A New Method for Determining the Stereochemistry of DNA Cleavage Reactions: Application to the *Sfi*I and *Hpa*II Restriction Endonucleases and to the MuA Transposase†

Kiyoshi Mizuuchi,*,‡ Timothy J. Nobbs,§ Stephen E. Halford,§ Kenji Adzuma,‡,⊥ and Jun Qin||,⊥

Laboratory of Molecular Biology, NIDDK, National Institutes of Health, Bethesda, Maryland 20892, Department of Biochemistry, School of Medical Sciences, University Walk, University of Bristol, Bristol BS8 1TD, U.K., and Laboratory of Biophysical Chemistry, NHLBI, National Institutes of Health, Bethesda, Maryland 20892

Received January 8, 1999; Revised Manuscript Received February 12, 1999

ABSTRACT: A new method was developed for tracking the stereochemical path of enzymatic cleavage of DNA. DNA with a phosphorothioate of known chirality at the scissile bond is cleaved by the enzyme in $H_2{}^{18}O$. The cleavage produces a DNA molecule with the 5′-[^{16}O,^{18}O,S]-thiophosphoryl group, whose chirality depends on whether the cleavage reaction proceeds by a single-step hydrolysis mechanism or by a two-step mechanism involving a protein−DNA covalent intermediate. To determine this chirality, the cleaved DNA is joined to an oligonucleotide by DNA ligase. Given the strict stereochemistry of the DNA ligase reaction, determined here, the original chirality of the phosphorothioate dictates whether the ^{18}O is retained or lost in the ligation product, which can be determined by mass spectrometry. This method has advantages over previous methods in that it is not restricted to particular DNA sequences, requires substantially less material, and avoids purification of the products at intermediate stages in the procedure. The method was validated by confirming that DNA cleavage by the *Eco*RI restriction endonuclease causes inversion of configuration at the scissile phosphate. It was then applied to the reactions of the *Sfi*I and *Hpa*II endonucleases and the MuA transposase. In all three cases, DNA cleavage proceeded with inversion of configuration, indicating direct hydrolysis of the phosphodiester bond by water as opposed to a reaction involving a covalent enzyme−DNA intermediate.

Determination of the stereochemistry of enzymatic phosphoryl transfer reactions has long been recognized as a powerful tool for elucidating the mechanisms of these enzymes (*1*, *2*). Because many phosphoryl transfer reactions of biological significance occur at phosphorus atoms with two nonbridging oxygens, chemical analogues with sulfur substituting for one of these oxygens have been widely used to generate the chiral center needed for stereochemical analysis (*3*−*5*). For reactions at DNA phosphodiesters in which the recipient of the transfer step is another polynucleotide, such as the strand transfer step of the Mu transposition and HIV integration reactions (*6*, *7*) and V(D)J recombination (*8*), a convenient method already exists for the determination of the stereochemical path (*3*). In these cases, the stereoconfiguration of the phosphodiester in the product can be assessed by using two nucleases that have complementary specificities for the diastereomers of phosphorothioates; snake venom phosphodiesterase cleaves only R_p phosphorothioate linkages but not S_p linkages, while nuclease P1 cleaves S_p but not R_p phosphorothioates (*9*, *10*). However, for DNA endonuclease reactions in which the phosphate moiety of interest gains an additional nonbridging oxygen atom from a water molecule, determination of the stereochemistry of the reaction products is often a difficult task (viz. *11*).

The previous method for determining the stereochemistry of endonuclease reactions not only requires a DNA substrate with a phosphorothioate linkage of defined chirality at the scissile bond, but also requires an adenosine residue to be located 5′ to the scissile bond (*12*). This DNA is cleaved first by the nuclease of interest in the presence of $H_2{}^{18}O$ and the reaction products then digested to mononucleotides. The mononucleotide carrying the target phosphorothioate, [^{16}O, ^{18}O, S]dAMP, is separated from the others by HPLC and reacted with adenylate kinase to yield [O^{18}]dATPαS. Its absolute configuration is then determined by ^{31}P NMR. This method is inapplicable to many enzyme reactions for technical reasons, such as the limitations on the substrate DNA sequence that can be used, the necessity for virtually 100% yields at each stage in the procedure, and the high demand for the purity and quantity of material. We have developed a new method that eliminates many of these technical difficulties and used it to study the mechanism of action of several restriction endonucleases and phage Mu transposase.

† Funded in part by the National Institutes of Health intramural AIDS targeted antiviral program (USA) and by the Wellcome Trust (U.K.).

* Corresponding author. Mail: Laboratory of Molecular Biology, National Institute of Diabetes and Digestive and Kidney Diseases, National Institutes of Health, Bldg. 5, Rm. 237, Bethesda, MD 20892. Phone: (301)496-8154. Fax: (301)496-0201. E-mail: kmizu@helix.nih.gov.

‡ Laboratory of Molecular Biology, NIH.

§ University of Bristol.

|| Laboratory of Biophysical Chemistry, NIH.

⊥ Present addresses: J.Q., Department of Biochemistry and Department of Cell Biology, Baylor College of Medicine, Houston, TX 77030; K.A., Rockefeller University, Box 256, 1230 York Ave., New York, NY 10021.

10.1021/bi990054p CCC: $18.00
Published on Web 03/20/1999

In this method, we prepare a substrate DNA for each nuclease to be studied with a phosphorothioate of known chirality at the scissile bond. The substrate is digested by the nuclease in the presence of $H_2{}^{18}O$, and the product is ligated to an appropriate oligonucleotide that contains a biotinylated thymidine analogue at the 3′-end. In the reaction pathway for DNA ligase (*13*), one specific oxygen atom is lost from the 5′-phosphate. The stereochemical selectivity of DNA ligase with a sulfur-containing 5′-end was determined in this study. From this, we can deduce the stereospecificity of the endonuclease reaction products by determining whether the ^{18}O that is incorporated during hydrolysis is retained or lost during ligation. We digest the ligation products with nuclease P1 that generates the biotinylated phosphorothioate dinucleotide of interest, while all other parts of the DNA substrates are converted to mononucleotides. We then purify the dinucleotide using matrix-bound avidin beads and measure its mass by MALDI-TOF.[1] We used *Eco*RI as a test for this method, because the stereochemistry of its reaction had been established previously by the NMR method (*12*). Our result agrees with the previous result, confirming the inversion of chirality during hydrolysis. The method was then applied to DNA cleavage reactions by the *Sfi*I and *Hpa*II restriction enzymes and by the transposase from phage Mu, the MuA protein.

MATERIALS AND METHODS

Isotopes. $H_2{}^{18}O$ (>97%) was obtained from Isotec. [γ-^{32}P]-ATP was from NEN. [α-^{32}P]ddATP and [γ-^{35}S]ATP were from Amersham Pharmacia Biotech.

Proteins. MuA transposase (*14*) and restriction endonucleases *Eco*RI (*15*) and *Sfi*I (*16*) were purified as before. Other enzymes were obtained from: endonuclease *Hpa*II, T4 ligase, polynucleotide kinase (New England Biolabs); P1 nuclease, terminal transferase (Amersham Pharmacia Biotech); calf intestine alkaline phosphatase and snake venom phosphodiesterase (Boehringer Mannheim). Streptavidin-coated magnetic beads (Dynabeads M-280 Streptavidin) were from Dynal.

