THE JOURNAL OF BIOLOGICAL CHEMISTRY Vol. 278, No. 18, Issue of May 2, pp. 16280–16288, 2003
*Printed in U.S.A.*

# Trapping HIV-1 Reverse Transcriptase Before and After Translocation on DNA*

Received for publication, December 18, 2002
Published, JBC Papers in Press, January 28, 2003, DOI 10.1074/jbc.M212911200

**Stefan G. Sarafianos‡, Arthur D. Clark, Jr.‡, Steve Tuske‡, Christopher J. Squire‡, Kalyan Das‡, Dequan Sheng‡, Palanichamy Ilankumaran§, Andagar R. Ramesha§, Heiko Kroth§, Jane M. Sayer§, Donald M. Jerina§, Paul L. Boyer¶, Stephen H. Hughes¶, and Eddy Arnold‡‖**

*From the ‡Center for Advanced Biotechnology and Medicine and Department of Chemistry and Chemical Biology, Rutgers University, Piscataway, New Jersey 08854-5638, the §Laboratory of Bioorganic Chemistry, NIDDK, National Institutes of Health (NIH), Bethesda, Maryland 20892-0820, and the ¶HIV Drug Resistance Program, NCI-Frederick Cancer Research and Development Center, NIH, Frederick, Maryland 21702-1201*

**A disulfide cross-linking strategy was used to covalently trap as a stable complex (complex N) a short-lived, kinetic intermediate in DNA polymerization. This intermediate corresponds to the product of polymerization prior to translocation. We also prepared the trapped complex that corresponds to the product of polymerization after translocation (complex P). The cross-linking method that we used is a variation of a technique developed by the Verdine and Harrison laboratories. It involves disulfide interchange between an engineered sulfhydryl group of the protein (Q258C mutation) and a disulfide-containing tether attached at the $N^2$ amino group of a modified dG in either the template or the primer strand of the nucleic acid. We report here a highly efficient synthesis of the precursor, bis(3-aminopropyl)disulfide dihydrochloride, used to introduce this substituent into the oligonucleotide. Efficient cross-linking takes place when the base pair containing the substituent is positioned seven registers from the dNTP-binding site (N site) and the N site is occupied. Complex N, but not complex P, is a substrate for the ATP-based excision reaction that unblocks nucleoside reverse transcriptase inhibitor (NRTI)-terminated primers and causes resistance to several NRTIs, confirming predictions that the excision reaction takes place only when the 3′-end of the primer is bound at the N site. These techniques can be used for biochemical and structural studies of the mechanism of DNA polymerization, translocation, and excision-based resistance of RT to NRTIs. They may also be useful in studying other DNA or RNA polymerases or other enzymes.**

HIV-1[1] reverse transcriptase (RT) is a complex molecular machine that uses several kinetically distinct steps to incorporate a nucleotide into a growing DNA strand. It is a heterodimer composed of a larger 560-residue subunit (p66) and a smaller subunit (p51) that contains the N-terminal 440 residues of p66. Both subunits contain subdomains that were named fingers, palm, thumb, and connection, because of the similarity of p66 to a right hand. The DNA polymerase active site is located in the p66 palm subdomain and the DNA binding cleft is formed primarily by the p66 fingers, palm, and thumb subdomains. The mechanism of polymerization by RT is similar to other polymerases and involves: 1) binding of the DNA substrate to the apo-enzyme; 2) binding of dNTP and divalent metal ions (required for catalysis) to the enzyme·DNA complex, followed by rate-limiting conformational changes; 3) formation of a phosphodiester bond between the 3′-OH primer terminus and the $\alpha$-phosphate of dNTP, followed by release of the pyrophosphate product; 4) translocation of the elongated DNA primer (for processive synthesis) from the dNTP-binding site (N site) to the priming site (P site) or release of the nucleic acid (distributive synthesis) (Fig. 1).

Extensive biochemical and crystallographic studies have enhanced our understanding of the details of the mechanism of DNA polymerization. However, the translocation step remains a poorly characterized step in DNA synthesis. This is primarily because there is limited biochemical and structural information on the pre-translocation intermediate that exists only transiently and is kinetically undetectable.

We present here a method that makes it possible to trap, purify, and crystallize this short-lived intermediate. The method is based on the cross-linking technique that the Harrison and Verdine laboratories used to obtain a stable covalent complex between RT, DNA, and dNTP. We have modified the protocol to extend/chain-terminate the template-primer with AZTMP. The nucleic acid is covalently trapped with the primer positioned at the N site (1). The covalent trapping prevents the primer from translocating to the P site. Covalent linking of the protein to DNA occurs by disulfide interchange between an engineered sulfhydryl group of the protein (Q258C mutation) and a disulfide-containing tether attached at the $N^2$ amino group of a modified dG in the primer strand of the nucleic acid. To synthesize the relatively large amounts of substituted oligonucleotides that are needed for biochemical, biophysical, and crystallographic studies, we developed novel and highly efficient procedures for preparing bis(3-aminopropyl)disulfide dihydrochloride and reacting it with a suitably protected 2-fluorodeoxyinosine residue in synthetic oligonucleotides (for sequences see Table I).

We have used this cross-linking strategy to probe the mechanism of nucleotide reverse transcriptase inhibitor (NRTI) re-

* This work was supported by National Institutes of Health (NIH) Grants AI 27690 (MERIT award) and GM 56609 (to E. A.) and by NIGMS, NIH and NCI, NIH (to S. H. H.). The costs of publication of this article were defrayed in part by the payment of page charges. This article must therefore be hereby marked "*advertisement*" in accordance with 18 U.S.C. Section 1734 solely to indicate this fact.

