 was derived from pMA9
by removal of the 80 bp XbaI-EcoRI fragment containing the ribosome binding site of
φ10 gene and the T7-Tag coding region. Expression plasmids for PolI derivatives having
mutations at the 5’-3’ exonucleolytic domain of the protein, were obtained by swapping
the 399 bp EcoRI-NheI restriction fragment from pMA10 for the corresponding PCR
mutagenic products. The sequence of the mutated polA genes was determined by
thermal cycle sequencing of the resulting recombinant plasmids using Taq Pol, the
Abitrin 377 automatic sequencer and the corresponding kit supplied by Applied
Biosystems Inc., at the Centro de Investigaciones Biológicas.

Mutagenesis of the polA gene

Introduction of the restriction sites EcoRI and XhoI at the pneumococcal polA gene was
carried out using ECR (5’-CAATGGTATTTTTTGAATTCTTTCCTTTATA-3’) and
XHO (5’-CTGGTACGAGGCTAAACTCGAGGGGGCTAGTCCTC-3’)
oligonucleotides, and the Altered site in vitro mutagenesis system from Promega.
Mutations of the 5’-3’ exonuclease domain coding region of polA gene were obtained
by using the high-yield method for site directed mutagenesis by PCR and three primers
as described by Steingberg et al. (28). The mutagenic primer used was 5’-
AAATTATTGATTNNNGGGTCTTCTGTAGCT-3’, where N corresponds to
degenerated position within the 10th codon, at which the four nucleotides can be
introduced. ECR and A-PCR (5’-AGGGATATTATCCGACTTACCCAACCCCA-3’)
oligonucleotides were used as primers for the amplification reaction. PCR products were
digested with _EcoRI_ and _NheI_ and subcloned into the expression plasmid pMA10. The resulting reaction mixture was introduced in _E. coli_ JM109 strain by electroporation and a set of the clones obtained were sequenced to determine the mutation efficiency. The PCR-mutagenic procedure was first carried out using Taq Pol during the amplification reaction. The analysis of 36 clones showed that 21 of them contained mutations at non-targeted positions of the _polA_ gene, such as insertions or deletions (resulting in frame-shift mutations), or multiple base substitutions. Only 4 of the 36 clones contained different mutations (GCG, CAG, CGU and UAG) at the 10\textsuperscript{th} codon (GAU), giving the amino acid substitutions D10A, D10Q, D10R and D10stop, respectively. One of the non-targeted mutations obtained during PCR reactions caused a change at codon 88 (GAG by AAG) resulting in the E88K amino acid substitution in Spn PolI. Then, a second round of PCR mutagenesis was carried out using Pfu Pol (enzyme with high fidelity of polymerization). Of the 37 clones analyzed, 17 were found to carry plasmids with mutations at the 10\textsuperscript{th} codon. Moreover, this analysis revealed 5 new mutations in which codon number 10 was replaced by AGU, GUG, ACU, AAA or GGC. Such changes correspond, respectively, to the amino acid substitutions D10S, D10V, D10T, D10K and D10G. Only one clone contained, in addition to D10A mutation, the amino acid substitution E114G (change of GAG by GGG at codon 114). Since E114 residue is highly conserved in the family of prokaryotic 5\textsuperscript{'-3\textsuperscript{'}} exonucleases (13), the mutation at this position was subcloned as a single amino acid substitution of Spn PolI into the expression plasmid pMA10, making use of _NdeI_ and _XhoI_ sites of _polA_ gene that are bracketing the 114\textsuperscript{th} codon.

**Induction protocol of Spn PolI and its derivatives**
Plasmid pMA10 and its derivatives were transferred to the *E. coli* BL21(DE3) overproducer strain. The expression of the wild-type PolI enzyme and its fusion derivatives was achieved by IPTG-induction of the host strain BL21(DE3) (27) containing the corresponding expression plasmid (pSM23 for wild-type enzyme (26) and pMA10 set plasmids for fusion proteins).

Cells containing the corresponding plasmid were grown in M9 medium supplemented with 200 µg/ml of ampicillin at 37 ºC to an absorbance of 0.45 at 600 nm and then, induced by addition of 0.75 mM IPTG. Samples were withdrawn at different induction times and crude extracts prepared to test enzymatic activities. The crude extracts were also examined by 0.1% SDS/8% PAGE. Quantification of the proportion of Spn PolI and mutant forms in crude extracts was performed by scanning the gels with the Molecular Analyst system (Bio-Rad Laboratories).

**Preparation of small-scale crude extracts**

Cell pastes obtained from 1.5 ml IPTG-induced cultures of the appropriate strain were washed by suspension in 1 ml of buffer I (10 mM Tris-HCl, pH 7.6, 3 mM β-mercaptoethanol), centrifuged and suspended in 58 µl of lysis buffer II (323 µg/ml lysozyme in buffer I). The resulting suspensions were incubated 5 min at 37 ºC, supplemented with 92 µl of buffer I and further incubated 5 min at 37 ºC. Then, samples were treated with 0.1% Triton X-100 for 5 min at 37 ºC and subjected to three cycles of freezing and thawing at −70 ºC and 37 ºC, respectively. The viscosity of the extracts was reduced by passage through a 0.36-mm-inner-diameter needle, then were centrifuged and the supernatant stored at -70 ºC.
**Purification of His-Tag fusion proteins**

Cells from 100 ml IPTG-induced cultures were harvested by centrifugation, washed with 50 ml of buffer A (0.5 M NaCl, 20 mM HPO$_4$Na$_2$, pH 7.6) and suspended in 10 ml of the same buffer. Cell lysis was achieved by incubation with 0.3 mg/ml lysozyme for 30 min at 0 °C and three cycles of freezing and thawing. The crude extracts were ultracentrifuged at 81000 x g, 50 min, and the soluble fraction recovered. After addition of 10 mM imidazole, the clarified extracts were added to 4 ml of Chelating Sepharose Fast Flow (Pharmacia Biotech), previously charged and equilibrated in buffer A plus 10 mM imidazole. The mixture was incubated by end-over-end rotation with gentle agitation for 30 min at 4 °C. The resin was sedimented by centrifugation at 2000 x g, 3 min, and washed three times by addition of five volumes of 10 mM imidazole-buffer A and incubation for 5 min at 4 °C as described above. After washing, the resin was again incubated with two volumes of 40 mM imidazole-buffer A, and the protein was eluted by incubation with an equal volume of 100 mM imidazole-buffer A. The imidazole was removed from the sample by dialysis against buffer A plus 50% of glycerol, allowing at the same time the concentration and equilibration in the storage buffer. Quantification of the proteins was carried out by gel fractionation of the sample and scanning of the gels with Molecular Analyst system (Bio-Rad Laboratories).

**Polymerase assay**

Polymerase activity was determined on activated calf thymus DNA following the method described previously (18). One unit of polymerase activity is defined as the
amount of enzyme catalyzing the incorporation of 10 nmol of dNTP into DNA in 30 min at 37 °C.

5'-3' exonuclease assays

The nuclease activity in DNA-containing 0.1% SDS/10% polyacrylamide gels after electrophoresis and removal of SDS was assayed as previously described by Rosenthal and Lacks (29). Exonuclease activity, assayed using salmon sperm DNA, was determined in the presence of 0.1 mM MnCl₂ after 30 min at 37 °C, as previously described (16). The salmon sperm substrate, (previously nicked with pancreatic deoxyribonuclease I) was labeled with [³H]dTTP using SpnPolc269 (an Spn PolI derivative that only contains the polymerase domain (18)). Exonuclease activity was also tested using a 5'-³²P-labeled (5'-CCAGTCACGACGTTGT-3') or 3'-³²P-labeled (5'-CCAGTCACGACGTTGTA-3') oligonucleotide annealed to M13mp2 ssDNA as described (23). The concentration of DNA substrate, MnCl₂ or MgCl₂ and enzymes as well as reaction time are indicated in Results. In the case of reactions with 5’-end labeled substrate, the concentration of mutant proteins used was as high as possible without exceeding the inhibitory concentration of 8% of glycerol in the reaction (enzymes were stored in 50% glycerol). One unit of exonuclease activity is defined as the amount of enzyme required for the release of 10 nmol of nucleotide from DNA in 30 min at 37 °C.