Oligonucleotides. All of the oligonucleotides used in this study were synthesized by the HHMI Biopolymer/Keck Foundation Biotechnology Resource Laboratory at Yale University. Oligonucleotides that contained biotinylated T (Glen Research, cat. no. 10-1038) at the 3′-ends were synthesized in one of two ways; the synthesis was started either with a 3′-phosphorylation column with the first nucleotide addition of biotinylated T followed by the remainder of the sequence, or with any of the standard nucleotide columns with first addition of the chemical phosphorylation reagent (Glen Research, cat. no. 10-1900) followed by the biotinylated T and the rest of the sequence. Either method produces oligonucleotides with the biotinylated base at the intended 3′-end position adjacent to a 3′-phosphoryl group, which was subsequently removed by alkaline phosphatase treatment. Phosphorothioate-containing duplexes were assembled from three oligonucleotides. One oligonucleotide was synthesized as a trityl-on derivative with a racemic phosphorothioate linkage instead of a phosphodiester between the 5′ and the penultimate nucleotides. HPLC purification of this derivative generally gave good separation of the two diastereoisomers, and their separation appears to be uninfluenced by the chain length as first noted by Stec et al. (*17*). Approximately 1 μmol of oligonucleotides was loaded onto a C-4 column (21.4 × 250 mm plus a guard column, Dynamax-300A, Rainin) at a flow rate of 7.5 mL/min of 100% solvent A (0.1 M triethylammonium acetate, pH 7), 0% solvent B (acetonitrile). One of the two buffer gradients (80% A + 20% B to 75% A + 25% B in 40 min at 7.5 mL/min or 75% A + 25% B to 70% A + 30% B in 40 min at 7.5 mL/min) was chosen for each sample by a test run of a small quantity of the sample. The retention time differed depending on the sequence and the gradient used. Typical separation between the two diastereomers was between 5 and 8 min. To determine the stereoconfiguration, a fraction of the sample from each HPLC peak was first detritylated and then labeled at its 5′ end with [γ-^{32}P]ATP and T4 kinase, prior to digestion with nuclease P1 or snake venom phosphodiesterase: digestion products were analyzed by electrophoresis through polyacrylamide (20% or 24% sequencing gel) and autoradiography. The early-eluting peak from the HPLC always contained the S_p phosphorothioate linkage, and the late-eluting peak, the R_p form. The oligonucleotide containing a phosphorothioate of known stereoconfiguration adjacent to its 5′ end was phosphorylated at its 5′ end, with unlabeled ATP and T4 kinase, and labeled at its 3′ end, with [α-^{32}P]ddATP and terminal transferase. This was subsequently ligated by T4 DNA ligase to the oligonucleotide that formed the other half of one strand of the endonuclease substrate, in the presence of a third (complimentary) oligonucleotide, the latter forming the second strand of the endonuclease substrate. The substrates for restriction endonuclease reactions were used after ethanol precipitation. The substrate for the MuA reaction was further purified through nondenaturing polyacrylamide gel electrophoresis before use. The sequences of the oligonucleotides used are as follows ("s" indicates the positions of phosphorothioate linkages, and "T^{bio}" represents biotinylated T analogue). Components of the *Sfi*I substrate: MM757 TsTGGCCGAAGAGATGCG, MM743 GAGTGGCCGTT, MM744 TTCGGCCAAAACGGCCACTC. *Sfi*I ligation partner: MM867 CGCATCTCTTCGGCCAAGAAGAGGGC-CCTTTTGGGCCCTCTTCT^{bio}. Components of the *Eco*RI substrate: MM758 GsAATTCGAAGAGATGCG, MM789 TGCCCG, MM791 TTCGAATTCCGGGCA. *Eco*RI ligation partner: MM865 CGCATCTCTTCGAATTAGAAGAGGGC-CCTTTTGGGCCCTCTTCT^{bio}. Components of the *Hpa*II substrate: MM759 CsCGGCAGAAGAGATGCG, MM749 AAAAGG, MM750 TTCTGCCGGCCTTTT. *Hpa*II ligation partner: MM866 CGCATCTCTTCTGCCGAGAAGAGGGC-CCTTTTGGGCCCTCTTCT^{bio}. Components of the MuA substrate: MM752 AAGTTTTCGCATTTATCGTGAAACG-CTTTCGCGTTTTTCGTGCGCCGCTTC, MM753 CCGAT-GAAGCGGCGCACGAAAAACGCGAAAGCGTTT-CACGATAAATGCGAAAACTT, MM760 AsTCGGTGAA-GAGATGCG, MM846 TCTCTTCA. (MM846 was added to the ligation reaction for the MuA substrate assembly without 5′-phosphorylation. Thus it was not incorporated into the substrate.) MuA ligation partner: MM832a CGCATCTCT-TCACCGAAGAAGAGGGCCCTTTTGGGCCCT-CTTCT^{bio}.

[1] MALDI-TOF, matrix assisted laser desorption ionization, time of flight; MS, mass spectrometry.

Reaction Conditions. The *Sfi*I reaction contained approximately 150 pmol of the substrate and 37.5 pmol of *Sfi*I in 200 μL of buffer containing 10 mM Tris-HCl (pH 7.9), 10 mM $MgCl_2$, 50 mM NaCl, 1 mM dithiothreitol, and 100 μg/mL BSA. The reactions contained either $H_2{}^{16}O$ or $H_2{}^{18}O$ (isotopic purity in the reaction ~95%). The reactions were overlaid with a drop of mineral oil and incubated at 50 °C for 19 h. After stopping the reaction with 20 mM EDTA, samples were extracted with phenol and chloroform and DNA was precipitated with ethanol. The *Eco*RI reaction contained 150 pmol of the substrate and 6000 units of *Eco*RI in 200 μL of buffer containing 100 mM Tris-HCl (pH 7.5), 10 mM $MgCl_2$, 50 mM NaCl, and 0.025% Triton X-100. The isotopic purity was ~96% for the reaction in $H_2{}^{18}O$. The reactions were overlaid with a drop of mineral oil and incubated at 37 °C for 19 h. Reactions were stopped and DNA precipitated as above. The *Hpa*II reaction contained 150 pmol of the substrate and 200 units of *Hpa*II in 200 μL of buffer containing 10 mM bis-Tris-propane-HCl (pH 7.0), 10 mM $MgCl_2$, and 1 mM dithiothreitol. The isotopic purity was ~87% for the reaction in $H_2{}^{18}O$. After incubation at 37 °C for 19 h, reactions were stopped and DNA precipitated as described above. DNA samples were annealed with approximately 300 pmol of appropriate ligation partners, and ligation reactions were carried out in 50 μL of buffer containing 50 mM Tris-HCl, 10 mM MgCl, 10 mM dithiothreitol, 1 mM ATP, 25 μg/mL BSA, and 400 units of T4 DNA ligase at 25 °C overnight. After ethanol precipitation and drying, half of the samples were digested with 10 μg/mL of nuclease P1 in 50 μL of buffer containing 20 mM sodium acetate (pH 5.2), 0.1 mM zinc acetate, and 10% glycerol at 56 °C for 60 min.