‖ To whom correspondence should be addressed. Tel.: 732-235-5323; Fax: 732-235-5788; E-mail: arnold@cabm.rutgers.edu.

[1] The abbreviations used are: HIV-1, human immunodeficiency syndrome, type 1; RT, reverse transcriptase; NRTI, nucleoside reverse transcriptase inhibitor; 3TC, 3′-thiacytidine; AZTTP, azidothymidine triphosphate; PMPA, 9-((*R*)-2-(phosphonomethoxy)propyl)adenine; HPLC, high performance liquid chromatography; IMAC, immobilized metal affinity chromatography; DTT, dithiothreitol; dd, dideoxy-; WT, wild-type; AZT, azidothymidine; AZTMP, azidothymidine monophosphate.

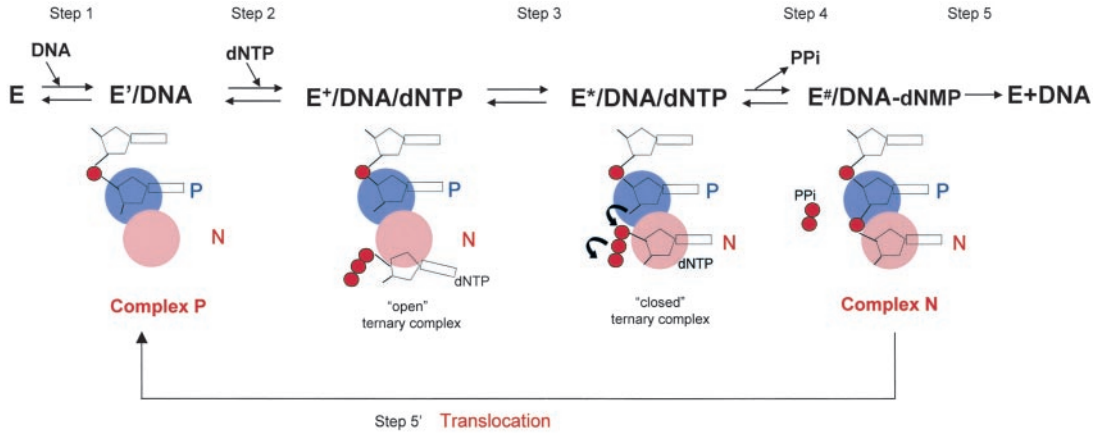



FIG. 1. **Schematic representation of the mechanism of DNA polymerization indicating the proposed occupancy of sites P ("priming" site in *blue*) and N ("nucleotide binding site" in *red*) at the 3′-end of the primer during the course of the reaction.** The novel non-translocated intermediate is the RT(N) (complex N), and the translocated intermediate is RT(P) (complex P). *Step 1*, binding of DNA to free enzyme E positions the 3′-primer end at the P site (Complex P). *Step 2*, binding of dNTP close to the N site to form "open" ternary complex. *Step 3*, formation of "closed" ternary complex. *Step 4*, phosphate bond formation accompanied by release of pyrophosphate to form the non-translocated complex RT(N). *Step 5*, in processive synthesis, translocation of primer from the N site to the P site. *Step 5′*, in non-processive synthesis, DNA falls off the enzyme.

sistance, which is based on enhanced excision. In this mechanism, AZT-resistant HIV-1 RT unblocks AZTMP-terminated primers by a mechanism that involves nucleophilic attack by a pyrophosphate donor (2), most likely ATP (3–6). We have recently proposed that several of the AZT-resistance mutations act primarily to facilitate the binding of ATP (3, 4, 7). We have also predicted that excision can occur if (and only if) the primer is positioned at the N site (3, 4, 7). Using the covalently trapped RT(P)·AZTMP and RT(N)·AZTMP complexes with the AZTMP primer terminus trapped at the P site or at the N site, we show that excision can occur only when the primer is positioned at the N site and not at the P site.[2] We also show that the product of ATP-based excision of AZTMP from RT(N)·AZTMP is a dinucleoside tetraphosphate. We present preparation and purification protocols that can be used to prepare dinucleoside tetraphosphates of various composition that could be used to probe the mechanism of ATP-based excision.

Finally, we established protocols for preparation, purification, and crystallization of the stable RT(P)·AZTMP and RT(N)·AZTMP complexes. Comparison of these crystal structures, which are published elsewhere (24), provides insights into the mechanism of translocation.

## MATERIALS AND METHODS

*Synthesis of Bis(3-aminopropyl)disulfide (N-tert-Butyloxycarbonyl)-3-aminopropyl Mesylate—tert*-Butyl-*N*-(3-hydroxypropyl) carbamate (700 mg, 4.0 mmol, 0.68 ml) from Aldrich Chemical Co. was dissolved in dichloromethane (20 ml), triethylamine (1.2 g, 12 mmol) was added, the solution was cooled to 0 °C, and methanesulfonyl chloride (500 mg, 4.1 mmol) was added slowly with stirring. The resultant mixture was stirred at 0 °C for 30 min and then stirred an additional 16 h at room temperature. Dichloromethane (50 ml) was added, and the resultant organic phase was washed with water and dried over magnesium sulfate. Concentration afforded the mesylate as a colorless oil (1.0 g, 99%). As previously described (8), the reactive mesylate was used without purification in the next step. $^1$H NMR (300 MHz, $CDCl_3$) δ 4.72 (s, br, 1H), 4.30 (t, 2H), 3.26 (q, 2H), 3.02 (s, $^3$H), 1.94 (q, 2H), 1.44 (s, 9H, Boc). MS (CI, $NH_3$) *m/z*: 252 (M-1), 174.

