DNA Hydrolytic Activity Associated with the *Ustilago maydis REC1* Gene Product Analyzed on Hairpin Oligonucleotide Substrates†

Saule Naureckiene‡ and William K. Holloman*

Hearst Microbiology Research Center, Department of Microbiology and Immunology, Cornell University Weill Medical College, New York, New York 10021

Received June 29, 1999; Revised Manuscript Received August 27, 1999

ABSTRACT: The *REC1* gene of *Ustilago maydis* functions in the maintenance of genome stability as evidenced by the mutator phenotype resulting from inactivation of the gene. The biochemical function of the Rec1 protein was previously identified as a 3′−5′-directed DNA exonuclease. Here studies on the mechanism of action of Rec1 were performed using radiolabeled oligonucleotide DNAs as substrates, enabling detection of single cleavage events after electrophoresis on DNA sequencing gels. The oligonucleotides that were utilized were designed to be self-annealing so that they formed hairpin structures. This simplified interpretation of the data since each molecule contained only one 3′-terminus. Analysis revealed that digestion proceeded by a distributive mode of action and that degradation of DNA was governed by an interplay between sequence context and conformation. The preferential substrate was DNA with a recessed 3′-end. It was discovered that the enzyme had abasic endonuclease activity, was capable of initiating at an internal nick, and had no preference for mismatched bases either internally or terminally. Endonucleolytic cleavage was 5′ to the abasic site.

The *REC1* gene of the fungus *Ustilago maydis* encodes a 3′−5′ exonuclease (*1*). Inactivation of the gene results in an extremely pleiotropic phenotype with defects in repair, recombination, genomic stability, mutation avoidance, and cell cycle checkpoint control (*2*, *3*). It is widely known from many studies in other systems that exonucleases play important roles in a number of these processes. For instance, exonucleases function in mutation avoidance by proofreading during DNA synthesis to remove misincorporated nucleotides, in repair by excision of tracts containing damaged or incorrectly base-paired residues, in recombination to generate invasive single-stranded ends, and in genome stability by preventing slippage of repeated sequences. The challenge posed by the *REC1* gene is understanding the molecular function in the array of genetic pathways affected by its inactivation.

Database searches have not turned up any relationship between Rec1 and other known exonucleases, but sequence similarity has been found between Rec1 and the Rad1 protein of *Schizosaccharomyces pombe* (*4*) and the Rad17 protein of *Saccharomyces cerevisiae* (*5*), both of which, like Rec1, function in cell cycle checkpoint control. Biochemical activity associated with these proteins has not yet been demonstrated, but there is indirect evidence that the *S. cerevisiae* Rad17 protein might contain nuclease activity (*5*). Mammalian homologues have also been identified (*6*−*9*), and preliminary evidence for exonuclease activity in the human homologue has been obtained (*8*), although there is not complete agreement on this issue (*7*). One possible clue about function has come from analysis of Rec1 protein structure by fold recognition and comparative modeling which has revealed an intriguing relationship to the sliding clamp PCNA[1] family of proteins (*10*). Given the multifarious use of PCNA sliding clamps in DNA metabolic processes, it is not hard to imagine that a protein with a similar structure might serve as a common DNA-tethering component of protein machines at work among several diverse molecular pathways.

Analysis of *REC1* gene expression in *U. maydis* has indicated a periodic mode of transcription in which the message level peaks in the S phase (*3*). This specific mode of expression is similar to that observed with genes dedicated to DNA synthesis, suggesting that the role of the gene is not in some constitutive function throughout the cell cycle. One possible explanation of a role for Rec1, albeit unorthodox, is that Rec1 might function as an external proofreading activity to some DNA polymerase perhaps dedicated to repair and recombination. This could account for the mutator activity of the mutant and the S phase-specific peak of message. On the other hand, certain properties of the *rec1* mutant are reminiscent of the phenotype of mismatch repair mutants of *S. cerevisiae*. Besides the elevated mutator

† This work was supported by Grant GM42482 from the National Institutes of Health.

* To whom correspondence should be addressed: Department of Microbiology, Cornell University Medical College, 1300 York Ave., New York, NY 10021. Telephone: (212) 746-6510. Fax: (212) 746-8587. E-mail: wkhollo@mail.med.cornell.edu.

‡ On leave from the Institute of Biochemistry, 2600 Vilnius, Lithuania.

[1] Abbreviations: buffer BB, 20 mM Tris-HCl (pH 8.5), 0.5 M NaCl, and 6 M guanidine hydrochloride; F, 1′,2′-dideoxyribofuranose; FPLC, fast protein liquid chromatography; IPTG, isopropyl β-D-thiogalactoside; NTA, nitrilotriacetic acid; PCNA, proliferating cell nuclear antigen; PCR, polymerase chain reaction; SDS, sodium dodecyl sulfate; T_m, melting temperature; TNE, 50 mM Tris-HCl (pH 8.5), 1 mM EDTA, and 0.1 M NaCl.

10.1021/bi991495b CCC: $18.00
Published on Web 10/07/1999

activity, *pms1* and *msh2* of *S. cerevisiae* are hyperactive for spontaneous allelic recombination, exhibit poor viability of meiotic progeny, and exhibit genomic instability (*11*, *12*), like *rec1* mutants (*2*). It is interesting to note that expression of genes in mismatch repair is also cell cycle-regulated and is S phase-specific (*13*). Other analogies are also intriguing. For instance, the *rad27* mutant of *S. cerevisiae*, which has a defective exo/endonuclease, exhibits elevated spontaneous recombination and is sensitive to DNA damage (*14*−*17*). Thus, there are a number of examples of provocative similarities in comparing the properties of *rec1* with those of mutants in other systems with defects in a variety of DNA metabolic processes.

Unfortunately, there are still no clear features of either the *REC1* gene, protein, or mutant that might serve to categorize the function firmly under one paradigm or another. Therefore, we thought it important to gather more biochemical information that might add weight to one or the other possible mechanisms for *REC1* gene function. In this study, we report a further analysis of the substrate specificity of the 3′−5′ exonuclease activity using a variety of oligonucleotide substrates. Our primary aim was to determine features of the DNA substrate that might provide hints about the function of the protein. Additionally, we explore the utility of oligonucleotides in a hairpin conformation as substrates for analysis.

