Nucleic Acids Research, 1994, Vol. 22, No. 21 ***4441–4448***

Methylphosphonodiester substitution near the conserved CA dinucleotide in the HIV LTR alters both extent of 3′-processing and choice of nucleophile by HIV-1 integrase

Abhijit Mazumder, Malini Gupta and Yves Pommier*
Laboratory of Molecular Pharmacology, Developmental Therapeutics Program, Division of Cancer Treatment, National Cancer Institute, Building 37, Room 5C25, Bethesda, MD 20892, USA

Received July 12, 1994; Revised and Accepted August 22, 1994

ABSTRACT

We present evidence suggesting that the 3′-processing activity of HIV-1 integrase is dramatically affected by electrostatic and/or steric perturbations 3′ to the conserved CA dinucleotide. When the phosphodiester bond 3′ to the scissile phosphodiester is replaced by a methylphosphonodiester linkage, 3′-processing decreases by two orders of magnitude. This block of cleavage can be somewhat overcome by increasing the pH of the reaction. Labeling of the substrates at the 3′-end revealed blockage of water and glycerol, but stimulation of the viral DNA 3′-hydroxyl, acting as the nucleophile with the methylphosphonodiester substrate. Interestingly, a circular trinucleotide was formed using the phosphodiester and methylphosphonodiester substrates when the terminal nucleotide was 3′-deoxyadenosine but not 2′-deoxyadenosine. Mutagenesis of the enzyme active site has previously been shown to alter the choice of nucleophile in the 3′-processing reaction. Taken together, the results in this study suggest that 'mutagenesis' of the DNA backbone can also alter the choice of nucleophile.

INTRODUCTION

Efficient replication of retroviral DNA requires establishment of the proviral state—i.e. the integration of a DNA copy of the viral genome, synthesized by reverse transcriptase, into a chromosome of the host cell. Integration is catalyzed by the viral integrase protein (1–4). Prior to integration, two nucleotides are excised from each 3′ end of the linear, blunt-ended, viral DNA (5–8). This 3′ processing reaction exposes the 3′-hydroxyl of a CA dinucleotide, which is conserved among all retroviruses (9,10). Each of these 3′-hydroxyl ends of the viral DNA are then joined to chromosomal DNA in the subsequent DNA strand transfer step (11–13). DNA strand transfer is an isoenergetic transesterification reaction (14). HIV-1 integrase catalyzes a nucleophilic attack of each 3′-hydroxyl group at the processed viral ends on a pair of phosphodiester bonds staggered by 5 bp in the target DNA (15–17). Completion of the integration process requires removal of the two unpaired nucleotides at the 5′ ends of the viral DNA and gap repair reactions that are thought to be accomplished by cellular enzymes. [See Katz and Skalka (18) and Vink and Plasterk (19) for recent reviews on retroviral DNA integration.]

It has previously been shown that integrases from several retroviruses can catalyze *in vitro* 3′-processing (5,7,8,20–23), strand transfer (11,12,15,20,22,23), and disintegration (24–26), an apparent reversal of the DNA strand transfer reaction. These reactions can be performed using short duplex oligodeoxyribonucleotides identical in sequence to either end of the viral DNA. The 3′-processing and DNA strand transfer reactions are one-step reactions (27), which do not involve a covalent intermediate between integrase and DNA (28). Although water is the most likely nucleophile *in vivo*, a variety of nucleophiles can participate in the 3′-processing reaction *in vitro*. For example, water, in an hydrolysis reaction, liberates a linear GT dinucleotide with 5′-phosphate and 3′-hydroxyl ends. Glycerol and other alcohols, in an alcoholysis reaction, liberate a product in which the alcohol is esterified to the 5′-phosphate of the linear GT dinucleotide (29,30). The 3′-hydroxyl of the viral DNA end can also act as a nucleophile, generating a cyclic GT dinucleotide (27).

The chemical mechanisms of the 3′-processing and DNA strand transfer reactions have been studied *in vitro*. The stereochemical course of the reactions catalyzed by HIV-1 integrase has been examined previously using an oligodeoxyribonucleotide containing a phosphorothioate of known chirality in the substrate DNA and determining the chirality in the reaction products (27). In this manner, it was determined that both reactions proceed via a one-step mechanism. The 3′-processing activity of HIV-1 integrase was explored further in the present report in light of previous findings with the *Eco*RI and *Eco*RV restriction endonucleases showing that the DNA substrate can hydrogen bond to the attacking water molecule, activating it (i.e. converting it into a better nucleophile) prior to attack on the scissile phosphodiester bond (31). The extent of 3′-processing via hydrolysis, alcoholysis, and circular nucleotide formation was examined using backbone-modified substrates. These substrates had a phosphodiester linkage (Fig. 1) replaced by a methylphos-

*To whom correspondence should be addressed

phonodiester at various positions near the conserved CA dinucleotide. Substitution of the negatively charged oxygen (of the phosphodiester 3′ to the scissile phosphodiester) by a neutral, hydrophobic methyl group (in a methylphosphonodiester) was found to block the 3′-processing reaction and to alter the choice of nucleophile. These effects could reflect disruption of electrostatic interactions or introduction of steric hinderance in the interaction of the substrate with the enzyme.

MATERIALS AND METHODS

Preparation of radiolabeled DNA substrates

The following oligonucleotides were purchased from Midland Certified Reagent Company (Midland, TX):

AE118, 5′-GTGTGGAAAATCTCTAGCAGT-3′;
AE117, 5′-ACTGCTAGAGATTTTCCACAC-3′;
RM22M, 5′-TACTGCTAGAGATTTTCCACAC-3′;
RMMP01, 5′-GTGTGGAAAATCTCTAGCA-X-GT-3′,

where X = methyl-phosphonodiester linkage;

RMMP02, 5′-GTGTGGAAAATCTCTAGCAG-X-T-3′;
RMMP03, 5′-GTGTGGAAAATCTCTAGC-X-AGT-3′;
RM020, 5′-GTGTGGAAAATCTCTAGCAG-3′.