*Bis[(N-tert-butyloxycarbonyl)-3-aminopropyl]disulfide*—The general procedure of Ramesha and Chandrasekaran (9, 10) was used for preparation of the disulfide with the exception that the product was extracted with ethyl acetate rather than ether. Ammonium tetrathiomolybdate (1.1 g, 4.1 mmol) from Aldrich Chemical Co. was added to a solution of benzyltriethylammonium chloride (2.1 g, 8.9 mmol) in acetonitrile (15 ml), and the mixture was stirred for 1 h. The above mesylate (1 g, 4 mmol in 5 ml of acetonitrile) was added to the resultant red solution, and the mixture was stirred for 18 h at room temperature. The reaction mixture was concentrated to ~50% of its original volume. Ethyl acetate (25 ml) was added to the dark mass, and the resultant slurry was filtered through Celite. Acetonitrile (5 ml) was added (to dissolve partly the remaining black residue in the flask), followed by ethyl acetate (25 ml), and the resultant slurry was filtered through the same Celite bed. Extraction of the solids was repeated as before, and the combined extracts were concentrated to dryness under reduced pressure. The residue was re-dissolved in ethyl acetate (100 ml), and the solution was washed with water to remove residual benzyltriethylammonium chloride. Standard workup provided the desired disulfide as an oil (684 mg, 90% yield) that crystallizes on standing (mp: 82–83 °C). $^1$H NMR (300 MHz, $CDCl_3$), δ 4.68 (s, br, 2H), 3.24 (m, 4H), 2.72 (t, 4H), 1.88 (m, 4H), 1.45 (s, 18H). MS (FAB) *m/z*: 381 (M+1), 325, 225, 102, 90. High resolution mass spectrometry *m/z* calculated for $C_{16}H_{32}O_4N_2S_2$, 380.1804; found, 380.1882.

*Bis(3-aminopropyl)disulfide Dihydrochloride*—The above bis(*N-tert*-butyloxycarbonyl)-3-aminopropyl disulfide (0.68 g, 50 mmol) was dissolved in absolute ethyl alcohol (25 ml), and dry HCl gas was passed into the solution for 10 min while the temperature was maintained below 50–60 °C. Upon concentration of the ethanol, the dihydrochloride precipitated. The white solid was triturated with ether, filtered, dried, and stored in a vacuum desiccator (0.44 g, 97% yield). $^1$H NMR (300 MHz, $CD_3OD$), δ 3.10 (t, 4H), 2.80 (t, 4H), 2.10 (q, 4H). The bis(3-aminopropyl)disulfide and its dihydrochloride (11, 12) are known compounds.

*Oligonucleotide Synthesis*—Fluorinated oligonucleotides were prepared essentially as described previously (13) on a 10-μmol scale, utilizing the 4,4′dimethoxytrityl protected phosphoramidite derived either from $O^6$-(2-*p*-nitrophenethyl)-2-fluoro-2′-deoxyinosine (13, 14) or the commercially available 4,4′dimethoxytrityl phosphoramidite from $O^6$-(trimethylsilylethyl)-2-fluoro-2′-deoxyinosine (15) (ChemGenes, Ashland, MA) as the precursor of the dG residue to be modified. A support-bound 27-mer oligonucleotide (precursor of *27-Tem-C2 and -C3, Table I) containing an $O^6$-(2-*p*-nitrophenethyl)-2-fluoro-2′-deoxyinosine residue at the eleventh position from the 5′-end was allowed to react as described previously (13) with a 90-fold molar excess of either bis(3-aminopropyl)disulfide dihydrochloride (3-carbon tether) or commercially available cystamine dihydrochloride (2-carbon tether) in 500 μl of water and 300 μl of triethylamine for 16 h at room temperature, followed by deprotection (3 days, 60 °C) with concentrated $NH_4OH$ containing 0.05 M tetrabutylammonium fluoride (deblocks $O^6$). Oligonucleotides were purified by HPLC on a Hamilton PRP-1 column (7 μm,

[2] The terms RT(P) and RT(N) are used to describe complexes of RT with double-stranded DNA with the 3′-end of the primer positioned at the P, or N site respectively. When the primer is terminated by an NRTI, the name of the NRTI monophosphate is added at the end of the name of the complex. For example, the complex RT(N)·AZTMP is a complex with an AZTMP-terminated primer at the N site. Also, the complex RT(ter)-ddN·dTTP is a ternary complex of RT with a ddNMP-terminated primer at the P site and dTTP as the incoming dNTP positioned at the N site.

10 × 250 mm) eluted at room temperature (3 ml/min) with a linear gradient that increased the concentration of acetonitrile in 0.1 M $(NH_4)_2CO_3$ buffer (pH 7.5) from 0% to 17.5% over 20 min: $t_R$ values, 16.2 min (*27-Tem-C2) and 16.8 min (*27-Tem-C3). HPLC of the purified material (Phenomenex Luna C18 column, 4.6 × 250 mm, eluted at 50 °C (1.5 ml/min) with a linear gradient that increased the concentration of acetonitrile in the above carbonate buffer from 5% to 11% over 20 min): *27-Tem-C2, $t_R$ 14 min; *27-Tem-C3, $t_R$ 16.6 min. MS (electrospray): *27-Tem-C2, calc., 8460; found, 8461; *27-Tem-C3, calc., 8488; found, 8489.