To measure the 5’-3’ exonuclease rates, experiments were performed using the 3’-³²P-labeled 17-mer oligonucleotide annealed to M13mp2 ssDNA as substrate, as previously described (23). The assays were carried out with the preferred divalent metal ion at the optimal concentration for the nuclease activity (0.1 mM of MnCl₂), and increasing
substrate concentrations. The amount of each enzyme was adjusted to obtain linear conditions and samples (2 µl) were removed at appropriate times during incubation at 37 ºC. The products were fractionated in a 20% polyacrylamide gel, where the degradation products are distinguished as discrete bands. A typical experiment is depicted in Fig. 1. The experimental objective was to determine the velocity of exonucleolysis (v_{n,n-1}) by measuring the proportion of reaction products of n and n-1 nucleotides in the gel. Since every primer that reached position n-1 also reached position n, the velocity at the site can be expressed by the following Equation (A1).

\[
v_{n,n-1} = \left( \sum_{i=m}^{n} \frac{I_i}{t} \right) \left( \sum_{i=m}^{n-1} \frac{I_i}{I_n} \right)
\]

(A1)

Where \( t \) is the reaction time, \( I_i \) is the integrated intensity at site \( i \) expressed as a percentage of total substrate and \( m \) is the length of the final product of the reaction. The velocity of the conversion reaction \( v_{15,14} \) and \( v_{14,13} \) was measured in order to minimize the contribution of both initial DNA:protein complex formation and the high dissociation rate at 11-mer product observed for all proteins analyzed (Fig. 1 and results not shown). The velocities remained essentially constant for \( t \) up to 30 min for all fusion proteins, indicating that the DNA:enzyme complexes at the point of analysis were in steady state (data not shown). In addition, the concentration of the 15-mer and 14-mer substrates remained lower than 5% of the total DNA, indicating that only one turnover was measured. The relationship of velocity and concentration of primed M13mp2 DNA substrate conformed to the Michaelis-Menten equation, as indicated by linearity in the Lineweaver-Burk plots (30) of [1/v] versus [1/ primed M13mp2] (Fig. S1). The double-reciprocal plots depicted in Figure S1 were fitted by a linear least-squares regression.
analysis and used to determine, from the intercepts, $V_{\text{max}}$ (corresponding to the maximum value of $I_{n,1}/I_n$), and $K_m$ (corresponding to the value of [primed M13mp2] when $I_{n,1}/I_n$ is at half maximum). The catalytic rate of the exonuclease reaction ($k_{\text{cat}}$) was calculated as $V_{\text{max}}/[E_{\text{total}}]$, ([E_{total}] being the concentration of enzyme used in the assay).

**Filter-binding assays**

Formation of DNA:Spn PolII derivative complexes was measured by using alkali-treated nitrocellulose filters (Millipore, type HAWP 45 µm) as described by McEntee et al. (31). The DNA substrate was obtained by PCR amplification of the BglII-XbaI 2239 bp polA gene fragment using ECR and A-PCR oligonucleotides and ($\alpha$-32P)-dCTP and Taq PolII enzyme, which generate dsDNA with 3´ protruding ends. The resulting 616 bp dsDNA was treated with NheI restriction enzyme and the 414 bp product, which contains only one 5´ protruding end, was purified by gel electrophoresis. As the presence of metal ion in the binding buffer was required to detect specific DNA retention on the filters (data not shown) it was necessary to standardize, conditions of the binding assay to minimize 5´-3´ exonuclease activity. The standard binding reaction was carried out in 30 µl of buffer B (10 mM Tris-ClH, pH 7.6, 1 mM DTT, 50 mM KCl, 2.6% glycerol, 0.1 mM MnCl$_2$) with 0.6 nM of the ($\alpha$-32P)-labeled 414 bp PCR product and increasing amounts of protein. The mixtures were incubated for 10 min at 15 ºC in order to reach DNA:protein equilibrium and the reactions stopped by addition of 150 µl ice-cold buffer B. Samples were filtered, washed with 9 ml of the same buffer, and dried, and their radioactivity measured by scintillation counting. The quantification of the amount of Spn PolII-derivatives:DNA complexes retained on the filters was
corrected by subtracting the non-specific retention of labeled DNA in the absence of PolI derivatives and in the presence of an equal BSA concentration.

**Molecular modeling**

A three-dimensional model of the putative 5’-3’ exonucleolytic domain of Spn PolI was built from its amino acid sequence, the 2.4 Å-resolution X-ray structure of the 5’-3’ exonucleolytic domain of Taq Pol (9), and 2.5 Å-resolution X-ray structure of the T5 5’ nuclease (10) by using knowledge-based protein modeling methods. Their cartesian coordinates were from the Brookhaven Protein Data Bank, with PDB codes 1TAQ and 1EXN, respectively. The structural conserved regions and the definition of the corresponding connecting loops were identified from the multiple alignment of the amino acid sequences of Spn PolI, Taq Pol, and T5 5’ nuclease (Fig. 6). This alignment was initially identified from the multiple alignment of the nuclease domains of 38 prokaryotic and eukaryotic proteins (Fig. S2). The overall conformation of the 5’-3’ exonucleolytic domain of Spn PolI was subjected to energy minimization until convergence, using a combination of steepest descent and conjugate gradients algorithms, was achieved. The energy calculations were carried out under the AMBER force field (33). Computations were performed on a Power Challenge R10000 by using the BIOSYM software package, release 95.0 (Molecular Simulations, Inc., San Diego, CA).
RESULTS

Production, expression, and purification of Spn PolI-(his) and its derivatives

*Spn PolI-(his).* After 3.5h IPTG-induction of *E. coli* BL21(DE3) containing pMA10, Spn PolI-(his) was the major protein product in cell extract, corresponding to about 9% of the total protein content (Fig. 2A, lane 4). This yield was very similar to the 10% obtained with cells harboring pSM23 (26), which encodes Spn PolI (Fig. 2A, lane 3). The nuclease activity in cell extracts from both BL21(DE3)[pMA10] and BL21(DE3)[pSM23] cultures, was detected *in situ* using a DNA containing polyacrylamide gel (Fig. 2B). In both extracts, a degradation band was detected corresponding to a polypeptide of approximately 100 kDa, the predicted size for Spn PolI and its fusion derivative (Fig. 2B, lanes 2 and 3). In addition, other bands of activity were observed that presumably corresponded to proteolytic fragments of the pneumococcal enzymes, since they were not detected in extracts carrying the pET5 vector (Fig. 2B, lane 1). Quantification of the amount of Spn PolI and Spn PolI-(his) (Fig. 2A, lanes 3 and 4) *versus* their nuclease activity (Fig. 2B, lanes 2 and 3) revealed that the fusion of the His-Tag did not affect the 5’-3’ exonuclease specific activity. However, this fusion resulted in a five-fold reduction of the polymerase activity of the pneumococcal enzyme (producing 60U of polymerase activity per mg of protein after 3.5h induction, compared with 318U for the vector coding for the wild-type Spn PolI, after correction for the endogenous polymerase activity (3U) of the host organism). Thus, the Spn PolI-(his) construct is appropriate for mutational analysis of the 5’-3’ exonuclease activity of the pneumococcal enzyme.

Spn PolI-(his) was purified from IPTG-induced cultures by binding to a chelating sepharose resin (Fig. 2C). By testing different concentrations of imidazole-eluting agent,
we developed a purification procedure yielding 0.37 mg of Spn PolI-(his) per L of induced culture. The purity of the protein sample was greater than 60% (Fig 2C, lane 8), with polymerase and exonuclease specific activities of 614 and 1049 U/mg of protein, respectively.