These oligonucleotides were purified by HPLC. The AE117 and AE118 oligonucleotides correspond to the U5 end of the HIV LTR. Oligonucleotides RMMP01, RMMP02, and RMMP03 are methylphosphonate analogs of AE118. To analyze the extents of 3′-processing and integration using 5′-end labeled substrates, AE118, RMMP01, RMMP02, or RMMP03 were 5′-end labeled using T_4 polynucleotide kinase (Gibco BRL) and γ-[^{32}P]-ATP (Dupont-NEN). The kinase was heat-inactivated and AE117 was added to the same final concentration as its complementary strand. The mixture was heated at 95°C and allowed to cool slowly to room temperature. The reaction was then run on a G-25 Sephadex quick spin column (Boehringer Mannheim) to separate annealed double-stranded oligonucleotide from unincorporated label. To determine the choice of nucleophile in 3′-processing using 3′-end labeled DNA, substrates were prepared in one of two ways. By the first method, each oligonucleotide was 3′-end labeled using terminal transferase (Dupont-NEN) and cordycepin (3′-deoxyadenosine) α-[^{32}P]-triphosphate (Dupont-NEN). The reaction was stopped by the addition of EDTA and RM22M was added to the same final concentration as its complementary strand. The reaction was then run on a G-25 Sephadex quick spin column as before. By the second method, each oligonucleotide was annealed to RM22M and radiolabeled at the 3′-end using Sequenase version 2.0 (Stratagene) and α-[^{32}P]-2′-deoxyadenosine triphosphate (Dupont-NEN) or α-[^{32}P]-2′,3′-dideoxyadenosine triphosphate (Amersham). The reaction was then run on a G-25 Sephadex quick spin column as before.

3′-processing and integration assays

Purified recombinant HIV-1 integrase was a generous gift of Dr R.Craigie, Laboratory of Molecular Biology, NIDDK, NIH, Bethesda, MD. The enzyme was incubated at a final concentration of 250 nM with 15 nM of the labeled oligonucleotide substrate in reaction buffer (50 mM NaCl, 1 mM HEPES, 50 μM EDTA, 50 μM dithiothreitol, 10% glycerol (wt/vol) unless otherwise specified, 7.5 mM $MnCl_2$ or 7.5 mM $MgCl_2$, 0.1 mg/ml bovine serum albumin, 10 mM 2-mercaptoethanol, 10% dimethyl sulfoxide, and 25 mM MOPS, pH 7.2 unless otherwise specified) at 30°C for 60 min. The 20-fold excess of enzyme was used in order for the 3′-processing to be detected more readily. The final reaction volume was 16 μl. Reactions were quenched by the addition of an equal volume of Maxam−Gilbert loading dye. An aliquot was electrophoresed on a denaturing 20% polyacrylamide gel. Gels were dried and exposed in a Molecular Dynamics Phosphorimager cassette. Gels were analyzed using a Molecular Dynamics phosphorimager (Sunnyvale, CA) or a Betascope 603 blot analyzer (Betagen, Waltham, MA).

Integrase-DNA binding assay

Reaction mixtures (10 μl) containing 250 nM HIV-1 integrase were incubated with 15 nM of the radiolabeled oligonucleotide in reaction buffer for 5 min at 30°C. Eight μl of the reaction mixture was then spotted on to a nitrocellulose filter that had been prewetted with 25 mM MOPS, pH 7.2. The filters were washed under vaccum with 5 ml of the same MOPS buffer and the radioactivity retained on the filter was counted to measure the amount of DNA retention. The data were corrected for non-specific binding of the free DNA to the filter (usually about 0.3%).

Calf intestinal phosphatase (CIP) or T_4 polynucleotide kinase (PNK) treatment of integrase reaction products

The integrase reaction products obtained in the presence of manganese and glycerol using 3′-end labeled DNA were mixed with the appropriate reaction buffer and either 3 units of CIP or 5 units of PNK and 50 μM ATP such that the integrase reaction was diluted three-fold. The incubation was continued at 37°C for an additional 1 h.

RESULTS

A neutral methylphosphonodiester linkage 3′ to the scissile bond blocks the 3′-processing reaction

In order to examine possible roles of the negatively charged, non-bridging oxygen of various phosphodiester linkages (see Fig. 1) in the HIV U5 DNA end, oligonucleotides were made with a methylphosphonodiester linkage at the cleavage site (oligo B) or 5′- (oligo A) or 3′- (oligo C) to the site of cleavage (Fig. 1). The chemical synthesis produces both the R and S diastereomers, each of which was purified, separated from the other, and investigated in integrase assays. The results obtained using either the R or S diastereomer were similar. For simplicity, the results obtained from only one set of diasteromers are reported. The absolute configuration of the diastereomers used in this study has not been assigned although it most likely is the R configuration where the methyl group is more solvent exposed (32).

Filter binding assays demonstrated that introduction of the methylphosphonodiesters did not significantly alter the ability of the enzyme to bind to any of the DNA substrates (Table 1). However, substitution of the phosphodiester 3′ to the scissile bond with a methylphosphonodiester resulted in a 100-fold reduction in the extent of 3′-processing (oligo C; Fig. 2, lane 9 and Table 1). 3′-processing was much less affected when the methylphosphonodiester was placed at (oligo B) or 5′ to (oligo A) the scissile bond (Fig. 2, lanes 5 and 7 and Table 1). Placing the methylphosphonodiester at the scissile bond (oligo B) resulted in two bands probably corresponding to two products, one having an hydroxyl terminus and one having a methylphosphonate terminus, as is expected from the hydrolysis of methylphosphonodiesters (33). Therefore, either the elimination of the negative charge of the phosphodiester 3′ to the scissile bond (oligo C) and/or the