Initially, coupling of bis(3-aminopropyl)disulfide (3-carbon tether) with the 20-mer oligonucleotide containing an $O^6$-(trimethylsilylethyl)-2-fluoro-2′-deoxyinosine residue at the fifth position from the 3′-end (precursor of *20-Pri-C3, Table I) was carried out in solution (15) as follows. The fluorinated 20-mer oligonucleotide was cleaved from the support with 0.1 M NaOH, and the resultant solution was neutralized with dilute acetic acid, evaporated to dryness, redissolved in $H_2O$ containing 0.2 M bis(3-aminopropyl)disulfide dihydrochloride and 0.75 M $Et_3N$, and allowed to react at room temperature for 3 days. At the end of this time, residual protecting groups were removed by heating at 60 °C for 16 h in a 1:1 ratio of concentrated $NH_4OH/H_2O$ containing 0.1 M tetrabutylammonium fluoride. Although DeCorte *et al.* (15) had observed more efficient coupling of 2-amino-2-phenylethanol with a fluorinated oligonucleotide after its cleavage from the support with NaOH as above, we subsequently observed that a separate NaOH cleavage step is not necessary with our less-hindered amine, because treatment (3 days at room temperature) of the glass beads with an aqueous solution of 0.9 M bis(3-aminopropyl)disulfide dihydrochloride and 4.5 M $Et_3N$ completely cleaves the oligonucleotide from the support concomitantly with replacement of the fluorine. The oligonucleotide, *20-Pri-C3, was purified by HPLC on the Hamilton column as above ($t_R$, 15 min). Although retention times varied depending on the amount of oligonucleotide and diamine injected, the desired oligonucleotide was well separated from multiple minor impurities, most of which eluted either much earlier or much later, as well as from bis(3-aminopropyl)disulfide ($t_R$, 13 min). Thus, preliminary dialysis (13, 14) to remove excess bis(3-aminopropyl)disulfide was unnecessary. The isolated yield of purified *20-Pri-C3 was 306 $A_{260\ nm}$ (~1.5 μmol). HPLC of the purified material on a Higgins DNA column (4.6 × 100 mm, Thomson Instrument Co., Clear Brook, VA) eluted at 40 °C (1.5 ml/min) with a linear gradient that increased the concentration of acetonitrile in the carbonate buffer described from 5% to 11% over 20 min, gave a single peak with a $t_R$ of 14.9 min. MS (electrospray): calc., 6232; found, 6233.

*Template-Primer Design*—All conventional DNA (non-cross-linkable) oligonucleotides were purchased from the DNA synthesis laboratory at the University of Medicine and Dentistry of New Jersey and were purified by HPLC. Oligonucleotide solutions (0.1 mM) containing 10 mM Tris/HCl, pH 7, and 50 mM NaCl were annealed by heating in a water bath for 5 min at 80 °C and allowed to cool slowly to ambient temperatures. The template-primer sequences are shown in Table I. The oligonucleotides are named as follows: The number indicates the length in nucleotides (*20-Pri is a 20-mer primer and *27-Tem is a 27-mer template). The asterisk indicates that the strand contains the modified G and can be cross-linked to the Cys-258. The suffix C3 or C2 on Tem or Pri indicates the length of the aliphatic chain that would link the protein with the enzyme. When we use the oligonucleotide sequence rather than the oligonucleotide name, the linker size is indicated by a subscript on the modified G (ex. $\mathbf{G}_3$ or $\mathbf{G}_2$). Finally, $(P)_A(N)_T$ indicates that a dAMP (or an analog of A) and/or a dTMP (or an analog of T) can be positioned at the P and N site, respectively, after cross-linking of the protein to the modified oligonucleotide.

*Preparation of Enzymes*—All RTs contained the C280S mutation in both p66 and p51 subunits to improve the stability of RT (16, 17) and minimize nonspecific cross-linking. The WT-1B1 enzyme also contained the Q258C mutation in both subunits. It was prepared by expressing p66 in *Escherichia coli* and allowing the conversion of p66 to p51 in *E. coli* extracts. The WT-1B1 enzyme was purified as described previously (18). The WT-Δ428 construct was prepared by expressing the p66 and p51 subunits from a plasmid that separately encodes p66 and p51. WT-Δ428 had the C terminus of p51 modified. The last HIV-1-encoded amino acid in p51 is 428, followed by two glycines and seven histidines. This allowed purification of RT-containing complexes using immobilized metal affinity chromatography. Introduction of the AZT-resistance mutations was performed in the WT-Δ428 construct in a subunit-specific manner as described previously (19). Hence, the AZT21-Δ428 AZT-resistant enzyme contained the C280S mutation in both p66 and p51 subunits, the Q258C mutation in the p66 subunit, and the M41L, D67N, K70R, T215Y, and K219Q mutations in the p66 subunit. The overexpressed WT-Δ428 and AZT21-Δ428 enzymes were purified by metal affinity chromatography.

*Cross-linking Reaction*—Cross-linking reactions of wild-type and mutant HIV-1 RTs contained 4-fold excess of DNA over enzyme in 10 mM Tris-HCl, pH 8, 150 mM NaCl, 8 mM $MgCl_2$. When needed, 1 mM dNTP and/or 1 mM AZTTP or other nucleoside reverse transcriptase inhibitor (NRTI) were added. The reaction was incubated for 3 h at 37 °C and then overnight at room temperature. The results were analyzed on non-reducing SDS-PAGE. Unless otherwise indicated, samples loaded for electrophoresis were not heated, because heating would cause a portion of the uncross-linked strand to separate from the cross-linked protein·DNA complex. Successful cross-linking was assessed by the appearance of a band corresponding to molecular weight of p66 plus DNA with a concomitant disappearance of the band corresponding to p66. The ratio of band intensities corresponding to cross-linked *versus* uncross-linked p66 subunits, ((p66·DNA)/(p66·DNA + p66)) × 100, was used to calculate the percent yield of the cross-linking reaction.