The avidin beads (10 mg/mL suspension) for isolation of the biotinylated mono- and dinucleotides were washed three times with 2 M ammonium acetate and resuspended in the original volume of the washing buffer before use. The nuclease P1 digested DNA samples dissolved in 100 μL of water were mixed with an equal volume of the avidin beads suspension at room temperature for 30 min with occasional gentle shaking. The beads were washed three times with 0.5 mL of 1 M ammonium acetate and then three times with 0.5 mL water. To liberate the nucleotides from the beads, the beads were washed with two 100 μL aliquots of concentrated ammonia, each for 45 min at room temperature, and the aliquots combined. The eluted nucleotide samples were dried by lyophilization. A small fraction of the sample at each stage was used to monitor the reaction by denaturing polyacrylamide gel electrophoresis and autoradiography.

MuA Reaction. The Mu end cleavage reaction by MuA contained approximately 240 pmol of the Mu end substrate and 600 pmol of MuA in 500 μL of buffer containing 25 mM Tris-HCl (pH 8.0), 10 mM MgCl, 130 mM NaCl, 15% glycerol, 15% DMSO, and 0.05% Triton X-100. The isotopic purity was ~80% for the reaction in $H_2{}^{18}O$. The reaction was incubated at 30 °C for 7.5 h and stopped by the addition of 0.2% SDS. The extent of the cleavage reaction was between 10 and 15%. Samples were boiled for 2 min, and DNA was ethanol precipitated and resuspended in 50 μL of 10 mM Tris-HCl (pH 7.6) 0.5 mM EDTA. One-half of the DNA sample was annealed with 200 pmol of the ligation partner, and the ligation reaction was carried out as described for the restriction endonuclease samples. DNA was ethanol precipitated, dried, digested by nuclease P1, and biotinylated nucleotides were isolated as described above.

Mass Spectrometry. A 1.2-meter MALDI-TOF mass spectrometer with delayed extraction (Voyager-DE, Perseptive Biosystems, Framingham, MA) was used. Data were collected with an external 2-GHz digital oscilloscope (Tektronix, Houston, TX). The working matrix solution was a 2-fold dilution of a saturated solution of 2,5-dihydroxybenzoic acid in acetonitrile/water (1:1). The lyophilized nucleotide samples were dissolved in 50% acetonitrile in water. Aliquots of 0.5 μL of nucleotide sample, 0.5 μL of the working matrix solution, and 0.5 μL of 50 mM NH_4HCO_3 were mixed on the sample plate and dried in air prior to MS analysis. The addition of 50 mM NH_4HCO_3 improved the quality of the mass spectrum. Spectra were taken in the negative mode; therefore, the measured peaks corresponded to m-H. The mass spectrometer was carefully tuned so that a unit resolution could be maintained. A high-resolution MALDI-TOF mass spectrometer with delayed extraction and a reflector (Voyager-STR, Perseptive Biosystems, Framingham, MA) was also used, but no spectrum could be obtained despite that many different matrixes were tested.

RESULTS AND DISCUSSION

For a large number of enzymatic phosphoryl transfer reactions, the stereochemical course of the reaction has been analyzed by using substrates containing either a chiral phosphorothioate or a chiral phosphate with ^{17}O and ^{18}O isotopes (*1*, *3*). All of the evidence currently available indicates that each transfer reaction at a phosphorus center proceeds with inversion of stereoconfiguration. Knowledge about the stereochemical course of a phosphoryl transfer reaction is thus a powerful tool in assessing whether the enzymatic reaction involves a covalent intermediate, linked to either the enzyme or a cofactor. A single-step substitution reaction without a covalent intermediate should lead to inversion of configuration, while a reaction involving a covalent intermediate should proceed with two inversions and thus result in overall retention of configuration.

We were particularly interested in the mechanism of DNA cleavage by two dissimilar endonucleases: a restriction endonuclease, *Sfi*I, and a transposition endonuclease, the MuA protein. *Sfi*I differs from orthodox restriction enzymes in that, instead of being a dimeric protein that attacks an individual recognition site (*18*), it exists as a tetramer that has to bind to two copies of its recognition site simultaneously before being able to cut the DNA (*19*). The demand for the pairing of two recognition sites, followed by the excision of the intervening loop of DNA (*20*), is reminiscent of site-specific DNA recombination reactions. Thus, even though the biological function of *Sfi*I appears to be in DNA restriction rather than rearrangements (*21*), it seems possible that this endonuclease might, at some time during its evolutionary history, have been involved in recombination reactions between pairs of recognition sites. If this were the case, then its reaction pathway may include a covalent DNA−protein intermediate, as in the strand exchange reactions of site-specific recombinases such as λ integrase (*22*, *6*). Other restriction enzymes, such as *Nae*I and *Hpa*II, are also activated by additional copies of their recognition sequences (*23*), though the way these enzymes utilize two

recognition sites differs from *Sfi*I (24). However, DNA cleavage by *Nae*I has been reported to involve a covalent protein-DNA intermediate (25).

In the case of MuA, a member of the transposase-retroviral integrase family of proteins, two dissimilar chemical reactions are catalyzed by the same active site of the protein (*26*, *27*). The first is the endonucleolytic cleavage of the 3′-ends of the transposon sequence, and the second is joining of the newly liberated 3′-OH groups to a target DNA, in a reaction termed DNA strand transfer. Previously, it has been shown that DNA strand transfer by MuA takes place by a single-step transesterification without the involvement of a covalent intermediate (*6*). However, the stereochemical course of the 3′ cutting has not been directly determined. For a closely related HIV1 integrase reaction, however, a side reaction was found in which a nearby 3′-OH, instead of a water molecule, participated in the 3′ cutting reaction. This reaction was shown to proceed with inversion of stereoconfiguration at the scissile phosphate, again indicating a single-step mechanism (*7*). While this result suggested that the normal 3′ cutting reaction also proceeds via a single-step mechanism, other interpretations were also possible. Therefore, it was of interest to directly examine the stereochemical course of the hydrolysis reaction at the 3′-ends of the transposon.

For the reactions of *Sfi*I (*24*) and MuA (*6*), it is difficult to generate sufficiently large quantities of the reaction products, in sufficiently pure form, for the determination of product stereochemistries by the NMR method. We therefore developed a new method based on the stereoselectivity and the known mechanism of action of DNA ligase to follow the stereochemistry of endonuclease reactions. In this method, the endonuclease of interest is used in $H_2{}^{18}O$ to generate a reaction product carrying a 5′-thiophosphoryl end and this is then joined by DNA ligase to an appropriate oligonucleotide. In the ligation process, we show below that of the two nonesterified oxygens, which are stereochemically distinct in the thiophosphate (*pro*S and *pro*R oxygens), one (*pro*S oxygen) is stereospecifically removed from DNA. We determine by MS whether or not the ^{18}O atom incorporated during DNA cleavage is retained in the ligation product. This in turn will tell us at which of the two possible positions in the original 5′-thiophosphoryl moiety the ^{18}O atom was located. Therefore, we can deduce the stereochemical path of the endonuclease reaction.