5′-3′ exonuclease mutant forms of Spn PolI-(his). The Asp10 of Spn PolI is highly conserved among the 5′ nucleases (7), and structural data support the involvement of this residue in the active site (9-11). Substitution of Asp10 by Ala in Spn PolI (22) or its equivalent residues in Eco PolI (7) or Mtb PolI (8) by Asn, yielded proteins with so little 5′ nuclease activity that they could not be further analyzed. To obtain a mutant protein that could be kinetically characterized, we mutated the pneumococcal polA gene at the 10th codon. Ten different mutant forms of Spn PolI-(his) were obtained, eight of them with different amino acid changes at the Asp10 (D10G, D10V, D10T, D10K, D10A, D10Q, D10S and D10R) and two of them containing either E88K or E114G amino acid substitutions. These carboxylate residues, Glu88 and Glu114, are also conserved among the prokaryotic 5′ nucleases (7). All mutant forms were overproduced in E. coli BL21(DE3) and purified as for Spn PolI-(his). Analysis of the protein production and polymerase activity after IPTG-induction at 30 or 37 ºC of two representative mutant derivatives (Spn PolID10A-(his) and Spn PolID10G-(his)), revealed that incubation for 3 h at 37 ºC was optimal induction, without significant insolubilization of the proteins (data not shown). This was surprising, since we have previously observed that 91% of Spn PolID10A was insoluble upon induction at 37 ºC (22). This procedure allowed us to purify the mutant proteins with a typical yield of 0.4 mg of protein per L of induced
culture at greater than 70% purity and 200-600 U of polymerase activity per mg of protein.

**Enzymatic activities of Spn PolII-(his) and its derivatives**

All the purified proteins were assayed for polymerase activity on activated DNA (Table 1). Similar specific activities were obtained for all mutant and control fusion proteins, indicating that the polymerase domain of Spn PolII was functionally unaffected by the single amino acid changes in the 5′-3′ exonucleolytic domain. Next, we determined the exonuclease activity of the enzymes on salmon sperm DNA substrate. Only E88K and the control protein showed a detectable activity, 890 and 1000 U/mg of protein, respectively. Due to the low exonuclease activity of the mutant proteins, it was necessary to use a more sensitive assay to determine the effect of the mutations introduced at the exonucleolytic domain. Therefore, a 5′-32P-labeled oligonucleotide annealed to M13mp2 ssDNA was used as substrate to compare the 5′-3′ exonuclease activity of Spn PolII-(his) and all the mutant proteins (Fig. 3). With this substrate we observed the first nucleolytic event in the exonucleolytic reaction. Spn PolII-(his) generated the same products as those previously obtained with Spn PolII (23), derived from both the 5′-3′ exonuclease (mononucleotides) and the 5′-end-dependent endonuclease activities (dinucleotides), which are present in all 5′ nucleases of the family (34). Of the ten mutant proteins, only those carrying the mutations E114G, E88K, D10G and D10A gave detectable activities, with that of the Spn PolID10A-(his) mutant only being visible after 16 h of incubation, as previously observed with the Spn PolID10A mutant (22). In addition, none of these four mutations altered the ratio of mononucleotide to dinucleotide products (approximately 15:1), in contrast to the
previously described behavior of Spn PolID190A, which showed a prevalence of endonucleolytic over exonucleolytic cleavages (23). In some reactions, products most likely generated by removal of mononucleotides at the 3’-end of the oligonucleotide substrate were also observed. These products were presumably produced by residual contamination of the samples with 3’-5’ exonucleases. Quantification of the experiments, shown in Fig. 3, allowed us to categorize the mutant proteins in terms of their exonuclease activity (Table 1). The E88K mutant possessed 26% of the exonuclease specific activity present in Spn PolI-(his), indicating that Glu_{88} is not an essential residue for this enzymatic activity. By contrast, the mutations introduced at Glu_{114} or Asp_{10} resulted in a decrease of more than 98% of the exonuclease activity compared to that of the control enzyme. In the case of the amino acid substitutions of Asp_{10}, 5’-3’ exonuclease activity was only measurable in D10G and D10A mutants. These results argue in favor of an essential role of the Asp_{10} and Glu_{114} residues in the exonuclease activity. Therefore, we proceeded to determine the role played by Asp_{10}, Glu_{88} and Glu_{114} residues by analyzing the mutant forms Spn PolID10G-(his) (the only mutant at Asp_{10} that retains significant exonuclease activity), Spn PolIE88K-(his) and Spn PolIE114G-(his).

**Metal-dependency of the 5’-3’ exonuclease activity of His-Tag-PolI derivatives**

The divalent metal ion requirements for the nuclease activity of the His-Tag Spn PolI fusion proteins, were determined using a 3’-^{32}P-end-labeled M13-primed substrate and measuring the 5’-3’ exonuclease activity at different Mn^{2+} or Mg^{2+} concentrations. The reaction products were fractionated in a denaturing 20% polyacrylamide gel and the degradation bands quantified as previously described (23).
The results obtained revealed that the Mn\(^{2+}\) and Mg\(^{2+}\) dependence of the exonuclease activity was similar for the four fusion proteins analyzed (Fig. 4), and correlated with that previously obtained for Spn PolII (23). Thus, neither the fusion of the His-Tag at the C-terminal end of the protein nor the amino acid changes introduced altered the metal dependency of the 5’-3’ exonuclease activity of the \textit{S. pneumoniae} enzyme. The His-Tag proteins showed maximum exonuclease activity over a wide range of Mn\(^{2+}\) concentration, ranging between 10 \(\mu\)M and 1mM (or even 5 mM) of MnCl\(_2\), but diminished dramatically at higher concentrations. In the case of Mg\(^{2+}\), the exonuclease activity of the four fusion proteins increased gradually from the lowest concentration tested (5 \(\mu\)M of MgCl\(_2\)), reaching its maximum value at 10 mM, above which the activity drastically decreased. The diminution of exonuclease activity observed at high metal ion concentration for all proteins was probably due to the metal binding to the substrate, rather than to an inhibitory effect on the protein.

**Comparative kinetic analysis of His-Tag-Spn PolI fusion proteins**

To investigate the role of the Asp\(^{10}\), Glu\(^{88}\) and Glu\(^{114}\) residues in the 5’-3’ exonuclease activity of Spn PolII, we determined the kinetic parameters of the D10G, E88K and E114G mutant enzymes and of Spn PolII-(his) control protein (Table 2). The apparent catalytic rate \((k_{cat})\) and \(K_m\) values were calculated from experiments performed using 3’-\(^{32}\)P-end-labeled 17-mer oligonucleotide annealed to M13mp2 ssDNA as substrate for the exonucleolytic reaction. The scheme of the reaction is as follows:

\[
E + D \rightleftharpoons E \cdot D \rightleftharpoons E \cdot D_{n-1} \rightleftharpoons E \cdot D_{n-2} \rightleftharpoons E \cdot D_{n-3} \rightleftharpoons E \cdot D_{n-4} \rightleftharpoons \ldots \rightleftharpoons E \cdot D_{n} \rightarrow E + D_n
\]
Where $E$ represents the enzyme, $D_n$ is the original 17-mer primer bound to DNA, $D_{n-1}, D_{n-2}, D_{n-3}$ and $D_{n-4}$ are the products of excision of one by one mononucleotide ($dNMP$), $Dm$ is the final product of the reaction. The side pathways (downwards pointing arrows) leading to dissociation of complexes $E·D_{n-1}, E·D_{n-2}, E·D_{n-3}$, with rate constants $k_a, k_b, k_c$. The reverse association of the enzyme and the reaction products ($D_{n-1}, D_{n-2}, D_{n-3}$) can occur via $k_a, k_b, k_c$ (upward point arrows).