*Purification of the Complex*—The cross-linked complex was purified using two affinity chromatography procedures combined into one step with the columns connected in tandem. The reaction mixture was loaded onto the immobilized metal affinity chromatography (IMAC) column linked to the heparin column. Because RT contains a hexahistidine peptide at the C terminus of p51, both the uncross-linked p66/p51 heterodimer and the cross-linked p66·DNA·p51 complex bound to the IMAC column (first column), whereas the excess of template-primer, nucleotides, NRTIs, and other reaction components flowed through the columns. Application of 100 mM imidazole releases the free and cross-linked proteins from the IMAC column. The free RT bound to the heparin column, whereas the cross-linked RT (p66·DNA·p51) flowed through the heparin column. The peak fractions containing the cross-linked RT p66·DNA·p51 were pooled and concentrated using Millipore Ultrafree 10K spin columns at 4 °C to a final concentration of ~1 mg/ml. For crystallographic studies the RT p66·DNA·p51 was brought to a final concentration of 10–20 mg/ml. Final p66·DNA·p51 concentration was determined by measuring absorbance at 280 nm and by interpolation from a standard curve of known RT p66·DNA·p51 concentrations.

*Unblocking and Extension of Primer*—To determine whether the N and P complexes were competent for excision we first prepared AZTMP-blocked covalent complexes of AZT21-Δ428 with DNA that had AZTMP positioned at the P and N sites using the protocol described above. For the covalent complex with AZTMP at the N position (AZT21-Δ428(N)·AZTMP) we used the template-primer 6 (Table I) in the protocol described above with a reaction mixture that contained dATP and AZTTP. For the covalent complex P with AZTMP at the P position (AZT21-Δ428(P)·AZTMP), we used the template-primer 7 (Table I) in the protocol described above with a reaction mixture that contained AZTTP. The cross-linked complexes were purified as described above. Subsequently, the purified N and P complexes were subjected to ATP-based excision. Excision reactions were performed in 50 μl of 50 mM Tris/HCl (pH 8.0), 6 mM ATP, 120 mM NaCl, 8 mM $MgCl_2$, containing 25 μg of each covalent complex. The reaction mixtures were incubated for 4 h at 37 °C. The excess of unreacted ATP and any tetraphosphate products of the excision reaction were removed by centrifuging the samples in Millipore Ultrafree 10K spin columns. For the subsequent extension reaction we added [α-$^{32}$P]dTTP (final concentration 0.1 μM, 1 μCi) in 50 mM Tris/HCl (pH 8.0), 8 mM $MgCl_2$, and 120 mM NaCl to the reaction. After incubation for 10 min the concentration of dTTP was adjusted to 2 μM. The reaction was incubated for an additional hour. The products were fractionated on an 11% SDS-PAGE and analyzed by autoradiography with photographic film.

*Enzymatic Synthesis of the Dinucleoside Tetraphosphate Product of Excision*—Complex N of AZT-resistant RT (AZT21-Δ428) with an AZTMP-terminated primer (template-primer 6) was prepared as described in the previous paragraph using dGTP and AZTTP for *in situ* primer extension and blocking. The bulk of the unreacted dGTP and AZTTP was removed by gel filtration through a G-50 column. The flow-through was adjusted to contain 50 mM Tris-HCl (pH 8.0), 6 mM ATP, 120 mM NaCl, 8 mM $MgCl_2$ and allowed to react for at least 4 h at 37 °C. The resulting RT·DNA complex (complex P) was removed by passing the reaction mixture through a nickel-charged IMAC column; the complex can be recovered for further synthesis of a tetraphosphate or other uses. The protein-free flow-through fraction was passed again through a G-50 column to remove most of the unreacted ATP. Final purification of AZTMP-ATP was by HPLC on a Higgins DNA column (10 × 250 mm) at 1 ml/min flow rate at room temperature with a linear gradient of 0% solvent B to 100% B over 30 min (solvent A, 0.1%

TABLE I
*Template-primers used in cross-linking studies*
Naming of the oligonucleotides is described under "Materials and Methods" section (template-primer design).

| | |
|---|---|
| 1) *27-Tem-C3 | 5′-ATG CAT CGG C**G**$_3$C TCG AAC AGG GAC TGT-3′ |
| 18-Pri | C GC G AGC TTG TCC CTG AC -5′ |
| 2) *27-Tem-C3 | 5′-ATG CAT CGG C**G**$_3$C TCG AAC AGG GAC TGT-3′ |
| 20-Pri | 3′- GCC GC G AGC TTG TCC CTG AC -5′ |
| 3) *27-Tem-C2 | 5′-ATG CAT CGG C**G**$_2$C TCG AAC AGG GAC TGT-3′ |
| 20-Pri | 3′- GCC GC G AGC TTG TCC CTG AC -5′ |
| 4) 27-Tem(P)$_A$ (N)$_T$ | 5′-ATG CAT GGC GC C CGA ACA GGG ACT GTG-3′ |
| *20-Pri-C3 | 3′- CCG C**G**$_3$G GCT TGT CCC TGA CA -5′ |
| 5) 27-Tem (P)$_T$ (N)$_A$ | 5′-ATG CTA GGC GC C CGA ACA GGG ACT GTG-3′ |
| *20-Pri-C3 | 3′- CCG C**G**$_3$G GCT TGT CCC TGA CA -5′ |
| 6) 27-Tem (P)$_G$ (N)$_T$ | 5′-ATG CAC GGC GC C CGA ACA GGG ACT GTG-3′ |
| *20-Pri-C3 | 3′- CCG C**G**$_3$G GCT TGT CCC TGA CA -5′ |
| 7) 27-Tem (P)$_T$ (N)$_C$ | 5′-ATG CGA GGC GC C CGA ACA GGG ACT GTG-3′ |
| *20-Pri-C3 | 3′- CCG C**G**$_3$G GCT TGT CCC TGA CA -5′ |

trifluoroacetic acid in $H_2O$; solvent B, 0.1% trifluoroacetic acid in acetonitrile).