Reactivity and Product Stereospecificity of T4 DNA Ligase with 5′-Thiophosphoryl DNA Termini. For the experimental approach outlined above to work, DNA ligase must be able to join a DNA end carrying a 5′-thiophosphoryl group to a ligation partner with a 3′-OH group. To test whether DNA ligase can join thiophosphoryl ends, an ^{35}S-labeled thiophosphoryl group was added to the 5′-end of a *Hin*dIII restriction fragment by using [^{35}S]ATPγS and T4 polynucleotide kinase prior to ligation with T4 DNA ligase. Ligase acted readily on this substrate. To determine the chirality of the resulting phosphorothioate linkage, the ligation product was digested with nuclease P1, which cuts the phosphorothioate linkages in their S_p but not in their R_p configuration. Nuclease P1 yielded an ^{35}S-labeled dinucleotide as its limit product (Figure 1, lanes i−k). Further digestion of this dinucleotide by snake venom phosphodiesterase, which can cut the R_p form of phosphorothioate linkages but not the S_p form, converted the dinucleotide to mononucleotides (Figure 1, lanes l, m). We therefore conclude that T4 DNA ligase can stereospecifically ligate 5′-thiophosphoryl termini to produce exclusively R_p phosphorothioate linkages. The phosphorothioate linkages formed by T4 ligase have the same properties as those generated by DNA polymerase with dNTPαS, which also yields the R_p diastereomer (*28*; also see Figure 1, lanes d−h). The strict stereoselectivity of T4 ligase was maintained when the ligation reaction was carried out in the presence of Mn^{2+} instead of Mg^{2+} (data not shown).

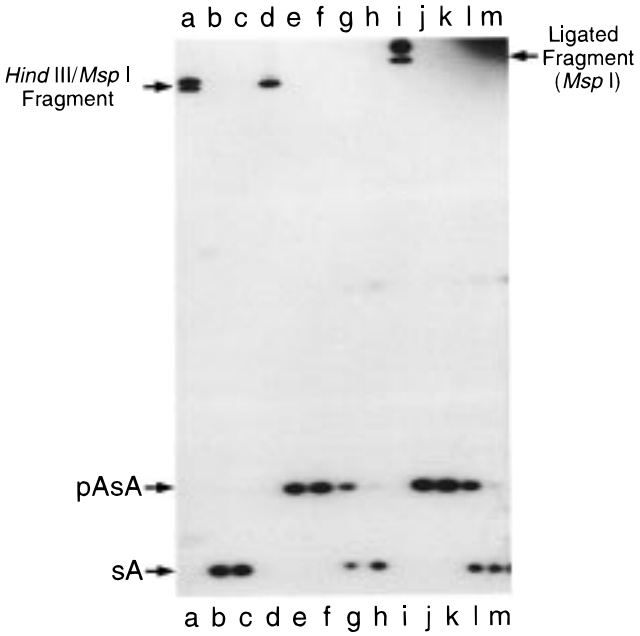

FIGURE 1: T4 DNA ligase joins a 5′-thiophosphoryl DNA end to generate an R_p phosphorothioate linkage. Autoradiogram of a 20% sequencing gel is shown. The pUC19 DNA was linearized by *Hin*dIII, dephosphorylated, and then thiophosphorylated using [γ-^{35}S]ATP and T4 polynucleotide kinase. Digestion by *Msp*I produced a 5′-thiophosphoryl 77-bp *Hin*dIII−*Msp*I fragment (lane a). The thiophosphorylated *Hin*dIII fragment was self-ligated using T4 DNA ligase in the presence of ATP, digested with *Msp*I, and run on a sequencing gel, from which a 111-bp fragment corresponding to the "head-to-tail" ligation product was purified (lane i). Radiolabeled phosphorothioate linkages were also made by extending the 3′-ends of the *Hin*dIII-linearized pUC19 DNA by Sequenase (Amersham Pharmacia Biotech) in the presence of [α-^{35}S]dATP and unlabeled dG,C,TTP. This DNA was also digested with *Msp*I and the 77(+3∼4)-bp fragment was purified (lane d). Digestion of the 5′-thiophosphoryl fragment (lane a) with nuclease P1 produced a labeled mononucleotide p[^{35}S]-dA (lanes b and c, 2 and 30 min reactions, respectively), as indicated by "sA." P1 digestion of the DNA polymerase filled-in fragment (lane d) produced a labeled dinucleotide p-dA-p[^{35}S]-dA (lanes e and f, 2 and 30 min reactions, respectively), as indicated by "pAsA", the results consistent with the fact that the polymerase produces an R_p phosphorothioate linkage exclusively. Further digestion of these nuclease P1 products with snake venom phosphodiesterase produced a labeled mononucleotide p[^{35}S]-dA (lanes g and h, 15 and 60 min reactions, respectively). The ligated fragment (lane i) was digested with nuclease P1 (lanes j, k) or P1 plus snake venom phsphodiesterase (lanes l, m) as in lanes e, f, or g, h. The DNA ligase-generated phosphorothioate linkage responded to the nucleases identically as did that made by DNA polymerase. (The top bands of the doublets of undigested fragments, lanes a, d, and i, most likely reflect double-stranded forms.)

DNA end joining by DNA ligase proceeds through a succession of phosphoryl transfer reactions (*13*, also see Figure 2). First, the enzyme is adenylated at its active site lysine by ATP, or NAD in the case of the *E. coli* enzyme.