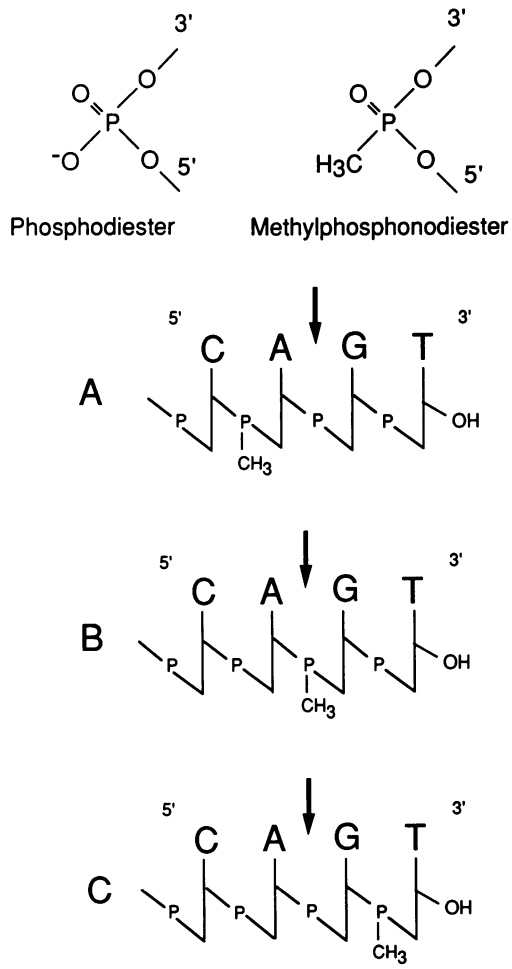

Figure 1. Structures and positions of the normal and methylphosphonodiester internucleotide linkages in oligonucleotide substrates A, B, and C. The sequence of the last four nucleotides in the methylphosphonodiester oligonucleotide substrates is shown. The scissile bond and methylphosphonodiester linkage in each oligonucleotide are indicated by an arrow and by a methyl (CH_3) group, respectively.

Table 1. DNA binding and 3′ processing of the phosphodiester and methylphosphonodiester U5 LTR oligonucleotides by HIV-1 integrase

Substrate	% DNA bound	3′-processing[a]
Normal	64.5 ± 12.2	14.79 ± 4.92
A	40.0 ± 10.8	2.1 ± 0.15
B	53.8 ± 5.4	6.38 ± 1.44[b]
C	8.4 ± 3.9	0.15 ± 0.10

[a]Expressed as percent of total counts.
[b]Total of the two products, hydroxyl- (19mer-OH) and methylphosphonate- (19mer-Mephos) terminated 19mers, obtained.

steric hinderance introduced by substitution of the methyl group dramatically affects 3′-processing.

Is the negative charge of the phosphodiester 3′ to the scissile bond involved in catalysis?

The very small but still detectable extent of 3′-processing of oligo C by integrase allowed quantification of the pH dependence of this reaction. The extent of 3′-processing for the normal

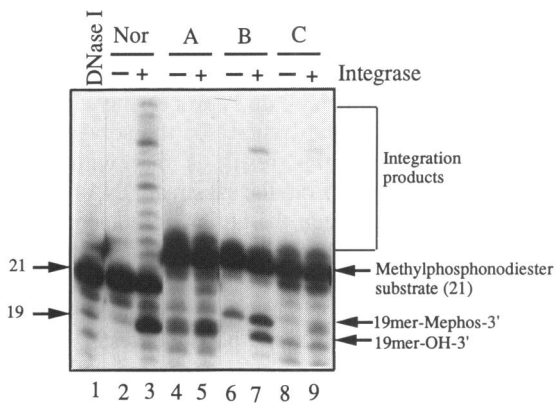

Figure 2. 3′-processing of the phosphodiester and methylphosphonodiester substrates. A phosphorimager picture of a 20% polyacrylamide denaturing gel is shown. The letters above the lanes correspond to the structures depicted in Figure 1. Lane 1, DNase I digest of the normal substrate; lanes 2, 4, 6, and 8, DNA alone; lanes 3, 5, 7, and 9, with HIV-1 integrase. The migrations of the methylphosphonodiester and normal substrates and of the two products, hydroxyl- (19mer-OH-3′) and methylphosphonate- (19mer-Mephos-3′) terminated, are indicated.

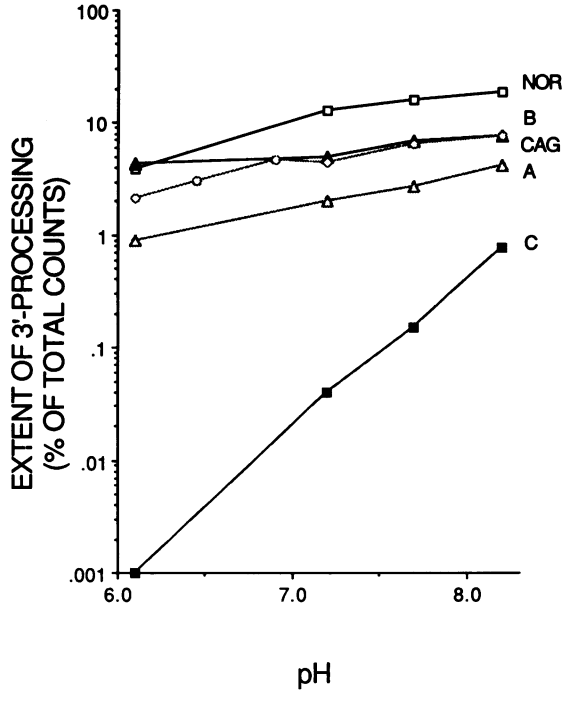

Figure 3. pH dependence of the 3′-processing reaction with either the normal (phosphodiester) or methylphosphonodiester substrates. The normal (phosphodiester), A, B, C (see Fig. 1), and CAG-terminated substrates are depicted by the open squares, open triangles, filled triangles, filled squares, and open circles, respectively.

phosphodiester substrate and substrates A and B (see Fig. 1) increased by a factor of 1.5−2 over the pH range of 7.2−8.2 (Fig. 3). However, the extent of 3′-processing for substrate C (where the methylphosphonodiester is 3′ to the scissile phosphodiester) increased by a factor of 19.5 over the same pH range (Fig. 3). The numbers used in this latter calculation were

very small but still in excess of background levels by four- to six-fold. Apparently, the block of 3′-processing can be partially overcome by increasing the basicity of the solvent.

One interpretation of this result is that with the normal oligonucleotide, assistance from the substrate DNA (via formation of a hydrogen bond between the attacking nucleophile and the negatively charged oxygen of the phosphodiester linkage 3′ to the scissile phosphodiester) results in deprotonation of a water or glycerol molecule or of the viral 3′-hydroxyl, prior to its nucleophilic attack on the scissile phosphodiester bond. In this case, hydroxide or alkoxide anion is generated within the active site, and exogenous hydroxide anion supplied by the increasing basicity of the solvent has little effect on the reaction. In contrast, with substrate C, where a methylphosphonodiester is 3′ to the scissile phosphodiester, such substrate activation is impossible. Endogenous hydroxide or alkoxide anion cannot be produced, and exogenous hydroxide anion must be supplied by the solvent.