EXPERIMENTAL PROCEDURES

Purification of the Rec1 Protein. For production of high levels of Rec1 protein, the expression plasmid pCM419, a derivative of pET14b (Novagen), was utilized. This plasmid was constructed so that it contained the open reading frame from the *REC1* gene which was modified by removal of the single 184 bp 3′-terminal intron. A 1.4 kb *Eag*I−*Fsp*I fragment from pCM450, which contains the *REC1* open reading frame deleted of the intron by a PCR method (*18*), was used to replace the intron-containing *Eag*I−*Fsp*I fragment of pCM391, a pET14b derivative containing the genomic version of the *REC1* gene described previously (*1*). Expression of the gene yields a Rec1 fusion protein containing a 20-residue leader sequence with a hexahistidine stretch. Cultures (500 mL) in LB medium containing 200 μg/mL ampicillin were started from a single colony and grown with vigorous aeration at 37 °C. When A_{600} reached 0.6, cells were induced by addition of 0.4 mM IPTG (isopropyl β-D-thiogalactoside) and harvested by centrifugation 3 h later. The cell pellet was resuspended in 25 mL of TNE [50 mM Tris-HCl (pH 8.5), 1 mM EDTA, and 0.1 M NaCl], and lysozyme was added to a concentration of 1.0 mg/mL. After 1 h at 0 °C, EDTA was added to a concentration of 10 mM, Triton X-100 was added to a concentration of 0.1%, and the mixture was incubated at 37 °C for 20 min. Cells were disrupted by three 30 s bursts of sonication with an immersion tip and centrifuged for 15 min at 20 000 rpm in the Sorvall SS34 rotor. The supernatant was discarded, and the pellet was washed by resuspension with the aid of a mechanical homogenizer in succession with 25 mL aliquots of TNE and 2 M NaCl, TNE and 2 M urea, and 0.1 M Tris-HCl (pH 8.5). The final pellet was redissolved in 25 mL of buffer BB [20 mM Tris-HCl (pH 8.5), 0.5 M NaCl, and 6 M guanidine hydrochloride] and loaded onto a column (10 mL bed volume) of the metal affinity matrix agarose Ni−NTA (nitrilotriacetic acid agarose, Qiagen) which was prepared by washing successively with 20 mL of 100 mM $NiSO_4$, 20 mL of H_2O, and 20 mL of buffer BB. After being loaded, the column was washed with 50 mL of buffer BB followed by washes of 20 mL each of buffer BB containing imidazole in steps of 300 mM and 1 M. An aliquot (1 mL) of the fraction containing Rec1 eluting with 300 mM imidazole was diluted with buffer BB to a protein concentration of 1 mg/mL as determined using BCA reagent (Pierce), and then DTT (dithiothreitol) was added to a concentration of 30 mM and Triton X-100 to a concentration of 0.01%. The sample was diluted with buffer A [50 mM Tris-HCl (pH 8.5), 1 mM EDTA, 1 mM DTT, 0.01% Triton X-100, and 10% glycerol] so that the guanidine hydrochloride concentration was reduced to 2 M. After 2 h, the sample was diluted with buffer A to a final protein concentration of 20 μg/mL. The sample was then loaded onto an FPLC MonoS column (Pharmacia) equilibrated with buffer A; the column was washed with 10 mL of buffer A and eluted with a 0 to 0.7 M NaCl linear gradient (10 mL total), and 0.5 mL fractions were collected. Fractions eluting at a NaCl concentration of approximately 0.5 M were pooled and used as the source of the enzyme in the experiments described herein. Enzyme preparations were stored at −70 °C in aliquots at a protein concentration of 200 ng/mL. Enzyme activity was assayed using as a substrate pBluescript II (Stratagene) plasmid DNA that had been cut with *Sau*3AI and labeled at the 3′-ends of the restriction fragements using [α-^{35}S]dCTP and the Klenow fragment of *Escherichia coli* DNA polymerase I as described previously (*1*). In general, reaction mixtures contained 80 fmol of 3′-^{35}S-labeled DNA ends. One unit of enzyme activity is defined as the amount of protein that renders 10 fmol of 3′-^{35}S-label soluble in 5% trichloroacetic acid in 30 min at 37 °C.

DNA Substrates. Oligonucleotides were obtained from Oligos, Etc. (Wilsonville, OR) and were labeled at the 5′-terminus with T4 polynucleotide kinase and [γ-^{32}P]ATP. In general, reaction mixtures (20 μL) contained 12 pmol of oligonucleotide and 20 units of T4 polynucleotide kinase (New England Biolabs, Beverly, MA) and were incubated at 37 °C for 1 h before processing. Unincorporated label was removed by fractionation on NENSORB cartridges (NEN Dupont). Plasmid DNA and 3′-^{35}S end-labeled restriction endonuclease fragments used as the substrate during purification of the Rec1 protein were prepared as described previously (*1*). ^{32}P-labeled oligonucleotides usually at a concentration of approximately 10 nM were self-annealed in 10 mM Tris-HCl, 1 mM EDTA, and 0.4 M NaCl by heating to 85 °C for 5 min and then incubating at 37 °C for 1 h. Oligonucleotides that were utilized in this study are listed in Table 1.

Exonuclease Assays. Reactions (20 μL) were carried out in 50 mM Tris-acetate (pH 8.5), 10 mM Mg^{2+} acetate, 1 mM dithiothreitol, 0.1 mM EDTA, 0.05 mg/mL bovine serum albumin, 0.3 pmol of oligonucleotide, and in general 0.5 unit of Rec1 exonuclease (as determined using 3′-^{35}S-labeled duplex DNA restriction fragments). Reactions were carried out at 37 °C unless otherwise indicated. Reactions were quenched by addition of 100 μL of a stop solution containing 95% formamide, 20 mM EDTA, 0.05% bromophenol blue, and 0.05% xylene cyanol. Samples were heated to 100 °C, chilled on ice, and loaded (20 μL) on a DNA