In order to minimize the contribution of both initial DNA:protein complex formation ($k_0/k_0$) and the high dissociation rate at 11-mer product ($k_{dis}$) (see Fig. 1), we measured the velocity of the conversion of 15-mer substrate to 14-mer product ($v_{15,14}$) and of 14-mer substrate to 13-mer product ($v_{14,13}$) (see details in Experimental Procedures). In addition, association of the enzyme and the substrates of the reactions analyzed ($D_{15}$ and $D_{14}$) was minimized, since at the time periods assayed the concentration of $D_{15}$ and $D_{14}$ was very small in relation to original primer ($D_{17}$) (see Fig. 1). Moreover, correction for further conversion of both substrate and product to shorter products (see details in Experimental Procedures), allow us to neglect association and dissociation rates downstream of the target reaction. The Lineweaver-Burk plots of $[1/v]$ versus $[1/\text{primed M13mp2}]$ are depicted in Figure S1 and they were used to calculate the kinetic constants shown in Table 2. The apparent $K_m$ and $k_{cat}$ values were similar for the two catalytic events analyzed for each protein except for the E88K mutant, where the rate of conversion of $D_{14}$ to $D_{13}$ (1.25 s$^{-1}$) was slightly higher than that of $D_{15}$ to $D_{14}$ (0.53 s$^{-1}$). This indicated that the kinetic constants were not significantly affected by the nature of the nucleotide present at the 5´ end of each substrate (dAMP or dGMP). The apparent $K_m$ of the E88K mutant was similar to that of Spn PolI-(his), and this mutation resulted in only a five fold reduction of the apparent $k_{cat}$. These results indicate that Glu$^{88}$ is not essential for the exonucleolytic reaction. By contrast, the D10G and E114G mutants
possessed $k_{cat}$ values ~6000-fold and ~100-fold lower than that of Spn PolI-(his), respectively. This sharp reduction of the apparent $k_{cat}$ values suggested an important role for Asp$^{10}$ and Glu$^{114}$ residues in the exonuclease catalytic event. The apparent $K_m$ value of D10G and E114G were lower than that of the control protein, implying that these mutations resulted in a higher DNA binding affinity. However, it is important to keep in mind that $K_m$ is not an equilibrium dissociation constant ($K_D$) because it is also affected by $k_{cat}$. Moreover, in our kinetic analysis of a catalytic even $x$, we can assume that the apparent kinetic constant measured is $K_m = (k_{-x} + k_{cat})/k_x$. Therefore, the fast catalytic rates of Spn PolI-(his) and E88K mutant could mask the actual affinity for the substrate yielding the high apparent $K_m$ values observed. Supporting this hypothesis, the $K_m$ previously obtained for Spn PolII ($K_m = 100$ nM, (23)) was also lower than that for Spn PolI-(his), probably due to the low $k_{cat}$ of wild-type enzyme ($k_{cat} = 0.11$ s$^{-1}$, (23)). The differences in $k_{cat}$ between both wild-type and Spn PolI-(his) enzymes can be explained because of the different purification procedures used. The His-Tag fusion proteins were obtained using a one day purification step, whereas Spn PolII had been previously purified over one week using a more complex protocol consisting of three chromatographic steps (23). Therefore, the inactivation of the wild-type enzyme was probably higher than that of the fusion proteins. This was borne out by the fact that the specific exonuclease activity of the enzymes prepared for the kinetic experiments were 14 U/nmol for Spn PolI-(his) and 1 U/nmol for Spn PolII. In conclusion, the kinetic analyses allowed the determination of the catalytic rate of the enzymes, but the apparent $K_m$ values were not a measure of the DNA binding capability of these His-tag fusion proteins. Therefore, we proceeded to determine the dissociation constant ($K_D$) to establish whether the substrate affinity was affected by the mutations.
DNA binding affinity of His-Tag fusion proteins

The effect of the mutations in Spn PolI-(his) on DNA:exonucleolytic domain dissociation equilibrium constant \( K_D \) was measured by filter binding assays using homogeneously \(^{32}\text{P}\)-labeled dsDNA substrate. The retention of DNA in the filters requires the interaction of the DNA with the protein. Since Spn PolI possesses two enzymatic domains, for which DNA is a substrate, either domain could bind DNA. To minimize interference of DNA interactions with the polymerase domain on binding of the substrate to the exonucleolytic domain of Spn PolI, we used assay conditions (see Experimental Procedures) in which no DNA binding was detected with Spn PolIc269 (not shown). This protein only contains the polymerase domain of Spn PolI and possesses the same \( K_m \) for DNA as the wild type protein (18).

Complex formation between each protein and DNA was measured as a function of protein concentration, and the saturation curves are depicted in Fig. 5. The data were fitted by direct nonlinear least-squares regression program to the equilibrium binding Equation A2 for independent sites:

\[
[C] = \left( \frac{[D][P]^n}{K + [P]^n} \right)
\]  

(A2)

where \([C]\), \([D]\) and \([P]\) represent concentrations of formed complex, DNA and free protein, respectively, \(K\) is the apparent dissociation rate constant \(K_D\) and \(n\) is the Hill coefficient. Transformation of Equation A2 into a linear form yielded Equation A3:

\[
\log \frac{[C]}{[D] - [C]} = n \log[P] - \log(K)
\]  

(A3)
The Hill plots of the data log([C]/([D]-[C])) versus log[P] are depicted in Fig. S3. The Hill coefficient \( n \) for the protein tested ranged from 0.8 to 1.2 indicating a stoichiometry of 1:1 for the protein:DNA complexes (as expected since the DNA substrate contains only one 5’ protruding end). Therefore, the apparent \( K_D \) of Spn PolI-(his) as well as of D10G, E88K, E114G and D10K (which displayed a non-measurable 5’-3’ exonuclease activity) mutant enzymes was calculated from Equation A4 at each protein concentration tested, and the averages are shown in Table 2.

\[
\log[K] = n \log[P] - \log \frac{[C]}{[D]-[C]} \quad \text{(A4)}
\]

Spn PolI-(his), and D10G, D10K and E88K mutant enzymes, showed a similar \( K_D \) value, ranging between 7 to 14 nM. These results show that differences of more than 6000 fold on the catalytic rate of the proteins are not reflected in the binding assay. Moreover, the results indicate that neither Asp\(^{10}\) nor Glu\(^{88}\) are involved in DNA binding. However, the E114G mutation produced a 13-fold increase of the \( K_D \), suggesting a direct involvement of Glu\(^{114}\) in DNA binding of Spn PolI through its exonucleolytic domain.
DISCUSSION

In recent years many structural and mutational data have emerged from studies on different 5’ nucleases, showing that catalysis is supported by divalent metal ions whose ligands are carboxylate residues that are highly conserved in all 5’ nucleases. A sequence alignment from ten bacterial DNA polymerases and related bacteriophage 5’ nucleases (13), revealed the presence of ten invariant or highly conserved carboxylate residues, nine of which appear to be important for the exonucleolytic reaction (reviewed in 1). The subsequent inclusion in the analysis of eight sequences of polymerase-dependent and independent 5’ nucleases from different bacteria plus the comparison of the 5’ nuclease domain of Eco PolI and the eukaryotic FEN-1 enzyme, revealed that only six invariant residues are present in all prokaryotic and eukaryotic nucleases (7). Moreover, the multiple alignment of the 5’ nuclease domain of 38 prokaryotic and eukaryotic proteins (Fig. S2) allowed us to confirm and further support the previously observed conservation patterns of acidic residues.

We therefore carried out a mutational analysis of the 5’ nuclease domain of Spn PolI-(his), to gain insights into the structure-function relationships in 5’ nucleases. Three carboxylate residues of the 5’ nuclease domain of Spn PolI were analyzed: Asp^{10}, Glu^{88} and Glu^{114}. Although Glu^{88} is not an invariant amino acid, in 82 % of the sequences aligned this position is occupied by an acidic residue (Fig. S2). In addition, some of the amino acids surrounding Glu^{88} are invariant and have been shown to be important for the 5’ nuclease activity of Eco PolI (35) and 5’ nuclease from T5 (36, 37). Therefore, it has been proposed that the Glu^{88} region could be involved in DNA binding (7). However, our results revealed that substitution of Glu^{88} by Lys did not drastically alter
either the catalytic constants or the DNA binding ability of Spn PolI. Thus, it seems that Glu\textsuperscript{88} specifically does not play an essential role in the exonucleolytic reaction.