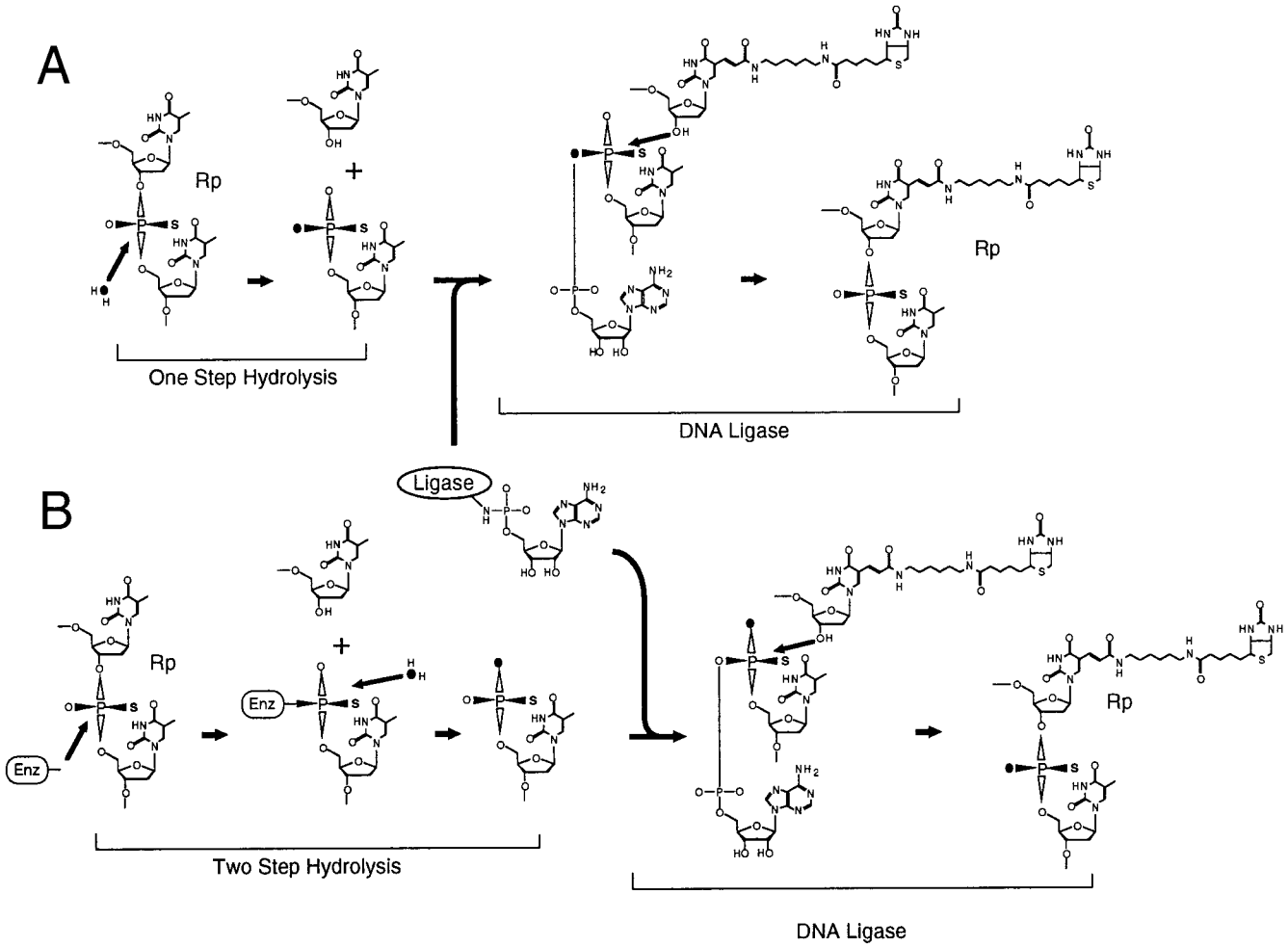

FIGURE 2: Expected stereochemical outcome of ligation of endonuclease products with a 5′-thiophosphoryl end (5′-T residue depicted here as an example). Panel A: A single-step hydrolysis of R_p phosphorothioate linkage in $H_2{}^{18}O$ incorporates the ^{18}O atom at the *pro*S position. DNA ligase first adenylates this thiophosphoryl group with the ^{18}O atom occupying the bridging position. When the 3′-OH of the ligation partner (shown here with a T^{bio} residue at the 3′-end) substitutes the adenyl group, the ^{18}O atom is released in the form of [^{18}O]AMP. Panel B: A two-step hydrolysis of R_p phosphorothioate linkage in $H_2{}^{18}O$ incorporates the ^{18}O atom at the *pro*R position. The adenylated ligation, intermediate with this thiophosphoryl end will have the ^{18}O atom at the nonbridging position. Thus, when the 3′-OH of the ligation partner substitutes the adenyl group, the ^{18}O atom will be retained within the ligated DNA. ● indicates ^{18}O.

The adenyl group is then transferred to the 5′-phosphoryl end of the DNA, forming a phosphoryl anhydride bond. In this process, one of the nonbridging oxygens in the 5′-phosphate acts as the nucleophile that attacks the adenyl-lysine bond and becomes the bridging oxygen between the two phosphorus atoms. Next, the 3′-OH of the second DNA fragment acts as the nucleophile that attacks the pyrophosphate bond and joins itself to the 5′-end of the first fragment, thus releasing AMP. In this last stage, the bridging oxygen in the pyrophosphate bond, which originates from the 5′-phosphate, is released with the AMP. With 5′-phosphoryl DNA, three nonbridging oxygens on the 5′-phosphate are equivalent and have an equal chance of becoming lost from the ligated DNA. On the other hand, when 5′-thiophosphoryl DNA carrying ^{16}O and ^{18}O atoms and a sulfur atom at the nonbridging positions are ligated, potentially either the ^{16}O or ^{18}O atom will become the bridging oxygen for the adenylated intermediate and thus will be lost from the final ligation product. Which one will be lost depends on the chirality of the 5′-thiophosphoryl group. According to the stereoselectivity of DNA ligase determined above, the ligation of a 5′-thiophosphoryl DNA results in the loss of the *pro*S oxygen, while the *pro*R oxygen is retained in the ligation product.

Figure 2 explains the positioning of the water-derived oxygen depending on the hydrolysis mechanism and its fate during the ligation reaction. If the R_p diastereoisomer of a phosphorothioate linkage is cut by an enzyme in $H_2{}^{18}O$ and the reaction proceeds by a one-step in-line mechanism, an ^{18}O atom will be incorporated at the *pro*S position as shown in panel A. On the other hand, an ^{18}O atom will be incorporated at the *pro*R position if the reaction proceeds through two in-line substitution steps, as would happen if the 5′-thiophosphoryl end is transferred first to an acceptor in the enzyme and only later to water. Thus, with an initial substrate for the nuclease of interest carrying an R_p phosphorothioate at the scissile bond, loss of ^{18}O upon ligation of the digestion product to the partner indicates a one-step mechanism, while retention of ^{18}O indicates a two-step scheme. Conversely, a substrate with an S_p phosphorothioate at the scissile bond yields the opposite result: retention of ^{18}O for a one-step hydrolysis and loss for a two-step reaction.

*Testing the Method with Eco*RI. We first tested the method with *Eco*RI that had previously been shown by the NMR

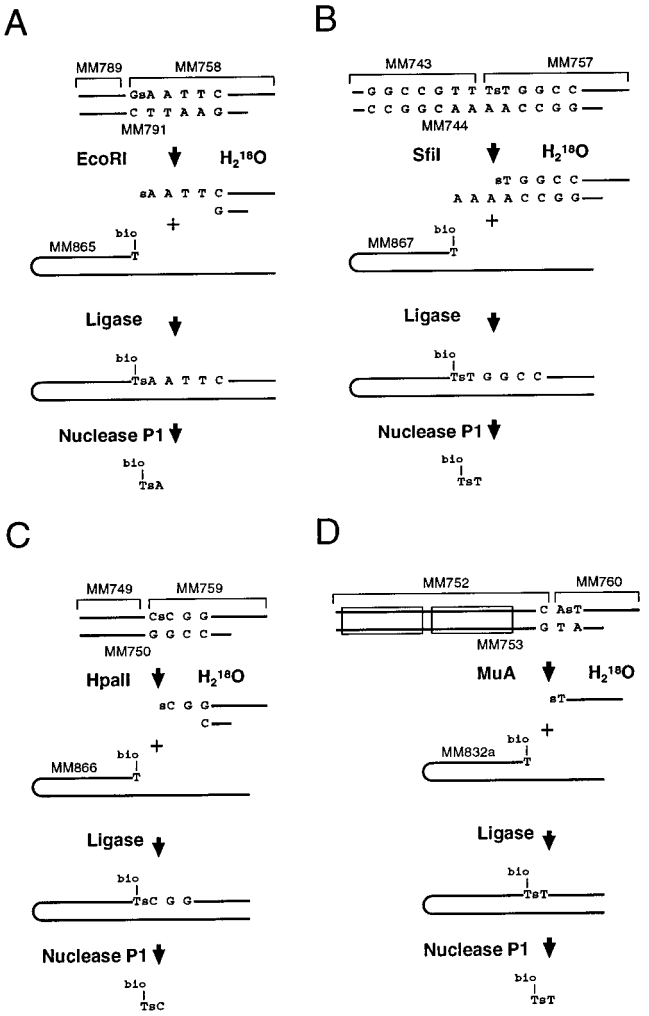

FIGURE 3: Endonuclease substrates and ligation partners used. The substrates of the endonuclease reactions were assembled from three oligonucleotides (see Materials and Methods). Nuclease digestion of the substrates in $H_2{}^{18}O$ produced a thiophosphorylated 5′-end with concomitant incorporation of the ^{18}O atom. The product strand that terminated with thiophosphoryl 5′-end was ligated with a partner DNA carrying a T^{bio} residue at the 3′-end of a hairpin segment adjacent to a single-stranded 5′-tail complementary to the strand to be ligated. Digestion of the ligation products with nuclease P1 generated biotinylated phosphorothioate containing dinucleotide. Mass analysis of this dinucleotide revealed retention or release of the ^{18}O atom from the ligation products.