An argument against this hypothesis is the fact that a significant increase in the extent of 3′-processing over the same pH range does not occur using an oligonucleotide in which the scissile phosphodiester is also the last internucleotide linkage (i.e. the last three bases are CAG). This oligonucleotide would not be able to form the hypothesized hydrogen bond with an incoming nucleophile because the appropriate phosphodiester linkage does not exist and would not, therefore, be expected to be able to use substrate assistance. Its pH profile should, then, be similar to that of oligo C. However, as seen in Fig. 3, its pH profile is analogous to that of the normal, A, and B substrates. Therefore, our data do not support the view that the phosphodiester bond 3′-to the scissile phosphodiester is actively involved in catalysis. Thus, the block to 3′-processing observed with substrate C is probably due to either steric hinderance or disruption of electrostatic interactions.

Backbone-modified DNA substrates alter the choice of nucleophile in the 3′-processing reaction

Because several nucleophiles can participate in the 3′-processing reaction (27,29), we determined whether one or all of the nucleophiles were blocked from attack using the methylphosphonodiester substrates. As depicted in Figure 4, alcoholysis, hydrolysis, and cyclic trinucleotide formation could occur, although cyclic nucleotides are usually not detected when one or three nucleotides are removed by integrase (29,30). In order to aid in identification of the reaction products from each pathway, assay conditions were used which do not allow formation of the glycerol or cyclic nucleotide product. For example, omitting glycerol from the reaction and using magnesium instead of manganese block the formation of the glycerol adduct and cyclic nucleotide product, respectively (30). In addition, the sensitivity of the reaction products to either phosphatase or polynucleotide kinase treatment should confirm the identity of each product. It has been previously demonstrated that the linear dinucleotide is phosphatase sensitive, the glycerol adduct can be inefficiently phosphorylated, and circular nucleotides are insensitive to either phosphatase or kinase treatment (27).

Due to the presence of the methylphosphonodiester, labeling of the terminal phosphodiester could not be accomplished using a DNA polymerase-like enzyme in all cases. Therefore, extension

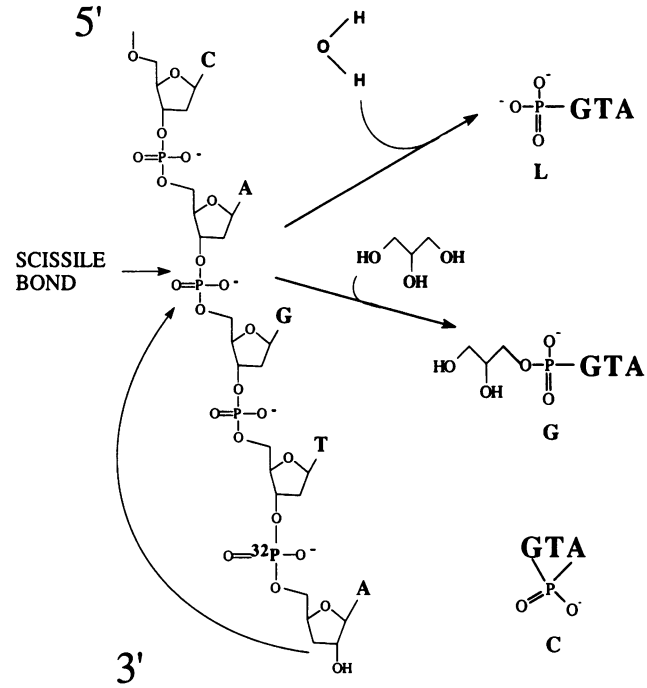

Figure 4. Schematic diagram showing the 3′-terminus of the oligonucleotide following incorporation of cordycepin (3′-deoxyadenosine) used to study the 3′-processing products. The location of the 3′-label is indicated by the bold ^{32}P in the last phosphodiester linkage. The possible nucleophiles in the 3′-processing reaction, and the expected products from each nucleophilic attack are shown to the right. L, linear trinucleotide from water attack; G, glycerol trinucleotide adduct from glycerol attack; and C, cyclic trinucleotide from DNA 2′-hydroxyl attack.

by one nucleotide beyond the GT dinucleotide was chosen in order to label the substrates at the 3′-end. In this case, a trinucleotide or trinucleotide adduct would be released. It has previously been shown that extension by 1 bp beyond the usual GT dinucleotide has only a moderate effect on the reaction efficiency (21). Labeling at the 3′-end was found to be most efficient using terminal transferase rather than a DNA polymerase (data not shown), probably due to the presence of the methylphosphonodiester.

The radiolabeled substrates were generated using terminal transferase and α-[^{32}P]-cordycepin (3′-deoxyadenosine) triphosphate and, therefore, have an extra base pair beyond the usual GT of the processed strands and are terminated by a 2′- instead of a 3′-hydroxyl (Fig. 4). These substrates are 22 bp in length and are designated N′, A′, B′, and C′ to distinguish them from the corresponding 21 bp duplexes. The results obtained are shown in Fig. 5. As seen in panel A (lane 3), three reaction products were formed using the phosphodiester substrate. The fastest migrating product comigrates with authentic linear trinucleotide and its migration rate is decreased after treatment with phosphatase (panel D, lane 1). The slowest migrating product has the expected mobility of a glycerol adduct (27,29) and was inefficiently phosphorylated by polynucleotide kinase (data not shown). Omitting glycerol from the reactions or using magnesium resulted in complete or significant loss, respectively, of the glycerol adduct (panels B, lane 1 and C, lane 1). The amount of linear trinucleotide product was not significantly diminished under either of these reaction conditions. The product of intermediate mobility was not sensitive to phosphatase (panel