The crystal structures of Taq Pol, T5 nuclease, and T4 RNase H, and the related eukaryotic flap endonuclease \textit{Mj}FEN-1 have been recently reported (9-11, 38). Although they differ in detail, all four share an active site made up of conserved residues that coordinate, at least, two divalent metal ions. Despite the high identity of the residues that constitute the active site, differences exist in the number of metals bound and the distances between them. The 5' nuclease from T5, T4 RNase H, and \textit{Mj}FEN-1, contain two divalent metal ions separated by 5-8 Å, whereas in the exonucleolytic domain of Taq Pol, three different metal binding sites were detected: site I (crystals soaked in Zn\textsuperscript{2+}), and sites II and III (crystals soaked in Mn\textsuperscript{2+}). Sites I and II are separated by about 5 Å, and are each about 10 Å from site III. The 5 Å distance between sites I and II in Taq Pol is similar to the 3.9 Å that separates the two metal ions bound to the 3'-5' exonucleolytic domain of the Klenow fragment (39). Accordingly, a two-divalent metal ion mechanism, analogous to that used by the 3'-5' proofreading exonuclease of Klenow fragment, was proposed for the 5'-3' exonuclease of Taq Pol (7, 9). Such a mechanism implies that two divalent metal ions (I and II) promote the formation of a hydroxyl ion that attacks the scissile phosphodiester bond (MgA) and stabilize the oxyanion leaving group (MgB) and the pentavalent transition state formed during the reaction (MgA and MgB).

The two divalent metal ions bound at the active site of T4 RNase H and 5' nuclease from T5 are further apart (7 and 8 Å, respectively). This greater distance makes it difficult to propose a mechanism similar to that of the 3'-5' exonucleases. It is more likely that the divalent metal ions act independently in the exonucleolytic reaction, one
of them (located at Me1 metal binding site) being essential for catalysis and the other (at the Me2 metal binding site) playing a more indirect role, probably in substrate binding (11).

The alignment of the 5′-3′ exonuclease domains of Spn PolI and Taq Pol, and the 5′ nuclease from bacteriophage T5 (derived from the multiple alignment of the 38 nuclease depicted in Fig. S2) revealed a high sequence homology, Spn PolI possessing a 62% and a 51% similarity with Taq Pol and T5 respectively (Fig. 6). This homology allowed us to build a three-dimensional model of the exonucleolytic domain of Spn PolI (residues 1-290) based on the crystal structures of the Taq Pol and the T5 enzyme. The model predicts that the 5′-3′ exonucleolytic domain of Spn PolI should adopt a overall conformation similar to that of those nuclease, consisting of a central β-sheet, surrounded by clusters of helices on either side, and organized in two subdomains. The superimposition of the 3-D model on the structure of the 5′-3′ exonucleolytic domain of Taq Pol (Fig. 7A) and T5 5′ nuclease (Fig. 7B), clearly revealed that most of the secondary structural elements are conserved in all three proteins (particularly between Spn PolII and Taq Pol) and that they are largely present in the same relative positions. Another interesting feature of the Spn PolII model is the presence of a helical arch similar to that of the T5 nuclease, which has been proposed as being important for the threading mechanism of the structure-specific endonuclease activity of these types of enzymes (10). In our model, the helical structure is composed of two helices made up of 36 residues, smaller than the three-helix arch of 44 residues present in the bacteriophage nuclease (the corresponding region in Taq Pol is crystallographically disordered (9)).

The predicted active site of the nuclease domain of Spn PolII (Fig 8) is located in a similar position to those of the crystallized enzymes. The nine acidic residues conserved
in the prokaryotic nucleases (Fig. 6 and Fig. S2), including the Asp\(^{10}\) and Glu\(^{114}\), cluster in a sphere of 11 Å radius that should accommodate the divalent metal ions. Similar radii are observed in the spheres containing the analogous residues in 5’ nuclease from T5 (9 Å) and Taq Pol (10.5 Å), indicating that, in an analogous manner, two or three metal ions could be bound at the active site of Spn PolI exonuclease.

From our model of Spn PolI, we can also infer that Asp\(^{190}\) (residue conserved among all prokariotic 5’ nucleases and analyzed previously by the mutation D190A (23)) is located at the active site of the exonucleolytic domain. Analysis of this mutant showed that the substitution of Asp\(^{190}\) by Ala produced a sharp reduction of the \(k_{\text{cat}}\) for the exonuclease activity, which was accompanied by a deficiency in catalytic DNA:protein complex formation as well as a marked preference of the protein for Mn\(^{2+}\) over Mg\(^{2+}\) and an alteration of the optimal concentration of Mn\(^{2+}\) for catalysis (23). These data suggest that Asp\(^{190}\) is an essential residue, being involved simultaneously in DNA binding and catalysis, and that its function is mediated by the metal coordination at the active site.

The crystal structure of T5 nuclease showed that Asp\(^{201}\) (counterpart of Asp\(^{190}\)) together with Asp\(^{153}\), Asp\(^{155}\) and Asp\(^{204}\) (counterparts of Asp\(^{139}\), Asp\(^{141}\) and Asp\(^{193}\) in Spn PolI) are within coordination distance of the metal ion located at the Me2 site, which has been implicated in substrate binding. Superimposition of the active site of the Spn PolI model on the structure of the T5 enzyme (not shown), revealed a RMS deviation between the Asp\(^{190}\)-Spn PolI residue and the bacteriophage nuclease Asp\(^{201}\) of only 2.6 Å, supporting the involvement of Asp\(^{190}\) in the coordination of the metal ion bound at the putative Me2 site in Spn PolI.

It can be also inferred from our model that Asp\(^{10}\) and Glu\(^{114}\) could coordinate the metal ions bound at the active site, so that these residues, which are present in all prokaryotic
and eukaryotic 5’ nucleases, could be essential for its nuclease activity. The available structural data indicate that these residues are involved in binding a divalent metal ion at the active site (Asp\textsuperscript{10} in Me1 of T5 nuclease and site I of Taq Pol, and Glu\textsuperscript{114} in Me1 of T5 and site III of Taq Pol). However, our results show that substitution of Asp\textsuperscript{10} or Glu114 by Ala drastically reduces exonuclease activity of Spn PolII without changing its Mn\textsuperscript{2+} and Mg\textsuperscript{2+} dependency. Thus, they indicate that these residues do not coordinate metal ions. Nevertheless, they do not rule out that, upon removal of one of these putative ligands, metal ions can remain bound to the active site of the exonucleolytic domain of Spn PolII through interactions with other acidic residues, although it should be expected an alteration of affinity for the metals as was the case for the D190A mutant of Spn PolII (23). Moreover, in our structural model of Spn PolII, Asp\textsuperscript{64}, Asp\textsuperscript{116} and Asp\textsuperscript{117} (metal ligands in other 5’ nucleases) cluster within a sphere of \(\sim 7.5 \text{ Å} \) radius (Fig. 8) and the conformational freedom of their side chains could allow an arrangement to bind a divalent metal ion at the putative Me1 site. Therefore, at least, we can infer that Asp\textsuperscript{10} and Glu\textsuperscript{114} do not have a crucial role in metal coordination at the nuclease active site of Spn PolII.