Table 1: Oligonucleotide Sets[a]

single-stranded vs double-stranded comparison	
DS-1	5′-ATTCCTAAGCAGCCAAGTGTTTTTACACTTGGCTGCTTAGGAGT-3′
SS-1	5′-ACACTTGGCTGCTTAGGAGT-3′
hairpin 44-mer oligonucleotides with terminal mismatches	
BR-1	5′-CTTATCTATTGTTGATAAACTTTTGTTTATCAACAATAGATAAG-3′
BR-2	5′-CTTATCTATTGTTGATAAACTTTTGTTTATCAACAATAGATAAA-3′
BR-3	5′-CTTATCTATTGTTGATAAACTTTTGTTTATCAACAATAGATAAT-3′
BR-4	5′-CTTATCTATTGTTGATAAACTTTTGTTTATCAACAATAGATAAC-3′
hairpin oligonucleotides with simple sequence tails	
GD-1	5′-TTTTTTTTTTCTTATCTATTGTTGATAAACTTTTGTTTATCAACAATAGATAAGTTTTTTTTTT-3′
GD-2	5′-TTTTTTTTTTCTTATCTATTGTTGATAAACTTTTGTTTATCAACAATAGATAAG-3′
GD-3	5′-CTTATCTATTGTTGATAAACTTTTGTTTATCAACAATAGATAAGTTTTTTTTTT-3′
RN-1	5′-CTTATCTATTGTTGATAAACTTTTGTTTATCAACAATAGATAAGAAAAAAAAAA-3′
RN-2	5′-AAAAAAAAAACTATTGTTGATAAACTTTTGTTTATCAACAATAGTTTTTTTTTT-3′
RN-3	5′-TTTTTTTTTTCTATTGTTGATAAACTTTTGTTTATCAACAATAGAAAAAAAAAA-3′
FA-2	5′-AAAAAAAAAACTATTGTTGATAAACTTTTGTTTATCAACAATAGTTTTTTTTTT-3′
FA-3	5′-TTTTTTTTTTCTATTGTTGATAAACTTTTGTTTATCAACAATAGAAAAAAAAAA-3′
DE-1	5′-CCCCCCCCCCCCTATTGTTGATAAACTTTTGTTTATCAACAATAGGGGGGGGGG-3′
DE-2	5′-GGGGGGGGGGCTATTGTTGATAAACTTTTGTTTATCAACAATAGCCCCCCCCCC-3′
hairpin oligonucleotides with abasic or internal mismatched site	
ZN-1	5′-CTTATCTATTGTTGATAAACTTTTGTTTATCAACAFTAGATAAGTTTTTTTTTT-3′
DL-2	5′-CTTATCTATTGTTGATAAACTTTTGTTTATCAACATTAGATAAGTTTTTTTTTT-3′
M	5′-TTTTGTTTATCAACAFTAGATAAGTTTTTTTTTT-3′

[a] The four-T residue stretch serving as a hinge is underlined.

sequencing gel of 8 to 15% acrylamide containing 7 M urea. Before loading, gels were prerun at 40 W (ca. 1600 V) until they were warmed to a surface temperature of ca. 60 °C on the glass plates. After electrophoresis, gels were fixed in 5% methanol and 5% acetic acid and then dried at 80 °C. Radioactivity was imaged in the dried gels using a Molecular Dynamics phosphorimager. Rates were determined by measuring the loss of radioactivity of the band representing the full-length substrate using ImageQuant software (Molecular Dynamics).

RESULTS

Enzyme Preparation. Initial efforts to purify the Rec1 protein after overexpression of the gene in *E. coli* indicated that it was sequestered in an insoluble form in inclusion bodies (*1*). An investigation into culture conditions that would allow production of the protein in soluble form was undertaken. Unfortunately, after an extensive analysis, no growth medium, temperature, or osmolality was found in which any soluble Rec1 protein was present in cell extracts prepared under nondenaturing conditions. Therefore, efforts focused on optimizing conditions for refolding the protein after solubilization of inclusion bodies in 6 M guanidine hydrochloride. A number of parameters were examined systematically, including protein concentration, reducing conditions, solvent composition, and the regimen of denaturant removal. Empirical testing indicated that for optimal exonuclease activity it was necessary to refold the protein in two steps: first, by holding for 2 h on ice in 2 M guanidine hydrochloride at a protein concentration between 20 and 200 μg/mL, and in alkaline buffer containing at least 30 mM DTT, 0.01% Triton X-100, and 10% glycerol; second, by diluting the solution an additional 10-fold with neutral buffer containing no denaturant. The protein was purified by chromatography on an immobilized metal affinity column after solubilization of inclusion bodies in 6 M guanidine hydrochloride, refolded by the regimen outlined above, and further purified by FPLC on a monoS column. Exonuclease activity was eluted from the column using a salt gradient and cochromatographed with the Rec1 protein as assessed by analysis of chromatographic fractions after SDS gel electrophoresis. We note that since enzyme activity was constituted after refolding denatured protein, no meaningful assessment of the level of active protein molecules that are present can be made. Therefore, amounts of protein added are indicated in terms of units of exonuclease activity as determined using 3′-^{35}S-labeled plasmid DNA as the substrate.

Experimental Design. The method of our attack was to use oligonucleotide substrates to enable examination of nucleolytic activity at the nucleotide level. We used a gel electrophoresis assay to measure relative rates of hydrolytic activity. DNA substrates were oligonucleotides labeled with ^{32}P at the 5′-end and were designed so that we could examine the activity of the enzyme in discriminating DNA conformation, determine the ability to recognize mismatched base pairs, gauge relative base specificity, and test for endonuclease activity. In all cases, the rate of reaction was determined by measuring the loss of full-length substrate and was monitored quantitatively by phosphorimaging after electrophoresis of the oligonucleotide on a sequencing gel that resolves single nucleotide differences in length as discrete bands.

We performed pilot studies with a variety of substrates formed by annealing complementary single-stranded oligonucleotides with different designs, but found that the yield of duplex product was low when oligonucleotides were equimolar in hybridization reactions. To ensure complete

hybridization of the ^{32}P-labeled oligonucleotide with unlabeled complementary strand, it was necessary to provide a large excess of complementary strand. This presented the problem of performing reactions with an excess of potentially competitive unlabeled single-stranded DNA present or else required a tedious gel purification step to separate heteroduplexed oligonucleotides from unannealed strands and generally resulted in low yields. As an expedient to circumvent these difficulties, we chose to use self-annealing oligonucleotides as substrates. Since intramolecular association to form hairpin structures is the favored reaction in this situation, duplex formation is efficient and equimolar with respect to complementary sequences. An additional advantage offered by such substrates was the presence of a single 3′-end, thus eliminating a potential source of complication in interpretation. The substrate of choice was an oligonucleotide with a sequence of two inverted complementary stretches 20 residues in length joined through a hinge of four T residues.