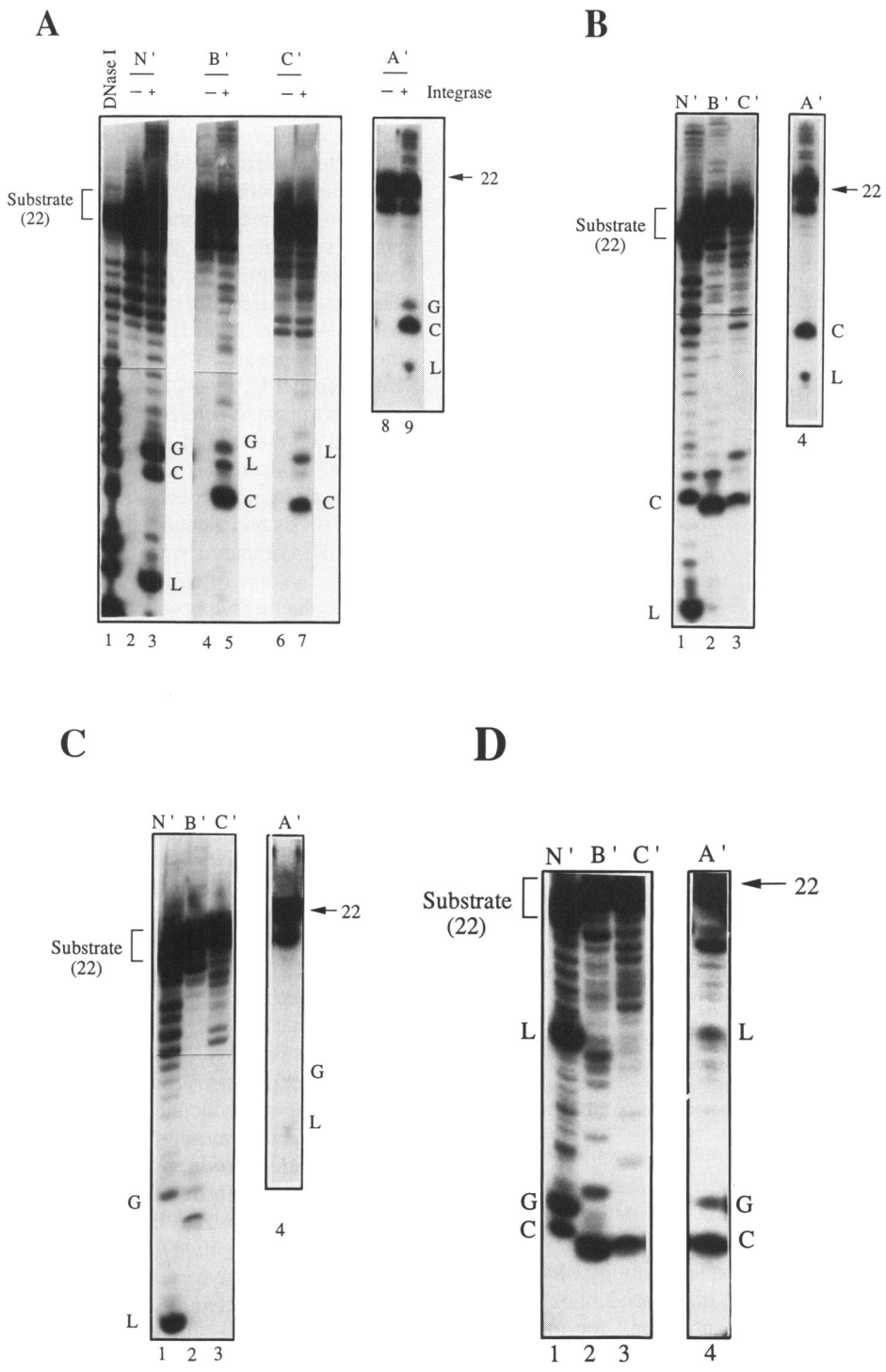

Figure 5. 3′-Processing by HIV-1 integrase of 3′-end labeled phosphodiester and methylphosphonodiester 22mer oligonucleotide substrates containing three nucleotides (5′-GTA-3′) after the conserved CA residue (see Fig. 4). The letters above the lanes correspond to the oligonucleotide used. The lanes corresponding to substrate A′ were loaded on a separate gel which was not electrophoresed as far. (**A**) Reactions performed in the presence of manganese and 10% glycerol. Lane 1, DNase I digest of the normal substrate; lanes 2, 4, 6, and 8, DNA alone; lanes 3, 5, 7, and 9, with integrase. The positions of the glycerol adduct (G) and cyclic (C) and linear (L) trinucleotides generated from the normal substrate and methylphosphonodiester substrates (22mer) are indicated. The extent of 3′-processing with oligo C′ (lane 7) is not reduced by two orders of magnitude compared with the normal N′ oligo (lane 3). We presume this difference is due to the fact that oligo C′ has 3 bp pairs beyond the CA dinucleotide (as opposed to 2 bp in the case of oligo C) and terminates with a 3′-deoxynucleotide. (**B**) Reactions performed in the presence of manganese and no glycerol. Lanes 1−4, with integrase. The positions of the substrates and of the cyclic (C) and linear (L) trinucleotides generated from the normal substrate and substrate A′ are indicated. (**C**) Reactions performed in the presence of magnesium and 10% glycerol. Lanes 1−4, with integrase. The positions of the substrates and of the glycerol (G) and linear (L) trinucleotides generated from the normal substrate and substrate A′ are indicated. (**D**) Calf intestinal phosphatase digest of the integrase reaction products obtained in the presence of manganese and glycerol. Lanes 1−4, with integrase and calf intestinal phosphatase. The positions of the substrates, the glycerol adduct (G), and of the cyclic (C) and linear (L) trinucleotides generated from the normal substrate and substrate A′ are indicated.