*Crystallization of P and N Complexes*—We used two protocols to crystallize the P and N complexes. The first protocol included a monoclonal antibody fragment (Fab28) that we have used previously to crystallize non-cross-linked complexes of HIV-1 RT with nucleic acid (20–22). Specifically, the peak fractions containing the p66·DNA·p51 fractions were pooled, concentrated to 20–25 mg/ml, and mixed with Fab28 at 1:0.8 mass ratio (23). Hanging drops were prepared by mixing equal volumes of the complex and crystallization solutions (100 mM cacodylate, pH 5.6, 31–34% saturated ammonium sulfate) at 4 °C. The second protocol was essentially that used previously to crystallize the catalytic complex of cross-linked RT·DNA with incoming dNTP (1): The p66·DNA·p51 fractions were concentrated to ~8 mg/ml and mixed with equal volumes of 50 mM bis-tris propane at pH 6.4, 10–12%, polyethylene glycol 8000 (w/v), 100 mM ammonium sulfate, 5% sucrose (w/v), and 5% glycerol (w/v).

## RESULTS

*Cross-linking Reaction and Purification of Complexes*—In our initial cross-linking experiments we used an enzyme that had the Q258C mutation in both subunits (WT-1B1). The enzyme cross-linked to the template strand of *27-Tem-C3/20-Pri (template-primer 2 in Table I) at yields comparable to those reported previously (1). However, the resulting cross-linked complex was unstable. Anion exchange chromatography (mono Q 10/10 fast-protein liquid chromatography column) of the reaction mixture resulted in the dissociation of p66·DNA from the p51 subunit. Further dissociation was observed during storage (Fig. 2). The uncross-linked enzyme was stable under these conditions (not shown).

To obtain more stable complexes we prepared an expression plasmid that could lead to the synthesis of an HIV-1 RT that has the Q258C mutation only in the p66 subunit (WT-Δ428). To facilitate the purification, the C terminus of p51 was modified to contain a stretch of seven histidines that simplifies the purification through an immobilized metal affinity chromatography step. The WT-Δ428 RT showed no apparent loss of enzymatic activity with respect to the wild-type enzyme. Furthermore, the cross-linked complex of WT-Δ428 with DNA was stable. The tandem chromatography protocol, described under "Materials and Methods," resulted in an efficient one-step purification of the cross-linked complex. A typical chromatogram for the purification of RT(N)·AZTMP (prepared using template-primer 4, dATP, and AZTTP) from the cross-linking reaction mixture is shown in Fig. 3*A*. Similar results were obtained for the purification of RT(P)·AZTMP (prepared using template-primer 5 and AZTTP), RT(P)-PMPA (prepared using template-primer 4 and PMPA diphosphate), and RT(N)-PMPA (prepared using template-primer 5 with dTTP and PMPA diphosphate).

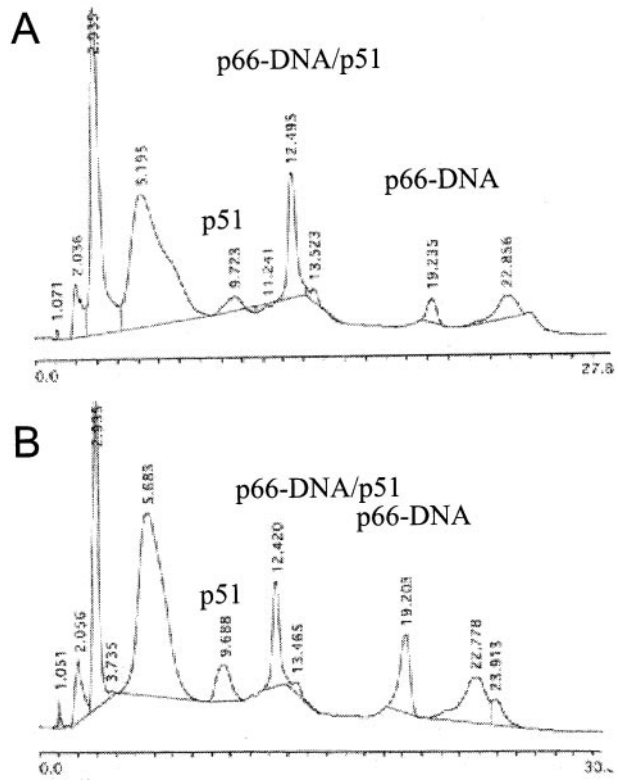



FIG. 2. **Chromatograms of cross-linking reaction mixtures that were loaded onto a Mono Q fast-protein liquid chromatography 10/10 column and eluted using a linear salt gradient from 50 mM NaCl to 1.0 M NaCl in 30 min (buffers were 40 mM Tris, pH 6.5, flow rate 0.5 ml/min).** *A*, mixture injected immediately after end of reaction. The ratio of the dissociated p66·DNA (without the p51 subunit) to the intact product peak p66·DNA·p51 was 0.2. *B*, the dissociation increased dramatically after a freeze-thaw cycle.