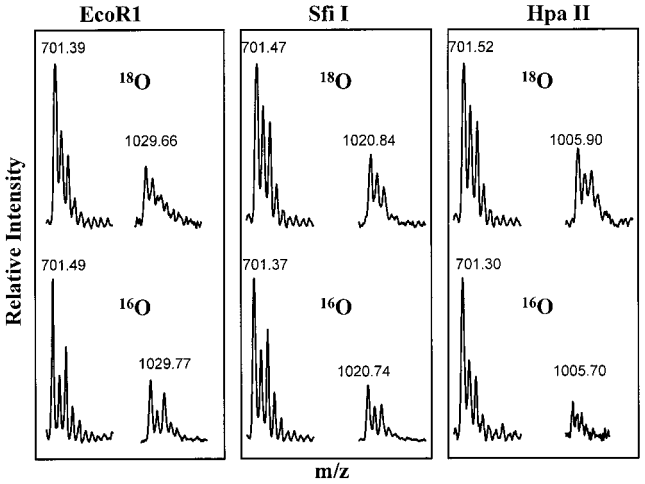

FIGURE 4: MALDI-TOF analysis of restriction endonuclease reaction products. The mass spectra of the area of interest, together with those corresponding to $T^{bio}MP$, which originates from the unligated partner fragments, are shown for the three restriction endonucleases studied. Left panel: The result for the *Eco*RI reaction in $H_2{}^{18}O$ (top) and in $H_2{}^{16}O$ (bottom). Middle panel: The result for the *Sfi*I reaction in $H_2{}^{18}O$ (top) and in $H_2{}^{16}O$ (bottom). Right panel: The result for the *Hpa*II reaction in $H_2{}^{18}O$ (top) and in $H_2{}^{16}O$ (bottom). The *m*/*z* value for the highest peak of each ion species is indicated. The expected molecular anion mass (m-H) for $T^{bio}MP$, p−T^{bio}-ps-dA, p−T^{bio}-ps-T, and p−T^{bio}-ps-dC made up without any heavier isotopic species are 701.26, 1030.31, 1021.30, and 1006.30, respectively.

method to cut DNA with inversion of stereoconfiguration at the scissile phosphate (*12*). The substrate for *Eco*RI shown in Figure 3 contained an R_p phosphorothioate at the site of cleavage on the top strand. In agreement with earlier studies (*29*), *Eco*RI cut this substrate readily, but no reaction was detected on the equivalent substrate with an S_p phosphorothioate at the scissile bond, with either Mg^{2+} or Mn^{2+} as the metal ion cofactor (data not shown).

The DNA was digested by *Eco*RI in the presence of either $H_2{}^{18}O$ or, for the control, $H_2{}^{16}O$. After denaturing the DNA, the ligation partner was added and ligated with T4 DNA ligase to the product from the *Eco*RI reaction (Figure 3). The ligation partner had a 5′ single-stranded tail complementary to the *Eco*RI-cut strand of interest with an unphosphorylated 5′-terminus and a hairpin that terminated with a biotinylated T residue at its 3′-end. The ligation product was subsequently digested with nuclease P1 to generate a biotinylated dinucleotide containing the phosphorothioate linkage. This particular dinucleotide cannot arise from either the fraction of the substrate not cleaved by *Eco*RI or the unligated DNA, which yields biotinylated TMP. The biotinylated components of the reaction mixture were adsorbed to avidin-coated magnetic beads and, after washing the beads extensively with 1 M ammonium acetate followed by water, eluted from the beads with concentrated ammonia. The washing steps also served to free the dinucleotide from nonvolatile counterions such as Na^+ or Mg^{2+}, because these can interfere with MS analysis of oligonucleotides (*30*). After lyophilization, the samples were subjected to mass analysis by MALDI-TOF.

Figure 4, left panel, compares the mass spectra of samples obtained from *Eco*RI reactions in the presence of either $H_2{}^{18}O$ or $H_2{}^{16}O$. The dinucleotide from the reaction in $H_2{}^{18}O$ displayed essentially the same mass distribution as that from the reaction in $H_2{}^{16}O$; in both cases the predominant species, which was the lightest, matched that predicted for the expected dinucleotide containing exclusively the lightest atomic species (m/z = 1030.3). The accuracy of the measurement was estimated to be within a single mass unit, and $T^{bio}MP$ present in each sample served to confirm the measurement consistency among different samples. The ^{18}O atom incorporated during phosphorothioate hydrolysis by *Eco*RI in $H_2{}^{18}O$ must therefore have been eliminated during the ligation reaction from the *pro*S position. Otherwise, the most abundant species of the dinucleotide should have shifted up by two mass units and the peaks near 1030 and 1031 should have been missing. The observed mass heterogeneity in both samples is principally due to naturally occurring heavy isotopes. Containing 38 C atoms with 1.1% ^{13}C, 11 N atoms with 0.4% ^{15}N, 15 O atoms with 0.2% ^{18}O, and 2 S atoms with 0.8 and 4.2% of ^{33}S and ^{34}S, nearly half of the dinucleotide is expected to have at least one heavy isotope. Hence, *Eco*RI cuts DNA with inversion of stereoconfigura-

tion at the scissile phosphate, indicative of a one-step hydrolysis reaction with H_2O as the primary nucleophile. This result agrees with the earlier observations (*12*) and confirms that the method developed here yields result that are consistent with the previous NMR method.

Stereochemical Course of the Endonuclease Reactions of SfiI and HpaII. We carried out the equivalent experiments to those noted above on both the *Sfi*I and *Hpa*II restriction enzymes (see Figure 3, panels B and C for substrate configuration). In the presence of either Mg^{2+} or Mn^{2+}, both *Sfi*I and *Hpa*II cleaved DNA at their respective recognition sequences with an R_p phosphorothioate at the scissile bond, but were unable to cut the S_p diastereomer with either Mg^{2+} or Mn^{2+} as cofactor (data not shown).[2] The data obtained for *Sfi*I and *Hpa*II were essentially the same as those for *Eco*RI (Figure 4, middle and right panels). In both cases, the mass distributions for the dinucleotides obtained from reactions in $H_2{}^{18}O$ were identical to those from reactions in $H_2{}^{16}O$. Hence, the ^{18}O atom incorporated during the hydrolysis of the substrates in $H_2{}^{18}O$ was lost during the ligation stage. Phosphodiester hydrolysis by both *Sfi*I and *Hpa*II thus proceeds with inversion of configuration at the phosphate, the same outcome as both *Eco*RI and *Eco*RV (*12*, *11*). All four restriction enzymes tested to date thus appear to follow one-step mechanisms in which the scissile bond is hydrolyzed directly by water, as opposed to the formation and subsequent hydrolysis of a covalent intermediate.