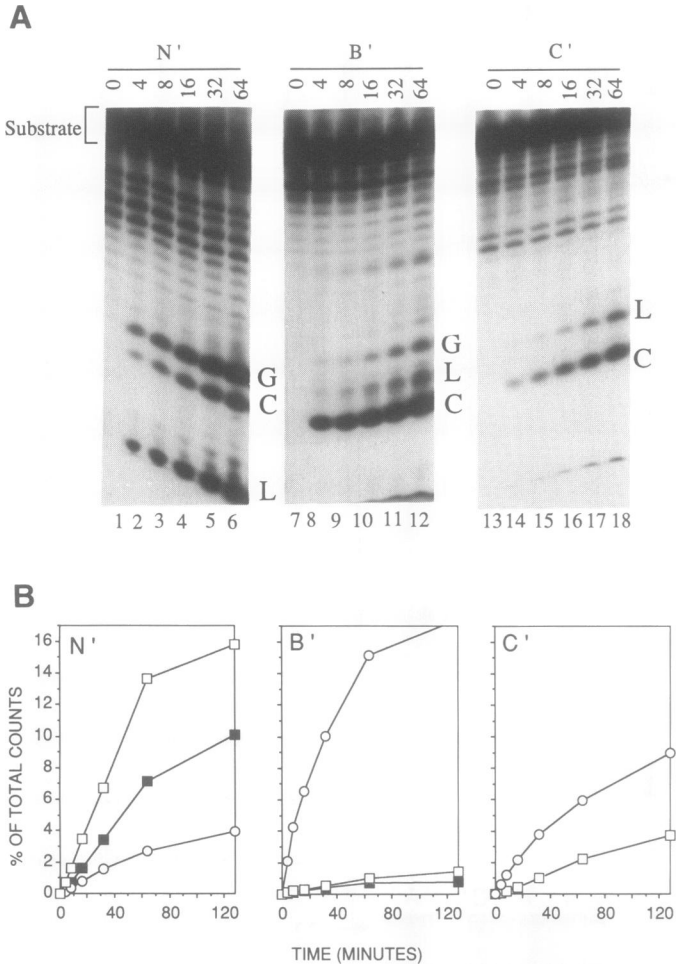

Figure 6. Kinetics of the 3′-processing reaction with the normal and methylphosphonodiester substrates. (**A**) The letters above the lanes correspond to the oligonucleotide used. Lanes 1, 7, and 13, DNA alone; lanes 2–6, 8–12, and 14–18, integrase incubated for the indicated time (in minutes). The positions of the substrate and of the glycerol adduct (G) and cyclic (C) and linear (L) trinucleotides generated from each substrate are indicated. (**B**) Quantitation of the amounts of product (expressed as percent of total counts) produced in Figure 6A. The linear trinucleotide (L), glycerol trinucleotide adduct (G), and circular trinucleotide (C) are depicted by the open squares, filled squares, and open circles, respectively.

D, lane 1) or kinase and had the expected mobility of a circular trinucleotide (27). This product, while not dependent on glycerol, disappeared when magnesium was substituted for manganese (panel C, lane 1). Furthermore, this product of intermediate mobility was not formed when other radiolabeled substrates with a 3′-hydroxyl terminus or a 2′,3′-dideoxy terminus were used, consistent with previous findings (27,29). Taken together, these data indicate that the intermediate product is the circular trinucleotide.

When the methylphosphonodiester was placed 5′ to the scissile bond (substrate A′), three products were also detected in the standard buffer containing manganese and glycerol (Fig. 5, panel A, lane 9). The slowest migrating product was glycerol-dependent (compare panels A, lane 9 and B, lane 4). The product with intermediate migration only disappeared when magnesium was used instead of manganese (panel C, lane 4). The fastest migrating product comigrated with authentic linear trinucleotide and its migration rate is decreased after treatment with phosphatase (panel D, lane 4). Therefore, the three products in the case of substrate A′ are, from top to bottom, the glycerol adduct (G) and circular (C) and linear (L) trinucleotides, respectively. The extents of hydrolysis and alcoholysis were decreased with this substrate compared with the normal phosphodiester substrate while the amount of circular nucleotide formation was increased.

When the methylphosphonodiester was placed at the scissile bond (substrate B′), three products were detected in buffer containing manganese and glycerol (Fig. 5, panel A, lane 5). The slowest migrating product was glycerol-dependent (compare panels A, lane 5 and B, lane 2). The fastest migrating product mostly disappeared when magnesium was used as the divalent cation (panel C, lane 2). The product of intermediate migation was phosphatase sensitive (panel D, lane 2). Thus, the three products in the case of substrate B′ are, from top to bottom, the glycerol adduct (G) and the linear (L) and circular (C) trinucleotides, respectively. As in the case of oligo A′, the extents of hydrolysis and alcoholysis were decreased with this substrate (oligo B′) compared with the normal phosphodiester substrate while the amount of circular nucleotide formation was increased.

When the methylphosphonodiester was placed 3′ to the scissile bond (substrate C′), only two products were detected in buffer containing manganese and glycerol (Fig. 5, panel A, lane 7). Neither product was glycerol-dependent (compare panels A, lane 7 and B, lane 3). The faster migrating product was phosphatase insensitive and disappeared when magnesium was used as the divalent cation (panel C, lane 3), consistent with it being the circular trinucleotide. The slower migrating product was phosphatase sensitive (panel D, lane 3), consistent with it being a linear trinucleotide. The extents of hydrolysis and, particularly, alcoholysis were markedly decreased with this substrate compared with the normal phosphodiester substrate while the amount of circular nucleotide formation was not significantly affected. Therefore, the block to 3′-processing observed when a methylphosphonodiester is placed 3′ to the scissile bond (Fig. 2) is primarily due to complete and marked blocks in the alcoholysis and hydrolysis reactions, respectively.

Kinetic experiments demonstrate that the results observed in Figure 5 are reproducible with reaction times between 4 and 128 minutes. Furthermore, the rate of formation of the circular trinucleotide product was markedly increased in the oligonucleotide substrate with the methylphosphonodiester at the scissile bond (substrate B′; Fig. 6, middle panels). In contrast, in this oligonucleotide (B′) the rate and yield of the linear trinucleotide and glycerol adducts were inhibited. The main effect with the oligonucleotide substrate with the methylphosphonodiester 3′ to the scissile bond (substrate C′; Fig. 6, right panels) was a five-fold reduction in the rate of formation of the linear trinucleotide and a complete inhibition of glycerol adduct formation.

DISCUSSION

Methylphosphonate oligonucleotides were originally developed for applications in antisense drug design (34). Recently, however, substitution of a phosphodiester with a methylphosphonodiester has become a powerful tool to study DNA–protein interactions involved in binding and catalysis and has been used to examine protein side chain–DNA phosphate contacts in several systems (35,36). Where a methylphosphonodiester is substituted for a

phosphodiester, the negative charge is removed and the local hydrophobic character of the internucleotide linkage is increased (Fig. 1A and B). Molecular dynamics simulations suggest that this analog is well-tolerated in B-DNA, mainly causing a significant reduction in local intrastrand and interstrand electrostatic repulsive interactions (37) and counterion distribution (38).