The cross-linking protocol was similar to the protocol used by the Verdine and Harrison laboratories to prepare the closed RT(ter)-ddNMP·dTTP complex (1). Linking of the protein to DNA occurred by disulfide interchange when the engineered Cys-258 in the αE helix of the p66 thumb of RT and the disulfide-containing tether attached at the $N^2$ amino group-modified dGMP in the primer (or in the template) were aligned

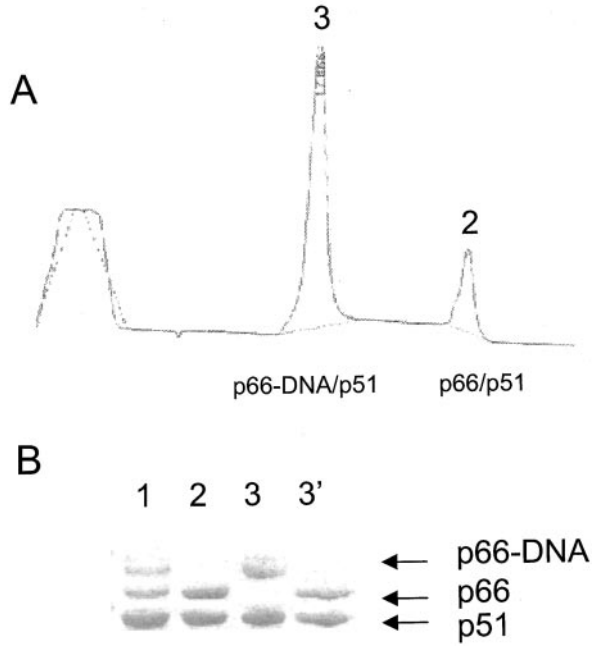
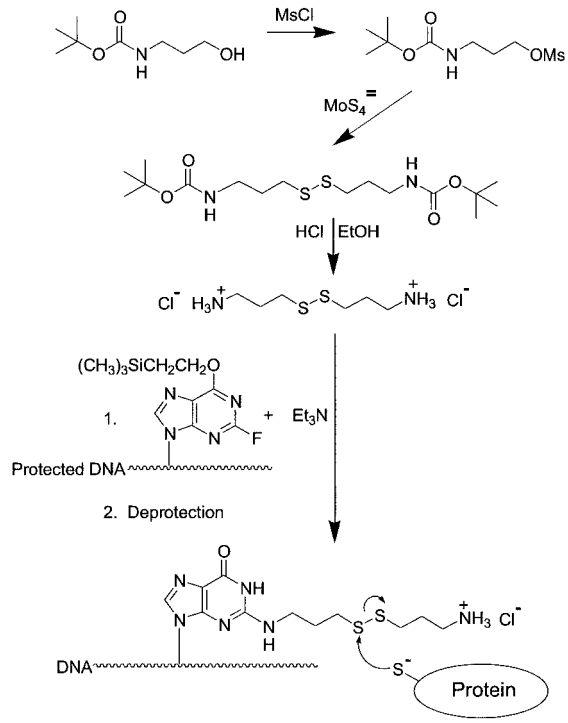



FIG. 3. **Purification of RT(N)·AZTMP and SDS-PAGE analysis.** *A*, chromatogram from the purification of a reaction mixture using an immobilized metal affinity chromatography column and a heparin column. *Peak 3* corresponds to the cross-linked product (p66-RT/p51) and *peak 2* to unreacted RT (p66/p51). *B*, a covalent complex of WT-Δ428 with 27-Tem-$(P)_A(N)_T$/*20-Pri was incubated with dATP and AZTTP as described under "Materials and Methods." The AZTMP-terminated primer was cross-linked and positioned at the N site of RT. The *lanes* shown in *B* are as follows: *lane 1*, cross-linking reaction mixture before purification; *lane 2*, pooled peak fraction containing uncross-linked p66 and p51 subunits (*peak 2* of chromatogram in *A*); *lane 3*, pooled peak fractions containing p66 cross-linked to the template-primer and p51 subunits (*peak 3* of chromatogram in *A*); *lane 3′* is identical to *lane 3* except that the material was incubated in the presence of 2 mM β-mercaptoethanol to cleave the covalent -S-S- bond between p66 and the primer strand.




SCHEME I

and in close proximity (Scheme I and Fig. 4). SDS-PAGE analysis of the cross-linking reaction mixture under non-reducing conditions shows a band that migrates more slowly than p66, accompanied by reduced intensity of the p66 band (Fig. 3*B*). The shifted band corresponds to p66 covalently bound to the modified DNA through a disulfide bond (p66·DNA). When 4 mM DTT was added to the loaded sample the p66·DNA band disappeared, because DTT cleaved the disulfide linkage between DNA and p66. Under the denaturing conditions of SDS-PAGE gel, the DNA is released (Fig. 3*B*, *lanes 3* and *3′*), and the p66 band reappears with similar intensity as the p51 band. Hence, appearance of this p66·DNA band can be used effectively to evaluate the extent of cross-linking for a given complex.

We first determined the optimal distance in nucleotides between the cross-linking site and the polymerase active site (P and N sites). We used as a starting template-primer the *27-Tem-C3/18-Pri (template-primer 1 in Table I) with the modified G at the third base pair from the P site (the 3′-end of the primer is assumed to be positioned at the P site). We performed separate polymerization reactions that were blocked by incorporation of a ddNMP after a defined number of reaction cycles (one, two, and three cycles). The polymerization cycles extended the 18-Pri by one, two, and three nucleotides resulting in complexes between dNTP, ddNMP-terminated DNA, and RT that had the base pair with the modified G at the fourth, fifth, and sixth positions from the P site.

Fig. 5*A* (*lanes 1–3*) shows that formation of a complex between RT, dNTP, and a primer (19-mer or 20-mer) that has the modified G at the fourth or fifth position from the P site (one and two cycles of polymerization, respectively) does not generate any cross-linking, as judged by the absence of a p66·DNA band in a non-reducing SDS-PAGE gel. After three cycles of polymerization, however, the primer has become a 21-mer and the modified G is at the sixth position from the P site. As shown in Fig. 4*A* (*lane 3*) when the G is in the sixth position about 50% of the p66 was cross-linked to DNA, as judged from the relative intensity of the bands corresponding to p66 and p66·DNA. This suggests that Cys-258 of p66 and the modified G are aligned optimally for disulfide interchange when the modified G is in the sixth position from the P site. The optimal distance is the same when the modified G is on the template strand (Fig. 4) or in the primer strand (not shown).