The essential role of Asp\textsuperscript{10} in exonuclease activity is unquestionable, since mutations at this position in all nucleases studied led to an almost total inactivation of the protein, as reported for D13N mutation in Eco PolII (7), D21N in Mtb PolII (8), D34A in \(h\text{FEN-1} \) (40), and D19N in T4 RNase H (12). We have demonstrated that in Spn PolII the Asp\textsuperscript{10} is involved exclusively in the catalytic event of the exonucleolytic reaction, because the substitution D10G produced a \(~6000\)-fold reduction of the catalytic rate without affecting the DNA binding capability. A similar role has been proposed for the analogous Asp\textsuperscript{34} residue in \(h\text{FEN-1} \), since its D34A mutant is inactive in cleavage but not in DNA binding (40). Our experimental results suggest that the Asp\textsuperscript{10} in Spn PolII is
not critical for metal coordination. Therefore, this residue could be essential by acting like a general base during the reaction, activating a water molecule and generating the hydroxyl group required for nucleophilic attack. For this purpose, Asp\textsuperscript{10} must be in close proximity to the metal ion bound to the putative Me1 site as shown by our nuclease domain model of the pneumococcal enzyme. This metal should stabilize the pentavalent transition state formed during the reaction and therefore would be essential for catalysis.

The role played by Glu\textsuperscript{114} in exonucleolytic reaction is less clear. This amino acid of Spn PolII, as well as the equivalent residue of hFEN-1 (40), seems to be involved in substrate binding, and has been considered less essential than other residues analyzed for the exonuclease activity of Eco PolII (7) and Mtb PolII (8). The structural data are highly contradictory, since the counterpart of Glu\textsuperscript{114} seems to be involved in binding of the catalytic metal ion (Me1) in T5 nuclease (10) and MjFEN-1 (38), does not interact with metal ions at the T4 RNase H active site (11), and is part of the metal binding site III in Taq Pol (for which a nonessential role has been proposed (9)). Thus, it is likely that Glu\textsuperscript{114} of Spn PolII will be at the exonuclease active site, as shown by the Spn PolII exonuclease conformational model, contributing to the extensive hydrogen-bonding/electrostatic network surrounding both metal ions, rather than interacting in a direct or indirect manner with any one of them. If this hypothesis is correct, substitution of Glu\textsuperscript{114} by Gly should not produce a loss of a coordination with one or two metal ions (as indicated by our results) but a reorientation of the metal(s) at the active site, making the catalytic center less accessible to the DNA substrate and reducing the catalytic efficiency of the reaction.

Summarizing the above results and the available data from other 5’ nucleases, it seems that the involvement of two metal ions playing different roles is the more plausible
mechanism for exonucleolytic reaction in 5’ nucleases. A water molecule (probably activated by Asp\textsuperscript{10} or an equivalent residue) would be responsible for the nucleophilic attack of the phosphodiester bond. The divalent metal ion bound to the Me1 site would be essential for the reaction, stabilizing the generated transition state, whereas the metal ion bound to the Me2 site positions the DNA correctly at the active site for cleavage (probably by interacting with the 5’-end). Thus, the role played by the Me2 site would be also important for the reaction, since an incorrect orientation of the DNA at the active site would lead to a drastic reduction of the exonuclease activity. Therefore its function may not be as indirect as previously proposed (8, 12). This mechanistic model does not necessary exclude the possibility of a reaction similar to that of the 3’-5’ exonuclease of the Klenow fragment proposed for Taq Pol. The 5’ nucleases are able to recognize a wide variety of DNA substrates with different structures and to catalyze endo- or exo-nucleolysis, thus they must possess a very flexible active site capable of adopting different conformations depending on the substrate and the reaction taking place. It is possible that the differences of active site architecture seen in distinct 5’ nucleases reflect the different functional conformations that these active sites can adopt. For example, in the case of Taq Pol, one metal binding site was detected in presence of Zn\textsuperscript{2+} whereas two different sites appeared in presence of Mn\textsuperscript{2+}. This could be taken as an indication of the flexibility of its active site, being capable of adopting in vivo two different conformations depending of the requirements of the enzyme. By contrast, is also possible that the differences between distinct nucleases are a consequence of the individual substrate specificities of the enzymes.
ACKNOWLEDGMENTS

We thank Dr M. Espinosa, Dr L. Blanco and Dr S.W. Elson for helpful discussions and critical reading of the manuscript. We are grateful to Dr T. A. Steitz and S. H. Eom for provided the refinement of the tertiary structure of Taq Pol and to Dr A. Ceska for the initial model of the active site of Spn PolI based in the three-dimensional structure of T5 nuclease and helpful discussion.
REFERENCES


FIGURE LEGENDS

Fig. 1. Exonucleolytic reaction catalyzed by Spn PolI-(his) on 3’-end-labeled oligonucleotide annealed to ssDNA from M13mp2. An autoradiogram is shown of the gel-fractionation of the reaction products obtained from the exonucleolytic reaction performed with 0.042 U of exonuclease activity from Spn PolI-(his) and 100 nM of DNA substrate. The experiment was carried out as described in Experimental Procedures withdrawing samples at the times indicated.

Fig. 2. Overexpression and purification of Spn PolI-(his). 15 µg of crude extracts from 3.5 h IPTG-induced cultures of E. coli BL21(DE3) carrying the indicated plasmids were fractionated in a 8% PA gel. The protein content was analyzed by Coomassie brilliant blue stain (A) and the nuclease activity of the extract was assayed by DNase gel assay (B). The purity of Spn PolI-(his) at each purification step was evaluated by 0.1% SDS/8% PAGE and staining the gel with Coomassie brilliant blue (C). Lane C, 225 ng of purified Spn PolI; lane 1, 1/1400 volume of crude extract; lane 2, 1/1400 volume of the sample applied to chelating sepharose resin; lane 3, 1/1400 volume of the sample recovered after application to the resin; lane 4, 5 and 6, 1/600 volume of each of the three wash steps with 10 mM imidazole; lane 7, 1/500 volume of the wash step with 40 mM imidazole; lane 8, 1/250 volume of the eluted fraction with 100 mM imidazole. St, mixture of polypeptides of known molecular weight.

Fig. 3. Exonuclease activity of His-Tag fusion proteins on 5’-end-labeled oligonucleotide annealed to M13mp2 ssDNA. Reaction mixtures contained 1.5 pmol of DNA and 1 mM of MnCl₂. Samples were removed at 10 min, 30 min, 1h, 2h and 16
The concentration of the enzymes in the reaction were: 0.05 pmol of Spn PolI-(his) (WT), 2.2 pmol of Spn PolIE114G-(his), 0.18 pmol of Spn PolIE88K-(his), 0.26 pmol of Spn PolID10S-(his), 0.62 pmol of Spn PolID10G-(his), 0.096 pmol of Spn PolID10V-(his), 0.12 pmol of Spn PolID10T-(his), 0.33 pmol of Spn PolID10K-(his), 0.91 pmol of Spn PolID10A-(his), 0.35 pmol of Spn PolID10Q-(his) and 0.21 pmol of Spn PolID10R-(his).

**Fig. 4. Metal-dependency of the exonuclease activity of fusion proteins.** Reactions were performed with 3 pmol of 3'-32P-end labeled oligonucleotide annealed to M13mp2 ssDNA and increasing concentrations of MgCl2 or MnCl2. To determine the exonuclease activity in the presence of Mg2+, the incubation time was 30 min for reactions catalyzed by Spn PolI-(his) (WT) and Spn PolIE88K-(his), or 2 h for those catalyzed by the Spn PolID10G-(his) and Spn PolIE114G-(his). Reaction time in the presence of Mn2+ was 10 min for Spn PolI-(his) (WT) and Spn PolIE88K-(his), 30 min for Spn PolIE114G-(his), and 2 h for Spn PolID10G-(his).

**Fig. 5. Detection of DNA:protein complex formation by filter assay.** The binding reactions with the indicated proteins were performed as described in Experimental Procedures. Concentrations of retained DNA:protein complexes obtained upon increasing protein concentration are depicted. Experimental data were treated through a non-linear regression analysis program (Curve Expert 1.3, Microsoft Corporation).