This technique can complement ethylation interference experiments (39,40) and offers several advantages over the latter. First, steric perturbations are dramatically smaller. A methylphosphonodiester introduces about a 0.5 Å perturbation while ethylation interference introduces about a 4.5 Å perturbation (36). Secondly, the R and S diastereomers could differentially influence hydrogen bonding schemes. Absence of such stereoselectivity limits interpretation of partial ethylation interference. Thirdly, incorporation of a methylphosphonodiester in discrete sites in an oligonucleotide allows quantitative determination of alterations in binding and catalysis.

Introduction of a methylphosphonodiester near the conserved CA dinucleotide affects the 3′-processing reaction

When a neutral methylphosphonodiester bond is substituted for a negatively charged phosphodiester bond at various positions near the conserved CA dinucleotide, binding of the integrase to the modified substrate was not significantly reduced as determined using filter binding. However, this assay cannot separate the non-specific and specific binding capacities of integrase. When this methylphosphonodiester is 3′ to the scissile bond, the extent of 3′-processing decreases by two orders of magnitude. The other diastereomer has also been tested and analogous blocks to hydrolysis have been observed (data not shown). Exogenous hydroxide anion, supplied by increasing the pH of the reaction, can partially overcome the block to 3′-processing. The block to 3′-processing presumably is due to the complete and marked blocks in the alcoholysis and hydrolysis reactions, respectively. Taken together, these results suggest that when a methylphosphonodiester is 3′ to the scissile phosphodiester, the change in electrostatic interactions and, more importantly, introduction of steric hinderance caused by the neutral, hydrophobic methyl group alters the interactions between HIV-1 integrase and its substrate DNA and ultimately changes the choice of nucleophile.

Ethylation interference experiments have previously been used to define functionally important interactions between integrase and substrate DNA (39). These experiments defined the phosphodiester bonds 5′ and 3′ to the scissile bond and the scissile bond itself as critical for integration to occur. Because the method used relied on analyzing integration products rather than physically assaying binding, these sites were determined to be necessary for catalysis. The results presented in this report are consistent with these findings. The extent of integration is slightly lowered using substrates A, A′, B, and B′ and significantly diminished using substrates C and C′ (Figs 2 and 5). This reduction can be explained by the inability to use glycerol or water as the nucleophile in the previous 3′-processing step, either due to the presence of a neutral methylphosphonodiester (as in this report) or of a neutral ethyl phosphotriester (39). Furthermore, the substrate which terminates with the sequence CAG-3′ (and is therefore missing the ultimate phosphodiester linkage and ultimate nucleotide) shows a three-fold reduction in 3′-processing, also consistent with the last phosphodiester being important in the 3′-processing and DNA strand transfer reactions.

The chemical nature of the 3′ end of viral DNA and the choice of nucleophile in the 3′-processing reaction

Retroviral integrases can use a variety of nucleophiles in the 3′-processing reaction (27,29). HIV-1 integrase prefers water or glycerol as the nucleophile when a di- or trinucleotide is released [(30) and present study]. However, FIV (feline immunodeficiency virus) (22) and RSV (rous sarcoma virus) (41) integrases (using different reaction conditions) have been found to use the 3′-hydroxyl of the viral DNA as the nucleophile. Therefore, the amino acids in the immediate vicinity of the active site probably play a role in the use of one nucleophile over the other.

In this report, it was demonstrated that 3′-processing by HIV-1 integrase can generate a circular trinucleotide (approximately 12% of the reaction products). This circular product is only formed when a 2′-hydroxyl, but not a 3′-hydroxyl, on the viral DNA is available. This result is consistent with the circular nucleotides having not been observed when one or three nucleotides are released (29,30). 3′-processing probably requires a distortion of the viral DNA end in order to bring the nucleophilic 3′-hydroxyl in close proximity to the scissile bond for the S_N2 reaction and formation of the circular trinucleotide. There may be less distortion required for alignment to occur if a 2′-hydroxyl were used as the nucleophile. This hypothesis can be supported by examination of the deoxyribose conformation in B-DNA. The sugar pucker in B-DNA is C-2′-*endo* where the C2′ and C3′ carbons are on opposite sides of the sugar plane defined by the C1′, O, and C4′ atoms (42). A 2′-hydroxyl would be spatially oriented in a different manner than a 3′-hydroxyl and bond distances and angles of subsequent linkages needed to form a circular trinucleotide could be more favorable. This could explain why a 2′-hydroxyl is more favorable than a 3′-hydroxyl for the formation of the circular trinucleotide.

The substrates A′ and B′ which contained methylphosphonodiester linkages 5′ to or at the scissile bond, respectively, resulted in the circular trinucleotide being the major reaction product. The increased hydrophobicity of the methylphosphonodiester and elimination of possible repulsive electrostatic interactions between this phosphodiester (located within the conserved CA dinucleotide) and integrase probably plays significant roles in determining the choice of the 2′-hydroxyl DNA terminus as a nucleophile for substrates A and A′. For substrates B and B′, charge neutralization in the scissile phosphodiester probably eliminates any electrostatic repulsion between this methylphosphonodiester and the attacking viral 3′- or 2′-hydroxyl, respectively, allowing it to serve as a good nucleophile. Therefore, several factors are important in the choice of nucleophile by integrase, including enzyme active site (30) and the DNA structure and resulting steric and electrostatic interactions between the DNA and amino acid residues of the enzyme (present study).

In summary, replacement of a phosphodiester with a methylphosphonodiester near the conserved CA dinucleotide can dramatically affect the 3′-processing and DNA strand transfer reactions. Alteration of either the electrostatic or steric nature at these sites can affect the efficiency of catalysis as well as the choice of nucleophile in the 3′-processing reaction. The design and future evaluation of oligonucleotides as inhibitors of HIV-1 integrase may profit from these results.