The *Sfi*I restriction enzyme has to interact with two copies of its recognition sequence, on either the same or separate DNA molecules, before being able to cut DNA (*16*, *19*), while the activity of *Hpa*II on DNA with one recognition site can be increased by adding a second DNA with an *Hpa*II site (*23*). The requirement for paired recognition sites in the case of *Sfi*I, or the utilization of a second site as an activator in the case of *Hpa*II, had raised speculations that such enzymes might be more closely related to the site-specific DNA recombinases than to the archetypal type II restriction enzymes, such as *Eco*RI and *Eco*RV, that act at individual DNA sites. However, the stereochemical outcomes of the *Sfi*I and *Hpa*II reactions effectively eliminate the possibility of covalent intermediates, of the type found with λ integrase (*6*), in their reaction pathways. The enzymes that mediate site-specific recombination generally conserve the energy of the phosphodiester bond by forming a covalent intermediate prior to transfer to the polynucleotidyl acceptor (*27*). The absence of such an intermediate from the *Sfi*I reaction thus explains why this enzyme fails to mediate DNA rearrangements (*21*).

Stereochemical Course of the Endonuclease Reaction of MuA Transposase. We also analyzed, by the same method, the donor DNA cleavage reaction in phage Mu transposition. The design of the substrate is shown in Figure 3, panel D. The MuA transposase protein was able to cut this substrate with the R_p phosphorothioate at the scissile bond but not with the S_p phosphorothioate (data not shown). Following digestion of the substrate containing the R_p diastereoisomer by MuA in the presence of $H_2{}^{18}O$ and ligation of the cleaved product to the biotinylated partner, MS analysis revealed that the ^{18}O atom was not retained in the ligation product (Figure 5). Hence, this reaction also proceeds with inversion of stereoconfiguration at the scissile phosphate.

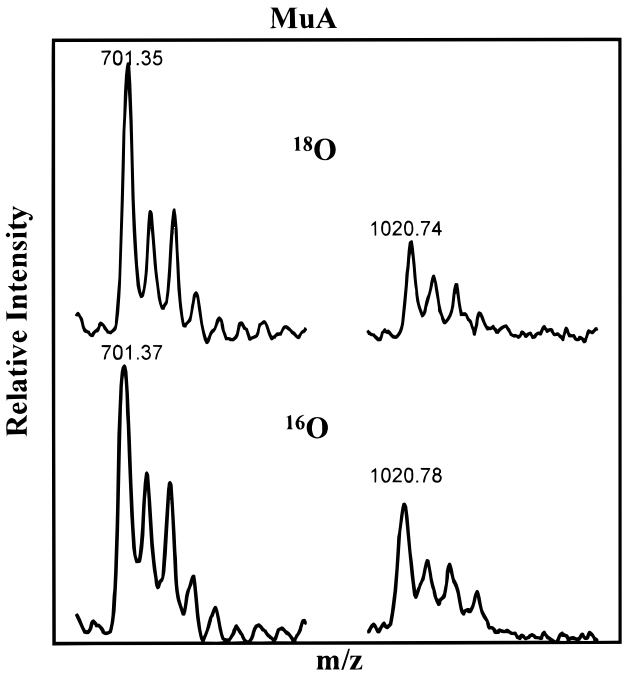

FIGURE 5: MALDI-TOF analysis of the MuA endonuclease reaction product. The result for the Mu end cleavage reaction by MuA transposase protein in $H_2{}^{18}O$ (top) and in $H_2{}^{16}O$ (bottom). The areas of mass spectra corresponding to $T^{bio}MP$ and pT^{bio}-ps-T are shown.

The initial chemical step in the transposition of Mu, DNA cleavage at the junction between transposon and host DNA, thus occurs by a one-step mechanism involving direct hydrolysis by water rather than via a covalent intermediate. This result agrees with the earlier experiment on the 3′-processing step of the HIV integrase reaction in which a nearby 3′-OH, instead of a water molecule, was used as the nucleophile for the cleavage reaction (*7*). This result also confirms the earlier assumption that the two chemical steps of the transpositional recombination of Mu are chemically equivalent reactions; each being a one-step nucleophilic attack at a phosphodiester bond in DNA (*26*, *27*). In the first step, a water molecule is the nucleophile, and in the second step, the 3′-OH that is generated in the first step becomes the nucleophile. While this notion provides a convenient basis for explaining why the same active site of the transposase and retroviral integrase can catalyze the two apparently dissimilar steps, it still remains to be understood how the change of substrate can take place within a stable protein−DNA complex in which the two steps take place successively. In the first step, the scissile phosphodiester is at the transposon 3′-end, and in the second step, it is replaced by the phosphodiester at a nonspecific DNA site that functions as the transposition target site. A question that still remains unanswered is whether the 3′-OH that functions as the nucleophile for the second step moves to the nucleophile binding pocket that had been occupied by a water molecule for the first step, or whether it remains in the leaving group pocket for the first step and the second step proceeds essentially as the reversal of the first step with respect to this moiety.

[2] In previous studies on *Sfi*I, a substrate with a *R*p phosphorothioate linkage at the scissile bond was cleaved with Mn^{2+} as the cofactor but not with Mg^{2+} (*23*). However, the recognition site for *Sfi*I, GGCCnnnnnGGCC, contains a 5 bp spacer of unspecified sequence. Changes to the spacer sequence modulate *Sfi*I activity, and the substrate used here has a different spacer from that used previously.