**Fig. 6. Amino acid sequence alignment of the 5' nuclease from T5 and the 5'-3' exonucleaseolytic domains of Spn PolI and Taq Pol.** This alignment was derived from the multiple sequence alignment of 38 exonucleaseolytic domains of the family A.
polymerases, and prokaryotic and eukaryotic nucleases (see Fig. S2). The numbers indicate the amino acid position relative to the N terminus of each sequence. Invariant residues are shown as white letters over black background whereas similarity is indicated on gray background. Conserved secondary structural motifs (α-helices as hollow rectangles and β-strands as arrows) are indicated at the top of the aligned sequences. Invariant acidic residues at the active site are indicated at the bottom of the aligned sequences. Symbols: (●) residues present in both prokaryotic and eukaryotic nucleases; (❍) residues present exclusively present in the prokaryotic enzymes and (■) acidic residue conserved in prokaryotic enzymes.

**Fig. 7. Structural modeling of the putative 5’-3’ exonuclease domain of Spn PolI.** Solid ribbon represents the optimally superimposed polypeptide backbone of the energy-minimized model of the 5’-3’ exonuclease domains of Spn PolI (red) and (A) Taq Pol (green) or (B) 5´ nuclease from T5 (cyan). The amino acid side chains bound to metal ions at the active site of Taq Pol (A) or T5 5´ nuclease (B) are drawn with solid sticks (yellow). Nt and Ct indicate the N- and C-terminal ends, respectively.

**Fig. 8. Close-up view of the active site of the structural model of Spn PolI 5’-3’ exonuclease domain.** The polypeptide backbone of Spn PolI is represented as a solid ribbon (cyan). The conserved-amino acid side chains are represented as solid sticks with the carbon atoms in green and the oxygen atoms in red. The seven residues present in both prokaryotic and eukaryotic nucleases are labeled in yellow and those exclusively present in the prokaryotic enzymes are labeled in white.
**Figure S1. Kinetic analysis of 5´-3´ exonuclease activity.** The Lineweaver-Burk plots of Spn PolI-(his) derivatives are depicted. Velocities of conversion of $v_{15,14}$ (■) and $v_{14,13}$ (▲) were measured at different substrate concentrations as indicated in Experimental Procedures using 22 nM Spn PolII-(his) (WT), 9 nM Spn PolIE88K-(his), 17 nM Spn PolIE114G-(his) or 41 nM Spn PolID10G-(his). The figures are the average of at least three independent experiments in which exonucleolytic rates were tested at four time points 2.5, 5.0, 7.5 and 10 min for WT and the E88K mutant, and 10, 15, 20 and 30 min for the E114G and D10G mutants. The double reciprocal plots were fitted by a linear least-squares regression analysis.

EMBL accession no. Q9ZIG3), and *Synechocystis* sp. (*Synechocystis_Pol*; EMBL accession no. Q55971), and the DNA polymerases I from *Bacillus stearothermophilus* gene *polG1*-encoded (Bst_Pol; EMBL accession no. Q45458) and gene *polA*-encoded (Bst_PolA; EMBL accession no. P52026). Nine additional 5’ nuclease were added to the alignment: the gene 6-encoded exonuclease from bacteriophage T3 (T3_exo; EMBL accession no. P20321), the RNaseH from bacteriophage T4 (T4_exo; EMBL accession no. P13319), the 5’ nuclease gene D15-encoded from bacteriophage T5 (T5_exo; EMBL accession no. P06229), the exonuclease gene 6-encoded from bacteriophage T7 (T7_exo; EMBL accession no. P00638), the exonuclease from *Mycobacterium tuberculosis* (Mtb_exo; EMBL accession no. Q07700), the exonuclease IX from *Escherichia coli* (Eco_exo; EMBL accession no. P38506), and 3 potential 5’-3’ exonucleases from *Bacillus subtilis* (Bsu_exo; EMBL accession no. P54161), *Mycoplasma genitalium* (Mge_exo; EMBL accession no. Q49406) and the gene *polA*-encoded from *Mycoplasma pneumoniae* (Mpn_exo; EMBL accession no. P75403). Additionally, the eukaryotic flap type I endonucleases from human (hFEN-1; EMBL accession no. P39748) and arquibacteria *Methanococcus jannaschii* (MjFEN-1; EMBL accession no. Q58839) were included. Residues identical in at least 80% of the sequences are highlighted on a black ground and those similar in at least 80 % are on a gray shaded ground. The resulting consensus sequence is shown.

**Figure S3. Hill plots of the data presented in Fig. 5.** Plots were generated by a linear least-squares fitting of the data. The Hill coefficient *n* were obtained for each protein from the slope and, the values are shown in the figure.
Table 1. Specific enzymatic activities of His-Tag fusion proteins. Polymerase activity was assayed on activated calf thymus DNA. The 5’-3’ exonuclease activity was determined from the experiments performed with 5’-32P-end-labeled oligonucleotide annealed to M13mp2 ssDNA. The enzymatic activities are expressed as units per µg of protein (U·µg⁻¹). The figures are the average of at least three independent experiments. Standard deviations are indicated.

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>Polymerase activity (U·µg⁻¹)</th>
<th>5’ Nuclease activity (U·µg⁻¹)</th>
<th>5’ Nuclease/Polymerase (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spn PolI-(his)</td>
<td>0.43±0.06</td>
<td>(7.2±0.8)·10⁻³</td>
<td>100</td>
</tr>
<tr>
<td>D10Q</td>
<td>0.42±0.03</td>
<td>&lt;1.0·10⁻⁵</td>
<td>&lt;0.14</td>
</tr>
<tr>
<td>D10S</td>
<td>0.46±0.04</td>
<td>&lt;1.4·10⁻⁵</td>
<td>&lt;0.08</td>
</tr>
<tr>
<td>D10V</td>
<td>0.23±0.02</td>
<td>&lt;3.7·10⁻⁵</td>
<td>&lt;0.96</td>
</tr>
<tr>
<td>D10T</td>
<td>0.42±0.03</td>
<td>&lt;2.9·10⁻⁵</td>
<td>&lt;0.41</td>
</tr>
<tr>
<td>D10K</td>
<td>0.30±0.07</td>
<td>&lt;1.1·10⁻⁵</td>
<td>&lt;0.22</td>
</tr>
<tr>
<td>D10R</td>
<td>0.29±0.03</td>
<td>&lt;1.7·10⁻⁵</td>
<td>&lt;0.35</td>
</tr>
<tr>
<td>D10G</td>
<td>0.32±0.07</td>
<td>(1.5±0.4)·10⁻⁴</td>
<td>4.69</td>
</tr>
<tr>
<td>D10A</td>
<td>0.25±0.04</td>
<td>(5.8±0.9)·10⁻⁶</td>
<td>0.14</td>
</tr>
<tr>
<td>E88K</td>
<td>0.66±0.11</td>
<td>(1.9±0.3)·10⁻³</td>
<td>17.2</td>
</tr>
<tr>
<td>E114G</td>
<td>0.38±0.03</td>
<td>(6.1±1.0)·10⁻⁵</td>
<td>0.95</td>
</tr>
</tbody>
</table>
Table 2. 5’-3’ exonuclease kinetic constants of His-Tag fusion proteins. Exonuclease rates were measured at different substrate concentrations as indicated in Experimental Procedures. The $K_m$ and $k_{cat}$ values are the average of those obtained for the two catalytic events analyzed in Figure S1. The $K_D$ values were determined from filter binding assays as described in the text. ND: no determined.