Conformational Preference. In previous work using as substrate heterogeneous sequence DNA in the form of a mixture of 3′-labeled restriction enzyme fragments derived from plasmid DNA, a 5-fold preference for heat-denatured versus native DNA was noted. We retested conformational specificity using a 20-mer oligonucleotide with a heterogeneous sequence which was determined to contain no self-annealing stretches and compared this to its duplex form which was generated by self-annealing of an oligonucleotide with the same sequence and inverted complement to form a hairpin as described above. In this case, only a 2-fold difference in activity on the single-stranded oligonucleotide compared to the duplex was observed (Figure 1A). As controls for DNA substrate conformation, the double-stranded specific exonuclease III and the single-stranded specific ϵ subunit of DNA polymerase III were employed (panels B and C of Figure 1). In both cases, the activity that was exhibited was in accord with known properties of these enzymes. Exonuclease III was at least 10-fold more active on the hairpin duplex substrate than the single-stranded oligonucleotide, and ϵ was at least 50 times more active on the single-stranded substrate. In the case of Rec1, the basis for the lower level of discrimination for single-stranded versus duplex DNA with oligonucleotide substrates compared to what was observed in the previous study is not clear. There are a number of possible reasons, among which are changes in the affinity for the shorter oligonucleotide substrates, differential base specificities unveiled through the use of defined sequences, and altered rates of catalysis due to differences in sequence context.

Pause Sites. Examination of the digestion products of a duplex hairpin upon extended digestion and electrophoresis (Figure 2) revealed two features of the mechanism. First, the formation of the gradually descending ladder of intermediates was indicative of a distributive mode of hydrolysis. Second, a nonuniformity in the steady state accumulation of truncated products was apparent. In particular, the accumulation of products 22−27 residues in length (barrier$_{22-27}$) suggested that the run of seven T residues interrupted by a single G and including the hairpin loop presented a barrier to digestion. An additional barrier to digestion was evident in the accumulation of products 13−15 residues in length (barrier$_{13-15}$). This barrier slightly preceded a run of four T residues interrupted by a G. Such barrriers slowing digestion could be due to a number of factors, including a response to changes in secondary structure as the duplex is eroded as well as the inherent differential base specificity of the enzyme.

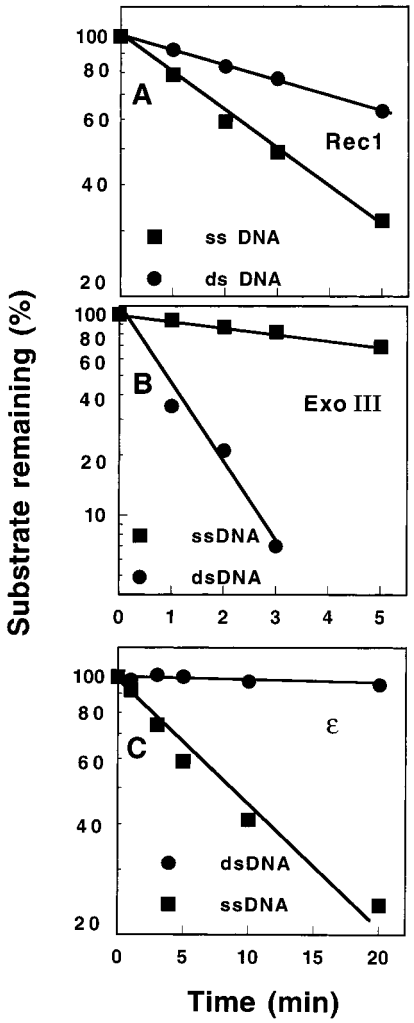

FIGURE 1: Specificity of exonucleases on oligonucleotides in single-stranded and hairpin duplex forms. 5′-^{32}P-labeled 20-mer single-stranded oligonucleotide SS-1 and duplex hairpin form DS-1 were digested with Rec1, *E. coli* Exo III, and the ϵ subunit of *E. coli* Pol III holoenzyme. Reactions were quenched at the indicated times, and digests were electrophoresed in a DNA sequencing gel under denaturing conditions. Radioactivity was digitized by phosphorimaging, and the time course of the reaction was determined by quantitating the level of the substrate band using ImageQuant software.

Tailed Hairpin Substrates. Following up the possibility that base specificity might influence hydrolysis, we investigated in more detail how homopurine or homopyrimidine tracts influenced the cleavage rate. The substrates that were tested were hairpin oligomers with 20 base pairs of duplex hinged through four T residues and containing 10-residue runs of homopurines and/or pyrimidines at the 5′- or 3′-end. When oligomers containing single-stranded runs of T residues at the 3′-terminus were examined (GD-1 and GD-3 in Figure 3A), it was evident upon analysis that there was little difference in the initial rate of digestion compared to that of the completely duplex hairpin control (BR-1). Addition of the 10-residue T stretch at the 5′-end of the hairpin (GD-2) resulted in a 2−3-fold increase in rate. This

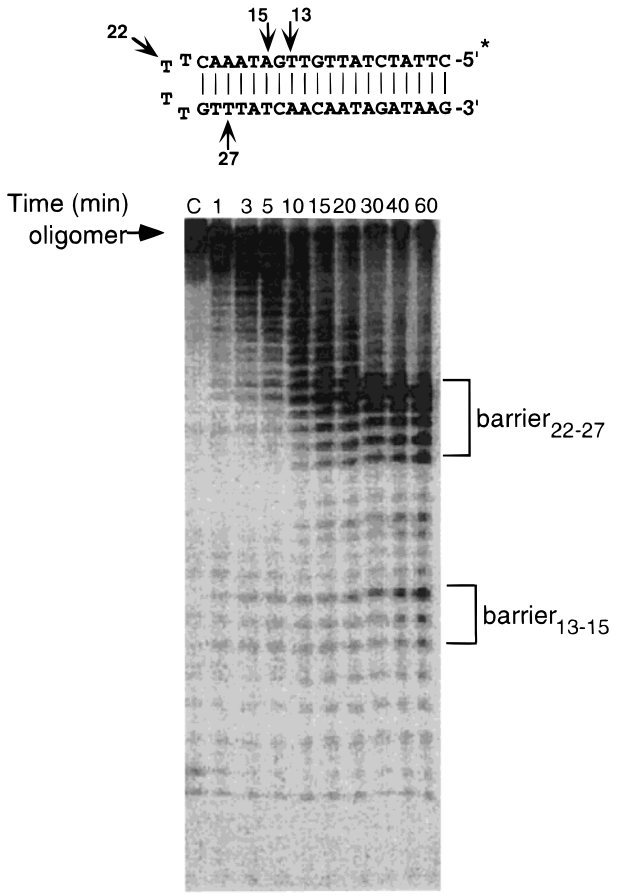

FIGURE 2: Distributive activity and nonuniformity of product accumulation during the course of digestion. A schematic illustration of the hairpin formed by self-annealing of oligonucleotide BR-1 with tandem inverted repeats joined through a run of four T residues. Digestions with Rec1 were performed with ^{32}P-labeled oligomer BR-1 for the indicated periods of time. The digitized phosphorimage is shown. Bands representing accumulated intermediates are bracketed, and their sites are denoted with arrows in the schematic.

indicates that a recessed 3′-end is a preferred substrate compared to a blunt end. Furthermore, a run of T residues in a single-stranded conformation does not appear to inhibit digestion in contrast to what was inferred from the pausing at T-rich runs in the duplex region of the hairpin.