Conclusions and Possible Applications. The approach for the product configuration determination used here is in principle similar to that for sn-glycerol 3-[^{18}O]phosphorothioate (*31*). In both cases an isotopically chiral thiophosphoryl group is converted to a diester form with assignable stereoconfiguration, with accompanying loss of ^{16}O or ^{18}O. The method can be applied to any phosphoryl transfer reaction that produces DNA 5′-terminal phosphate, including polynucleotidyl kinase reactions, by using [γ-^{18}O]ATPγS. Given the proper design of substrate and ligation partners, this method should be applicable to the study of many sequence-specific DNA endonucleases; not only those involved in restriction and recombination, as examined here, but also many of those involved in replication or repair. It may also be possible to apply this method to nucleases that show no DNA sequence specificity, provided that partial reaction conditions can be found in which a discrete product cleaved at a particular phosphodiester bond is generated in sufficient quantity. The principal limitations of the method are that the nuclease of interest must be able to cleave a phosphorothioate linkage in DNA, in either its R_p or its S_p configuration or both, and it must leave a 5′-thiophosphoryl end for utilization by DNA ligase. For example, we had hoped to determine the stereochemical path for *Nae*I, because this restriction enzyme also requires two recognition sites for DNA cleavage and it had been reported to form a covalent intermediate (*25*). However, earlier studies on *Nae*I had revealed no cleavage of a DNA substrate containing the racemic mixture of phosphorothioates at the scissile bond (*32*). Using stereochemically pure substrates for *Nae*I, we were also unable to find reaction conditions under which this enzyme could cut DNA at either R_p or S_p phosphorothioate linkages. We were thus unable to study this reaction. None of the nucleases examined in this study could cleave both diastereomers of the phosphorothioate substrates. Thus, we could observe only the absence of ^{18}O atom in the ligation products. If we had developed a method to purify the AMP that is produced during the ligation reaction, we should have been able to detect the transfer of the ^{18}O atom to AMP. However, we have not investigated the feasibility of this analysis. Therefore, one should be wary of possible loss of the ^{18}O by an unexpected side reaction that mediates thiophosphoryl−oxygen exchange with $H_2{}^{16}O$ during the experiment, instead of the stereospecific loss to the AMP molecule during ligase reaction. However, we can rule out such a side reaction by DNA ligase. In an earlier experiment in which we used T^{bio}pTpTpT as the ligation partner for the *Sfi*I product and carried out mass analysis without P1 nuclease digestion, we performed ligase reaction in $H_2{}^{18}O$ for an extended period. Loss of ^{18}O was observed in this experiment also (data not shown). Thus, we conclude that isotope washout by the solvent water does not take place during the DNA ligase reaction.

This new method has several advantages over the previous methods in which the 5′-nucleotide generated by endonuclease cleavage of the DNA was converted by a succession of enzyme reactions to a mononucleotide triphosphate, prior to examination by ^{31}P NMR (*12*, *3*). Probably the most significant advantage is the small quantity of the material needed for the analysis. The MS analysis can be carried out with < 100 pmole of starting material while the NMR approach requires ~ 1 μmole of the sample to be analyzed by NMR. While the sensitivity of the previous method can be improved to ~ 10 nmol by using fast atom bombardment mass spectrometry for the analysis of [^{18}O](d)ATPγS, instead of ^{31}P NMR (*33*), this method is still less sensitive compared to that described here. A further advantage comes from the fact that the presence of unreacted substrate does not interfere with the MS analysis, because unreacted substrate cannot give rise to the biotinylated dinucleotide. In contrast, one consequence of the conversion of the DNA to mononucleotides in the NMR method is that the reaction of the nuclease under study should proceed to near 100% completion; otherwise, the mononucleotide can arise either from the cleaved DNA or the intact substrate, necessitating product purification steps. However, many endonucleases cleave phosphorothioate-containing substrates so inefficiently that complete cleavage is difficult to achieve (*11*, *24*). The NMR method then requires the rigorous purification of the cleaved DNA from the intact substrate. Finally, given the current interest in both ribozyme and protein reactions on RNA (*5*), it would be of interest to find out if a similar method could be devised to examine the stereochemistry of RNA reactions by using RNA ligase. This has not been investigated here.

ACKNOWLEDGMENT

We thank Bernard Connolly and Mark Szczelkun for comments on the manuscript.

REFERENCES

1. Knowles, J. R. (1980) *Annu. Rev. Biochem. 49*, 877−919.
2. Gerlt, J. A. (1993) in *Nucleases* (Linn, S. M., Lloyd, R. S., and Roberts, R. J., Eds.) pp 1−34, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
3. Eckstein, F. (1985) *Annu. Rev. Biochem. 54*, 367−402.
4. Frey, P. A. (1989) *Adv. Enzymol. Relat. Areas Mol. Biol. 62*, 119−201.
5. Verma, S., and Eckstein, F. (1998) *Annu. Rev. Biochem. 67*, 99−134.
6. Mizuuchi, K., and Adzuma, K. (1991) *Cell 66*, 129−140.
7. Engelman, A., Mizuuchi, K., and Craigie, R. (1991) *Cell 67*, 1211−1221.
8. van Gent, D. C., Mizuuchi, K., and Gellert, M. (1996) *Science 271*, 1592−1594.
9. Burgers, P. M. J., and Eckstein, F. (1979) *Biochemistry 18*, 592−596.
10. Potter, B. V. L., Connolly, B. A., and Eckstein, F. (1983) *Biochemistry 22*, 1369−1377.
11. Grasby, J. A., and Connolly, B. A. (1992) *Biochemistry 31*, 7855−7861.
12. Connolly, B. A., Eckstein, F., and Pingoud, A. (1984) *J. Biol. Chem. 259*, 10760−10763.
13. Engler, M. J., and Richardson, C. C. (1982) in *The Enzymes* (Boyer, P. D. Ed.) Vol. 15B, pp 3−29, Academic Press, New York.
14. Baker, T., Mizuuchi, M., Savilahti, H., and Mizuuchi, K. (1993) *Cell 74*, 723−733.
15. Luke, P. A., and Halford, S. E. (1984) *Gene 37*, 241−246.
16. Wentzell, L. M., Nobbs, T. J., and Halford, S. E. (1995) *J. Mol. Biol. 248*, 581−595.
17. Stec, W. J., Zon, G., and Uznanski, B. (1985) *J. Chromatogr. 326*, 263−280.
18. Roberts, R. J., and Halford, S. E. (1993) in *Nucleases* (Linn, S. M., Lloyd, R. S., and Roberts, R. J., Eds.) pp 35−88, Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
19. Nobbs, T. J., Szczelkun, M. D., Wentzell, L. M., and Halford, S. E. (1998) *J. Mol. Biol. 281*, 419−432.
20. Wentzell, L. M., and Halford, S. E. (1998) *J. Mol. Biol. 281*, 433−444.

21. Bilcock, D. T., and Halford, S. E. (1999) *Mol. Microbiol. 31* 1243−1254.
22. Craig, N. L., and Nash, H. A. (1983) *Cell 35*, 795−803.
23. Oller, A. R., Broek, W. V., Conrad, M., and Topal, M. D. (1991) *Biochemistry 30*, 2545−2549.
24. Nobbs, T. J., Williams, S. A., Connolly, B. A., and Halford, S. E. (1998) *Biol. Chem. 379*, 599−604.
25. Jo, K., and Topal, M. D. (1995) *Science 267*, 1817−1820.
26. Mizuuchi, K. (1992) *Annu Rev. Biochem. 61*, 1011−1051.
27. Mizuuchi, K. (1997) *Genes Cells 2*, 1−12.
28. Burgers, P. M. J., and Eckstein, F. (1979) *J. Biol. Chem. 254*, 6889−6893.
29. Connolly, B. A., Potter, B. V. L., Eckstein, F., Pingoud, A., and Grotjahn, L. (1984) *Biochemistry 23*, 3443−3453.
30. Fitzgerald, M. C., and Smith, L. M. (1995) *Annu. Rev. Biophys. Biomol. Struct. 24*, 117−140.
31. Pliura, D. H., Schomburg, D., Richard, J. P., Frey, P. A., and Knowles, J. R. (1980) *Biochemistry 19*, 325−329.
32. Conrad, M., and Topal, M. D. (1992) *Nucleic Acids Res. 20*, 5127−5130.
33. Connolly, B. A., Eckstein, F., and Grotjahn, L. (1984) *Biochemistry 23*, 2026−2031.

BI990054P