<table>
<thead>
<tr>
<th>Enzyme</th>
<th>$K_m$ (nM)</th>
<th>$k_{cat}$ (s$^{-1}$)</th>
<th>$K_D$ (nM)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Spn PolI-(his)</td>
<td>721±35</td>
<td>5.1±0.6</td>
<td>8.9±3.4</td>
</tr>
<tr>
<td>Spn PolI E88K-(his)</td>
<td>777±65</td>
<td>0.9±0.3</td>
<td>6.6±2.9</td>
</tr>
<tr>
<td>Spn PolI E114G-(his)</td>
<td>328±50</td>
<td>(5.0±1.0)·10$^{-2}$</td>
<td>113±46</td>
</tr>
<tr>
<td>Spn PolI D10G-(his)</td>
<td>206±8</td>
<td>(8.7±1.4)·10$^{-4}$</td>
<td>11±5.2</td>
</tr>
<tr>
<td>Spn PolI D10K-(his)</td>
<td>ND</td>
<td>ND</td>
<td>14±6.9</td>
</tr>
</tbody>
</table>
Figure 1

Time (min): 0 2.5 5.0 7.5

17-mer

11-mer
Figure 2

A

B

C

Spn PolI or Spn PolI-

Spn PolI or Spn PolI-

Spn PolI-

Figure 2

A

B

C

Spn PolI or Spn PolI-

Spn PolI or Spn PolI-

Spn PolI-

Figure 2

A

B

C

Spn PolI or Spn PolI-

Spn PolI or Spn PolI-

Spn PolI-

Figure 2

A

B

C

Spn PolI or Spn PolI-

Spn PolI or Spn PolI-

Spn PolI-

Figure 2

A

B

C

Spn PolI or Spn PolI-

Spn PolI or Spn PolI-

Spn PolI-

Figure 2

A

B

C

Spn PolI or Spn PolI-

Spn PolI or Spn PolI-

Spn PolI-

Figure 2

A

B

C

Spn PolI or Spn PolI-

Spn PolI or Spn PolI-

Spn PolI-

Figure 2

A

B

C

Spn PolI or Spn PolI-

Spn PolI or Spn PolI-

Spn PolI-

Figure 2

A

B

C

Spn PolI or Spn PolI-

Spn PolI or Spn PolI-

Spn PolI-

Figure 2

A

B

C

Spn PolI or Spn PolI-

Spn PolI or Spn PolI-

Spn PolI-

Figure 2

A

B

C

Spn PolI or Spn PolI-

Spn PolI or Spn PolI-

Spn PolI-

Figure 2

A

B

C

Spn PolI or Spn PolI-

Spn PolI or Spn PolI-

Spn PolI-

Figure 2

A

B

C

Spn PolI or Spn PolI-

Spn PolI or Spn PolI-

Spn PolI-

Figure 2

A

B

C

Spn PolI or Spn PolI-

Spn PolI or Spn PolI-

Spn PolI-

Figure 2

A

B

C

Spn PolI or Spn PolI-

Spn PolI or Spn PolI-

Spn PolI-

Figure 2

A

B

C

Spn PolI or Spn PolI-

Spn PolI or Spn PolI-

Spn PolI-

Figure 2

A

B

C

Spn PolI or Spn PolI-

Spn PolI or Spn PolI-

Spn PolI-

Figure 2

A

B

C

Spn PolI or Spn PolI-

Spn PolI or Spn PolI-

Spn PolI-

Figure 2

A

B

C

Spn PolI or Spn PolI-

Spn PolI or Spn PolI-

Spn PolI-

Figure 2

A

B

C

Spn PolI or Spn PolI-

Spn PolI or Spn PolI-

Spn PolI-

Figure 2

A

B

C

Spn PolI or Spn PolI-

Spn PolI or Spn PolI-

Spn PolI-

Figure 2

A

B

C

Spn PolI or Spn PolI-

Spn PolI or Spn PolI-

Spn PolI-

Figure 2

A

B

C

Spn PolI or Spn PolI-

Spn PolI or Spn PolI-

Spn PolI-

Figure 2

A

B

C

Spn PolI or Spn PolI-

Spn PolI or Spn PolI-

Spn PolI-

Figure 2

A

B

C

Spn PolI or Spn PolI-

Spn PolI or Spn PolI-

Spn PolI-

Figure 2

A

B

C

Spn PolI or Spn PolI-

Spn PolI or Spn PolI-

Spn PolI-

Figure 2

A

B

C

Spn PolI or Spn PolI-

Spn PolI or Spn PolI-

Spn PolI-

Figure 2

A

B

C

Spn PolI or Spn PolI-

Spn PolI or Spn PolI-

Spn PolI-

Figure 2

A

B

C

Spn PolI or Spn PolI-

Spn PolI or Spn PolI-

Spn PolI-

Figure 2

A

B

C

Spn PolI or Spn PolI-

Spn PolI or Spn PolI-

Spn PolI-

Figure 2

A

B

C

Spn PolI or Spn PolI-

Spn PolI or Spn PolI-

Spn PolI-

Figure 2

A

B

C

Spn PolI or Spn PolI-

Spn PolI or Spn PolI-

Spn PolI-

Figure 2

A

B

C

Spn PolI or Spn PolI-

Spn PolI or Spn PolI-

Spn PolI-

Figure 2

A

B

C

Spn PolI or Spn PolI-

Spn PolI or Spn PolI-

Spn PolI-

Figure 2

A

B

C

Spn PolI or Spn PolI-

Spn PolI or Spn PolI-

Spn PolI-

Figure 2

A

B

C

Spn PolI or Spn PolI-

Spn PolI or Spn PolI-

Spn PolI-

Figure 2

A

B

C

Spn PolI or Spn PolI-

Spn PolI or Spn PolI-

Spn PolI-

Figure 2

A

B

C

Spn PolI or Spn PolI-

Spn PolI or Spn PolI-

Spn PolI-

Figure 2

A

B

C

Spn PolI or Spn PolI-

Spn PolI or Spn PolI-

Spn PolI-

Figure 2

A

B

C

Spn PolI or Spn PolI-

Spn PolI or Spn PolI-

Spn PolI-

Figure 2

A

B

C

Spn PolI or Spn PolI-

Spn PolI or Spn PolI-

Spn PolI-

Figure 2

A

B

C

Spn PolI or Spn PolI-

Spn PolI or Spn PolI-

Spn PolI-

Figure 2

A

B

C

Spn PolI or Spn PolI-

Spn PolI or Spn PolI-

Spn PolI-
Figure 3
Figure 4

![Graph showing 5' Nuclease activity (U/nmol of enzyme) vs. [MgCl$_2$] (mM) and [MnCl$_2$] (mM). The graph includes lines for WT, E88K, E114G, and D10G.]
Figure 5

WT

E88K

D10G

D10K

E114G
Figure 6
Figure 7
Figure 8
Figure S1

- **WT**
  - 1/V (s⁻¹) vs 1/[primed M13mp2] (nM⁻¹)
  - Line 14-13, 15-14

- **E88K**
  - 1/V (s⁻¹) vs 1/[primed M13mp2] (nM⁻¹)
  - Line 14-13, 15-14

- **D10G**
  - 1/V (s⁻¹) vs 1/[primed M13mp2] (nM⁻¹)
  - Line 14-13, 15-14

- **E114G**
  - 1/V (s⁻¹) vs 1/[primed M13mp2] (nM⁻¹)
  - Line 14-13, 15-14
Figure S3

WT

E88K

D10G

D10K

E114G

$n = 0.98 \pm 0.11$

$n = 0.77 \pm 0.16$

$n = 0.91 \pm 0.03$

$n = 1.22 \pm 0.21$

$n = 1.17 \pm 0.08$
Biochemical analysis of point mutations in the 5'-3' exonuclease of DNA polymerase I of Streptococcus pneumoniae: functional and structural implications
Monica Amblar, Mario García de Lacoba, Maria A. Corrales and Paloma López

J. Biol. Chem. published online March 7, 2001

Access the most updated version of this article at doi: 10.1074/jbc.M008678200

Alerts:
- When this article is cited
- When a correction for this article is posted

Click here to choose from all of JBC's e-mail alerts

This article cites 0 references, 0 of which can be accessed free at
http://www.jbc.org/content/early/2001/03/07/jbc.M008678200.citation.full.html#ref-list-1