Hairpin oligomers with 3′-runs of A and T residues in single-stranded and duplex conformations were tested for activity (Figure 3B). The initial rate of digestion of the hairpin oligomer with a single-stranded tail of 10 T residues at the 3′-end (GD-3) was almost 20 times faster than that of a completely duplex form of the oligonucleotide containing a complementary stretch of A residues (RN-2). This indicates that a run of T residues in the duplex conformation is more refractory to digestion than in the single-stranded conformation. Moreover, comparison of the initial rates of digestion of the heterogeneous hairpin duplex BR-1 in Figure 3A with that of the duplex homopolymer-tailed hairpin RN-2 in Figure 3B indicates that the run of T residues in the duplex form impedes digestion and is consistent with the pausing noted at T-rich sites in the heterogeneous sequence duplex control shown in Figure 2.

When the initial rates of digestion of oligomers terminating at the 3′-end in a single-stranded or duplex run of A residues (RN-1 and RN-3, respectively) were compared (Figure 3B), it was apparent that there was a slight preference for the

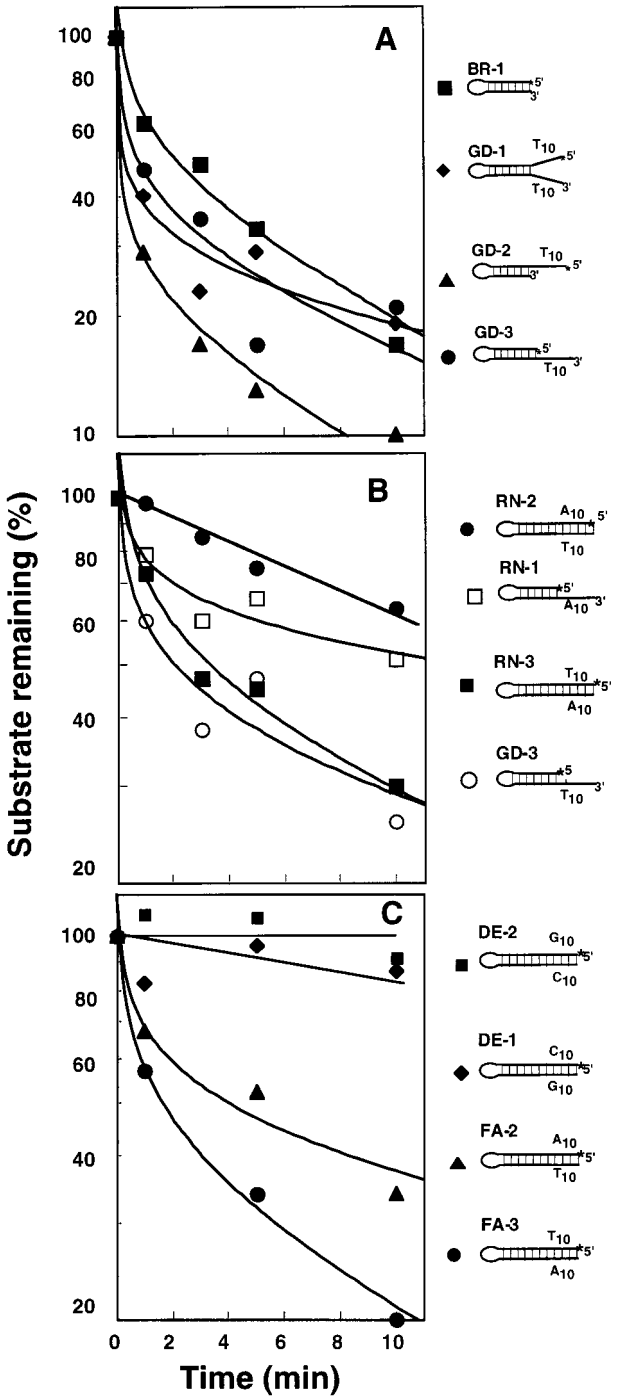

FIGURE 3: Hairpin substrates with terminal homopolymer stretches. Digestions were performed with ^{32}P-labeled hairpin oligonucleotides as shown schematically, products analyzed on DNA sequencing gels, and time courses determined from the digitized phosphorimages. All reactions whose results are depicted in a particular frame (A−C) were perfomed in parallel.

oligonucleotide with the A residue stretch in the duplex form (RN-3). When hairpins containing 3′-terminal duplex runs of G or C residues were tested for activity as substrates, it was found that the initial rates of digestion were slower than one-tenth of those observed with the hairpins containing terminal duplex runs of A or T residues (Figure 3C). In summary, these results indicate that there is an interdependent relationship of both sequence context and DNA conformation to Rec1 exonuclease activity.

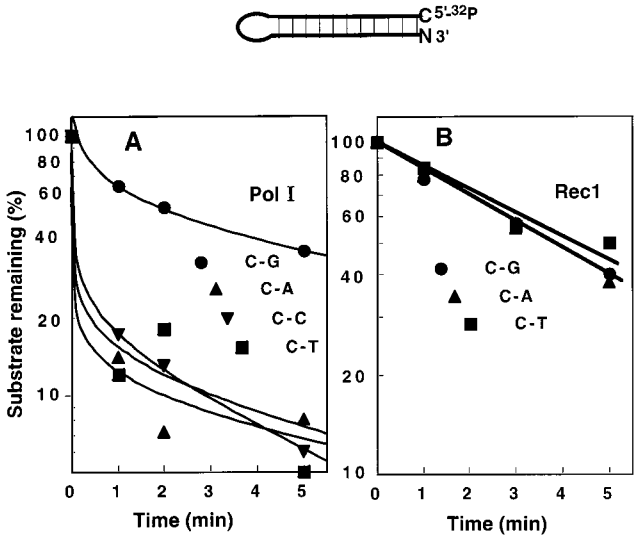

FIGURE 4: Hairpin substrates with terminal mismatched base pairs. Schematic illustration of hairpins formed with mismatched terminal base pairs. N represents any base paired with C. The asterisk represents the ^{32}P label. Digestions were performed with ^{32}P-labeled hairpins BR-1, BR-2, BR-3, and BR-4 using *E. coli* Pol I Klenow fragment (A) and Rec1 (B) and the products examined by electrophoresis on a DNA sequencing gel. Time courses were determined from the digitized phosphorimage.