ACKNOWLEDGEMENTS

We would like to thank Drs Robert Craigie and Alan Engelman (Laboratory of Molecular Biology, NIDDK) for generously

providing us with purified HIV-1 integrase and A.Engelman for critical reading of this manuscript. We would also like to thank Dr Kurt Kohn (Chief, Laboratory of Molecular Pharmacology, NCI) for stimulating discussions during the course of these experiments.

REFERENCES

1. Donehower,L.A. and Varmus,H.E. (1984) *Proc. Natl Acad. Sci. USA*, **81**, 6461–6465.
2. Schwartzberg,P., Colicelli,J. and Goff,S.P. (1984) *Cell*, **37**, 1043–1052.
3. Panganiban,A.T. and Temin,H.M. (1984) *Proc. Natl Acad. Sci. USA*, **81**, 7885–7889.
4. Hippenmeyer,P.J. and Grandgenett,D.P. (1984) *Virology*, **137**, 358–362.
5. Katzman,M., Katz,R.A., Skalka,A.M. and Leis,J. (1989) *J. Virol.*, **63**, 5319–5327.
6. Roth,M.J., Schwartzberg,P.L. and Goff,S.P. (1989) *Cell*, **58**, 47–54.
7. Sherman,P.A. and Fyfe,J.A. (1990) *Proc. Natl Acad. Sci. USA*, **87**, 5119–5123.
8. Vora,A.C., Fitzgerald,M.L. and Grandgenett,D.P. (1990) *J. Virol.*, **64**, 5656–5659.
9. Kulkosky,J. and Skalka,A.M. (1990) *J. Acquired Imm. Deficiency Synd.*, **3**, 839–851.
10. Varmus,H.E. and Brown,P.O. (1989) *Mobile DNA*, Eds., Berg, D. and Howe, M., 53–108, Am. Soc. Microbiol., Washington, DC.
11. Craigie,R., Fujiwara,T. and Bushman,F.D. (1990) *Cell*, **62**, 829–837.
12. Katz,R.A., Merkel,G., Kulkosky,J., Leis,J. and Skalka,A.M. (1990) *Cell*, **63**, 87–95.
13. Fujiwara,T. and Mizuuchi,K. (1988) *Cell*, **54**, 497–504.
14. Brown,P.O., Bowerman,B., Varmus,H.E. and Bishop,J.M. (1987) *Cell*, **49**, 347–356.
15. Bushman,F.D., Fujiwara,T. and Craigie,R. (1990) *Science*, **249**, 1555–1558.
16. Vincent,K.A., York-Higgins,D., Quiroga,M. and Brown,P.O. (1990) *Nucleic Acids Res.*, **18**, 6045–6047.
17. Vink,C., Groenink,M., Elgersma,Y., Fouchier,R.A.M., Tersmette,M. and Plasterk,R.H.A. (1990) *J. Virol.*, **64**, 5626–5627.
18. Katz,R.A. and Skalka,A.M. (1994) *Annu. Rev. Biochem.*, **63**, 133–173.
19. Vink,C. and Plasterk,R.H.A. (1993) *Trends Genet.*, **9**, 433–437.
20. Bushman,F.D. and Craigie,R. (1991) *Proc. Natl Acad. Sci. USA*, **88**, 1339–1343.
21. Vink,C., van Gent,D.C., Elgersma,Y. and Plasterk,R.H.A. (1991) *J. Virol.*, **65**, 4636–4644.
22. Vink,C., van der Linden,K.H. and Plasterk,R.H.A. (1994) *J. Virol.*, **68**, 1468–1474.
23. Katzman,M. and Sudol,M. (1994) *J. Virol.*, **68**, 3558–3569.
24. Chow,S.A., Vincent,K.A., Ellison,V. and Brown,P.O. (1992) *Science*, **255**, 723–726.
25. Sherman,P.A., Dickson,M.L. and Fyfe,J.A. (1992) *J. Virol.*, **66**, 3593–3601.
26. Donzella,G.A., Jonsson,C.B. and Roth,M.J. (1993) *J. Virol.*, **67**, 7077–7087.
27. Engelman,A., Mizuuchi,K. and Craigie,R. (1991) *Cell*, **67**, 1211–1220.
28. Mizuuchi,K. (1992) *J. Biol. Chem.*, **267**, 21273–21276.
29. Vink,C., Yeheskiely,E., van der Marel,G.A., van Boom,J.H. and Plasterk,R.H.A. (1991) *Nucleic Acids Res.*, **19**, 6691–6698.
30. van Gent,D.C., Oude Groeneger,A.A. and Plasterk,R.H.A. (1993) *Nucleic Acids Res.*, **21**, 3373–3377.
31. Jeltsch,A., Alves,J., Wolfes,H., Maass,G. and Pingoud,A. (1993) *Proc. Natl Acad. Sci. USA*, **90**, 8499–8503.
32. Ferguson,D.M. and Kollman,P.A. (1991) *Antisense Res. Dev.*, **1**, 243–254.
33. Murakami,A., Blake,K.R. and Miller,P.S. (1985) *Biochemistry*, **24**, 4041–4046.
34. Miller,P.S. and Ts'o,P.O.P. (1987) *Anti-Cancer Drug Des.*, **2**, 117–128.
35. Hamy,F., Asseline,U., Grasby,J., Iwai,S., Pritchard,C., Slim,G., Butler,P.J., Karn,J. and Gait,M.J. (1993) *J. Mol. Biol.*, **230**, 111–123.
36. Botfield,M.C. and Weiss,M.A. (1994) *Biochemistry*, **33**, 2349–2355.
37. Swarnalatha,Y. and Yathindra,N. (1993) *J. Biomol. Struct. Dynam.*, **10**, 1023–1045.
38. Hausheer,F.H., Singh,U.C., Palmer,T.C. and Saxe,J.D. (1990) *J. Am. Chem. Soc.*, **112**, 9468–9474.
39. Bushman,F.D. and Craigie,R. (1992) *Proc. Natl Acad. Sci. USA*, **89**, 3458–3462.
40. Chow,S.A. and Brown,P.O. (1994) *J. Virol.*, **68**, 3896–3907.
41. Muller,B., Jones,K.S., Merkel,G.W. and Skalka,A.M. (1993) *Proc. Natl Acad. Sci. USA*, **90**, 11633–11637.
42. Saenger,W. (1984) *Principles of Nucleic Acid Structure*, Springer, New York.