Hairpin Oligonucleotides with Mispaired Termini. In light of the mutator activity of the *rec1* mutant, it was interesting to know if Rec1 has any intrinsic specificity for mismatched base pairs. A set of hairpin-forming oligomers with either a perfectly base paired C·G terminus or the three other possible combinations of mismatches with C residues were tested as substrates. The course of digestion was determined, and it was apparent that there was little difference in the activity of Rec1 on any of the hairpins regardless of the state of terminal base pairing (Figure 4B). As a control to ensure that the hairpin substrate was indeed terminated in mismatched base pairs, the Klenow fragment of DNA polymerase I was used. As expected, the Klenow fragment was highly active on all three hairpin oligonucleotides with mispaired termini compared to the completely base paired control (Figure 4A). Regardless of whether reactions with Rec1 were performed at 37, 30, or 25 °C, there was no change in the specificity. These results indicate that Rec1 exonuclease has little or no inherent preferential activity on mispaired residues.

Apurinic Endonuclease Activity. No endonuclease activity was observed in Rec1 preparations using covalently closed circular DNA as a substrate. However, we did detect nicking activity using the same DNA held for a few minutes at 70 °C in mild acid. Since such treatment promotes depurination, this suggested there was apurinic endonuclease activity. To confirm this observation, a hairpin oligonucleotide substrate containing a single abasic site was generated by synthesis with a 1′,2′-dideoxyribofuranose moiety replacing an internal deoxyribonucleoside. The abasic site was at residue 36 in the sequence of a 54-mer in the conformation of a hairpin duplex with a 3′ single-stranded tail of 10 T residues. As an additional control for endonuclease activity at a site of helix distortion, an identical substrate containing a mismatched T·T base pair was designed. Pyrimidine-pyrimidine base pairs are the most structurally disturbing combinations in terms of destabilizing and disrupting base stacking (*19*, *20*).

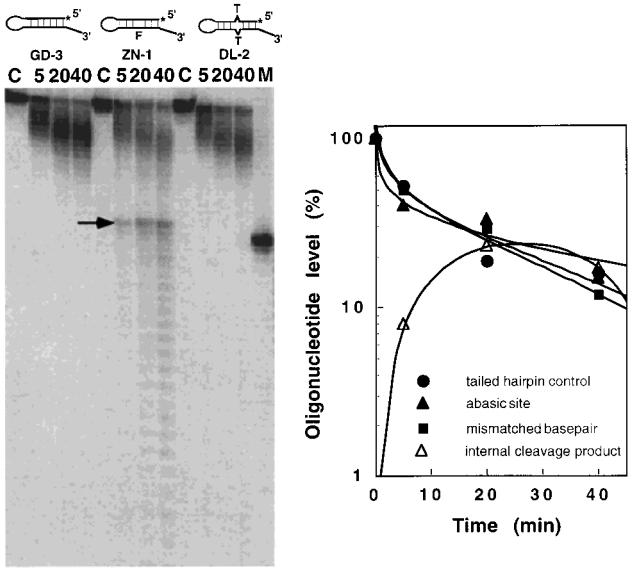

FIGURE 5: Rec1 has endonuclease activity at abasic but not at mismatched base pair sites. Hairpin oligonucleotides GD-3, ZN-1, and DL-2 are shown schematically. All have single-stranded tails of 10 T residues. GD-3 has a completely complementary sequence in the duplex region. ZN-1 has a 1′,2′-dideoxyribofuranose moiety (represented by F) replacing an internal deoxyribonucleoside at residue 36. DN-2 has a T·T mismatch corresponding to the same residues in ZN-1 corresponding to the base pair disrupted by the absence of a base at residue 36. Reactions were performed and analyzed as described previously. The arrow represents the 35-mer accumulating during the course of the reaction that was quantitated and termed internal cleavage product (△) in the quantitation of the phosphorimage. M is a 34-residue oligonucleotide marker.

During the course of digestion, it was observed under denaturing conditions that an oligonucleotide intermediate of 35 residues rapidly accumulated and then was itself degraded in a distributive fashion concomitant with degradation of the longer substrate and intermediates (Figure 5). The absence of a corresponding band in either the control or the mismatched base pair substrate indicates that Rec1 has abasic endonuclease activity. The length of the fragment of 35 residues establishes that cleavage is 5′ to the abasic site. The ladder of degraded intermediates less than 35 residues in length that appears concomitant with the cleavage at the internal abasic site indicates that Rec1 can act exonucleolytically at an internal nick as well as remove a dideoxyribofuranosyl phosphate residue. The Rec1 protein purified after overexpression of the gene in an *E. coli* strain without the gene encoding exonuclease III had abasic endonuclease activity with identical properties as described above.

DISCUSSION

There are several notable features of the nuclease activity associated with the Rec1 protein that have emerged from this study. First, DNA sequence and conformation govern exonuclease activity in a complex, interdependent manner. With heterogeneous sequence DNA as the substrate, there is a slight preference for single-stranded versus double-stranded DNA. When acting upon substrates with homopolymer tails, the enzyme is more active on runs of T in the single-stranded form and runs of A in the double-stranded form than it is on runs of T in the double-stranded form and runs of A in the single-stranded form. Runs of G and C in

the double-stranded form are poor substrates. The cumulative contribution of these effects in heterogeneous sequence DNA is most likely responsible for the nonrandom distribution of transient intermediate digestion products observed during the course of the reaction. Such characteristics are not without precedents. Exonuclease III of *E. coli* has been noted to release mononucleotides from duplex DNA at variable rates depending on sequence context. It was noted in a study using mouse satellite DNA as a substrate that the summation of even rather minor differences in the rates of cleavage of different nucleotide residues led to transient accumulation of discrete fragment intermediates during the course of digestion (*21*). Analysis of the *Drosophila melanogaster* Rrp1 exonuclease on simple sequence oligonucleotide substrates revealed orders-of-magnitude variations in the rates of removal of 3′-terminal nucleotides depending on the specific nucleotide and sequence context (*22*). In previous studies on the biochemistry of the Rec1 exonuclease (*1*), it was determined that activity was several times greater when the 3′-labeled substrate was denatured heterogeneous sequence DNA rather than native DNA, and it was concluded that the enzyme preferred single-stranded DNA over double-stranded DNA. However, such a generalization might not always be appropriate or accurate. The studies presented here raise the caveat that rates of removal of individual nucleotide residues by Rec1 exonuclease vary with the context of the sequence, the identity of the base, and the conformation of the residue.

A second feature of the nuclease activity that addresses aspects of a speculative model of the in vivo function of Rec1 is the absence of any intrinsic specificity for exonucleolytic removal of mismatched base pairs. Certain biochemical features of the Rec1 exonuclease activity observed previously such as polarity of digestion, the preference for denatured over native plasmid DNA, the activity on phosphorothioate linkages, and the inhibition by deoxyribomononucleotides were reminiscent of known properties of ϵ, the proofreading exonuclease component of *E. coli* DNA polymerase III (*23*−*25*). Other observations such as the tentative alignment of stretches of the protein sequence with three signature sequence motifs present in the proofreading exonuclease domain of DNA polymerases, the periodic expression of the gene, and S phase specificity, together with the mutator phenotype of the *rec1* mutant, led to the speculation that Rec1 might function as an external proofreading component of an as yet unidentified DNA polymerase dedicated to DNA repair (*18*). While not all of the phenotypic properties of the mutant could be well rationalized by such a model, we found the similarities with ϵ intriguing enough to warrant further testing. The results from this study indicate that Rec1 has no apparent intrinsic specificity for mismatched bases pairs as is the case with ϵ. We observed that the rate of removal of a mispaired 3′-terminal residue was no faster than the rate of removal of perfectly base paired residues. If we consider what appears to be the more widespread case that the specificity of DNA polymerases in recognizing mispaired bases lies in the architecture of the protein itself which excludes duplex DNA from entering the exonuclease site (*26*), it might be argued that a function for Rec1 in proofreading cannot formally be ruled out. However, it seems clear that the inherent ability of ϵ to preferentially remove mismatched base residues is not a feature that is evident in Rec1, and thus, it is unlikely that Rec1 functions as an editing exonuclease by a mechanism similar to that exemplified by the bacterial enzyme.

Another model for Rec1 that we have considered is one in which it might play a role in mismatch repair. Mismatch repair involves action of a set of proteins that recognize mispaired residues and then removes a tract of nucleotides spanning the site of mismatch (*27*). In studies on *E. coli* mismatch repair in vitro and in vivo, molecular studies have implicated three exonucleases in mismatch repair, including ExoI, ExoVII, and RecJ (*28*−*30*), with the possibility of others that have yet to be discovered. Degradation in both 5′−3′ and 3′−5′ directions is possible. Eukaryotic exonucleases involved in mismatch repair have not been unequivocally identified, although strong candidates have been identified. In *Sc. pombe*, mutants defective in the *exo1* gene, which encodes a 5′−3′ exonuclease (*31*) that prefers double-stranded DNA, exhibit a mutator phenotype during mitotic growth and marker effects at the *ade6* locus during meiotic recombination, consistent with the loss of long-patch repair of hybrid DNA (*32*). The Exo1 orthologue in *S. cerevisiae* also prefers double-stranded DNA over single-stranded DNA (*33*) and was found by two-hybrid and co-immunoprecipitation experiments to interact with the Msh2 protein, the bacterial MutS homologue that recognizes mispaired bases (*34*). Inactivation of the *S. cerevisiae EXO1* gene leads to defects in mitotic recombination (*33*) and a mutator phenotype (*34*). It seems likely that *EXO1* plays a role in both mismatch repair and recombination. Biochemical studies of mismatch repair in eukaryotes indicates strand excision can proceed bidirectionally, implying there could be an array of exonucleases as found in prokaryotes (*35*, *36*). While the eukaryotic candidates for mismatch repair exonucleases differ from the bacterial exonucleases in exhibiting no specificity for single-stranded DNA, there does not seem to be any reason a priori to expect that the exonucleases involved should be strictly single-strand specific since specificity is inherent within the MutS−MutL (or their homologues in eukaryotes) mispair recognition complex. However, one might expect a candidate exonuclease in eukaryotes to initiate at a nick since that could be the signal that confers directional repair. Thus, the third significant finding in this work, namely, that Rec1 can initiate exonucleolytic degradation at an internal nick, is consistent with such a notion.

The fourth important observation from this study is the discovery of endonuclease activity in Rec1 directed at abasic sites. This raises the issue that Rec1 could play a direct role in DNA repair through recognition of damaged residues. This possibility had been investigated before in connection with the extreme UV sensitivity of the *rec1* mutant. One explanation for this was that the radiation sensitivity might be due to an inability to excise pyrimidine dimers. However, when it was found that the rate of removal of pyrimidine dimers in *rec1* was no different from that of the wild type (*37*), it was concluded that Rec1 most likely did not play a direct role in DNA repair, and other mechanisms had to be considered. The discovery of the cell cycle checkpoint defect in the *rec1* mutant provided a framework for rationalizing the radiation sensitivity, at least in part (*3*). Survival of *rec1* cells irradiated with UV or ionizing radiation was greatly enhanced if mitosis was delayed for several hours. This and

other observations indicated that cell death was due in part to a failure of the cells to halt in the cell cycle before proceeding through mitosis. At high UV doses, the recovery of viability achieved through artificial checking of cells in G2 with addition of a microtubule inhibitor, while significant, never reached a level of more than 10% compared to that of the wild type. This was puzzling in light of the substantially higher fraction of survival noted in a similar experiment performed with the *rad1-1* checkpoint mutant of *Sc. pombe* (*38*), which is defective in a gene related to *REC1* (see below). One explanation for this could be that besides its function in cell cycle checkpoint control, Rec1 does indeed have a direct role in DNA repair. However, it is still not apparent how the loss of an abasic endonuclease activity might result in a large increase in spontaneous mutator activity given that AP endonuclease mutants of *S. cerevisiae* are not mutators (*39*−*41*).

We note that as yet there is no accumulated biochemical data to support the sliding clamp mechanism for Rec1 proposed by Thelen et al. (*10*). As shown in the study presented here, the degradation of oligonucleotides is distributive in manner, contrary to what one would expect for an exonuclease with DNA clamp function, and as shown in a previous study (*1*). Rec1 in solution does not appear to self-associate. Of course, it is possible that the oligonucleotides used in this study are too short to enable a processive mode of degradation, or that the use of a recombinant protein that is tagged and renatured might not reflect the action of the native protein, or that a clamp loader or different solvent conditions might be necessary for assembly of a Rec1 clamp on DNA. More studies are needed to illuminate this hypothesis.

Homologues of Rec1 have been identified in *S. cerevisiae* (*5*) and *Sc. pombe* (*4*) as products of the *RAD17* and *rad1*$^+$ genes, respectively. Inactivation of these genes results in cell cycle checkpoint defects, radiation sensitivity, and abnormal recombination. The recent identification of a human Rec1 homologue (*6*−*9*) and mapping of the structural gene to a region associated with the loss of heterozygosity in several human cancers, notably lung cancer, provide a compelling reason to understand the biological role of Rec1. Given the similarity in the biochemical properties of the exonuclease activities associated with Rec1 and p53, the guardian of the genome (*42*), it seems obvious that continued investigation of the biochemical properties as well as analysis of the genetic and molecular properties of the *rec1* mutant will lead to insight into the cellular function.

REFERENCES

1. Thelen, M. P., Onel, K., and Holloman, W. K. (1994) *J. Biol. Chem. 269*, 747−754.
2. Holliday, R., Halliwell, R. E., Evans, M. W., and Rowell, V. (1976) *Genet. Res. 27*, 413−453.
3. Onel, K., Koff, A., Bennett, R. L., Unrau, P., and Holloman, W. K. (1996) *Genetics 143*, 165−174.
4. Long, K. E., Sunnerhagen, P., and Subramani, S. (1994) *Gene 148*, 155−159.
5. Lydall, D., and Weinert, T. (1995) *Science 270*, 1488−1491.
6. Dean, F. B., Lian, L., and O'Donnell, M. (1998) *Genomics 54*, 424−436.
7. Freire, R., Murguia, J. R., Tarounas, M., Lowndes, N. F., Moens, P. B., and Jackson, S. P. (1998) *Genes Dev. 12*, 2560−2573.
8. Parker, A. E., Van de Weyer, I., Laus, M. C., Oostveen, I., Yon, J., Verhasselt, P., and Luyten, W. H. M. L. (1998) *J. Biol. Chem. 273*, 18332−18339.
9. Udell, C. M., Lee, S. K., and Davey, S. (1998) *Nucleic Acids Res. 26*, 3971−3978.
10. Thelen, M. P., Venclovas, C., and Fidelis, K. (1999) *Cell 96*, 769−770.
11. Alani, E., Reenan, R. A. G., and Kolodner, R. D. (1994) *Genetics 137*, 19−39.
12. Williamson, M. S., Game, J. C., and Fogel, S. (1985) *Genetics 110*, 609−646.
13. Kramer, W., Fartmann, B., and Ringbeck, E. C. (1996) *Mol. Gen. Genet. 252*, 275−283.
14. Reagan, M. S., Pittenger, C., Siede, W., and Friedberg, E. C. (1995) *J. Bacteriol. 177*, 364−371.
15. Sommers, C. H., Miller, E. J., Dujon, B., Prakash, S., and Prakash, L. (1995) *J. Biol. Chem. 270*, 4193−4196.
16. Symington, L. S. (1998) *Nucleic Acids Res. 26*, 5589−5595.
17. Tishkoff, D. X., Filosi, N., Gaida, G. M., and Kolodner, R. D. (1997) *Cell 88*, 253−263.
18. Onel, K., Thelen, M. P., Ferguson, D. O., Bennett, R. L., and Holloman, W. K. (1995) *Mol. Cell. Biol. 15*, 5329−5338.
19. Patel, D. J., Kozlowski, S., Ikuta, S., and Itakura, K. (1984) *Fed. Proc. 43*, 2663−2670.
20. Bhattacharyya, A., and Lilley, D. M. J. (1989) *J. Mol. Biol. 209*, 583−597.
21. Linxweiler, W., and Hörz, W. (1982) *Nucleic Acids Res. 10*, 4845−4859.
22. Sander, M., and Benhaim, D. (1996) *Nucleic Acids Res. 24*, 3926−3933.
23. Scheuermann, R. H., and Echols, H. (1984) *Proc. Natl. Acad. Sci. U.S.A. 81*, 7747−7751.
24. Brenowitz, S., Kwack, S., Goodman, M. F., O'Donnell, M., and Echols, H. (1991) *J. Biol. Chem. 266*, 7888−7892.
25. Griep, M. A., Reems, J. A., Franden, M. A., and McHenry, C. S. (1990) *Biochemistry 29*, 9006−9014.
26. Derbyshire, V., Freemont, P. S., Sanderson, M. R., Beese, L., Friedman, J. M., Joyce, C. M., and Steitz, T. A. (1988) *Science 240*, 199−201.
27. Modrich, P., and Lahue, R. (1996) *Annu. Rev. Biochem. 65*, 101−133.
28. Lahue, R. S., Au, K. G., and Modrich, P. (1989) *Science 245*, 160−164.
29. Harris, R. S., Ross, K. J., Lombardo, M. J., and Rosenberg, S. M. (1998) *J. Bacteriol. 180*, 989−993.
30. Viswanathan, M., and Lovett, S. T. (1998) *Genetics 149*, 7−16.
31. Szankasi, P., and Smith, G. R. (1992) *J. Biol. Chem. 267*, 3014−3023.
32. Szankasi, P., and Smith, G. R. (1995) *Science 267*, 1166−1169.
33. Fiorentini, P., Huang, K. N., Tishkoff, D. X., Kolodner, R. D., and Symington, L. S. (1997) *Mol. Cell. Biol. 17*, 2764−2773.
34. Tishkoff, D. X., Boerger, A. L., Bertrand, P., Filosi, N., Gaida, G. M., Kane, M. F., and Kolodner, R. D. (1997) *Proc. Natl. Acad. Sci. U.S.A. 94*, 7487−7492.
35. Holmes, J., Clark, S., and Modrich, P. (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 5837−5841.
36. Thomas, D. C., Roberts, J. D., and Kunkel, T. A. (1991) *J. Biol. Chem. 266*, 3744−3751.
37. Unrau, P. (1975) *Mutat. Res. 29*, 53−65.
38. Rowley, R., Subramani, S., and Young, P. G. (1992) *EMBO J. 11*, 1335−1342.
39. Ramotar, D., Popoff, S. C., Gralla, E. B., and Demple, B. (1991) *Genes Dev. 12*, 3137−3143.
40. Johnson, R. E., Torres Ramos, A. A., Izumi, T., Mitra, S., Prakash, S., and Prakash, L. (1998) *Genes Dev. 12*, 3137−3143.
41. Bennett, R. A. O. (1999) *Mol. Cell. Biol. 19*, 1800−1809.
42. Mummenbrauer, T., Janus, F., Müller, B., Wiesmüller, L., Deppart, W., and Grosse, F. (1996) *Cell 85*, 1089−1099.

BI991495B