We also asked whether the cross-linking efficiency is affected by the presence or absence of the incoming dNTP at the N site. We evaluated the efficiency of cross-linking between the enzyme and template-primers extended by a defined number of cycles as described above but in the absence of an incoming dNTP. In the absence of the incoming dNTP the cross-linking yield is very low, even when Cys-258 of p66 and the modified G are aligned optimally for disulfide interchange with the enzyme (three cycles of polymerization on the 18-Pri) (Fig. 5*B*, *lanes 4–6*). To minimize accumulation of byproducts and simplify the system, we increased the length of the primer strand to 20 nucleotides (*27-Tem-C3/20-Pri) (template-primer 2 in Table I) and placed the modified G at the fifth base pair from the 3′-end of the primer. This allowed preparation of P complexes by a single polymerization cycle. Similarly, the RT(N)·AZTMP complex was prepared by *in situ* incorporation of one dNTP and AZTMP.

To evaluate the effect of the length of the linker on the cross-linking yield and the stability of the covalent complex, we cross-linked WT-Δ428 and template-primers with tethers of different sizes (*27-Tem-C3/20-Pri, three-methylene linker; *27-Tem-C2/20-Pri, two-methylene linker) (template-primers 2 and 3, Table I). The cross-linking yields and protein stabilities of the two covalent complexes were comparable (not shown).

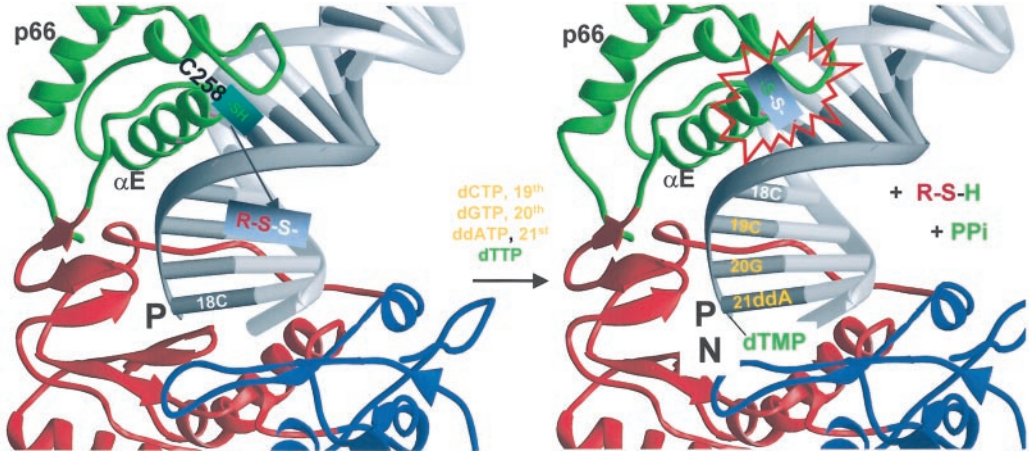



FIG. 4. ***In situ* cross-linking *via* disulfide interchange between Cys-258 of the p66 thumb (in *green*) of RT and the disulfide-containing tether attached at the $N^2$ amino group of the modified dGMP.** In this case the cross-linkable strand is the template strand (*27-Tem-C3, see Table I) (in *light gray*); the annealed primer is 18-Pri (see Table I) (in *dark gray*). *In situ* addition of dCMP, dGMP, and ddAMP extends the primer to 21 nucleotides. This advances the nucleic acid bringing the sulfur-containing G into close proximity and favorable alignment with Cys-258 of the p66 thumb for the disulfide interchange reaction (*R* is $-CH_2-CH_2-CH_2-NH_2$). For efficient cross-linking, the N site should be occupied (in this case by dTTP or AZTTP). The original primer (18-mer) contains the modified G at the third base pair from the P site. After primer extension the modified G is at the sixth base pair from the P site (or at the seventh from the N site), which leads to efficient cross-linking. For preparation of an AZTMP-terminated N complex with this template-primer ddATP was replaced by dATP and dTTP with AZTTP. The palm and fingers subdomains of p66 of RT are shown in *red* and *blue*, respectively.

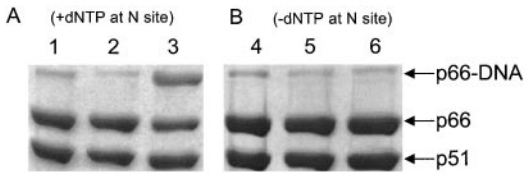



FIG. 5. **Preparation of P site complex in the presence (*A*) or absence (*B*) of dNTP at the N site.** All reactions contained *27-Tem-C3/18-Pri as the starting template-primer. The modified G was at the third base pair (on the template strand) from the 3′-primer end, which is positioned the P site. The reactions also contained: 1) ddCTP and dGTP at the N site; 2) dCTP, ddGTP, and dATP at the N site; 3) dCTP, dGTP, ddATP, and dTTP at the N site; 4) ddCTP; 5) dCTP and ddGTP; and 6) dCTP, dGTP, and ddATP. The reactions extend the 18-mer primer to a 19-mer (*lanes 1* and *4*), 20-mer (*lanes 2* and *5*), and 21-mer (*lanes 3* and *6*) changing the distance between the P site (3′ primer end) and the modified G to four, five, and six nucleotides in *lanes 1* and *4*, *lanes 2* and *5*, and *lanes 3* and *6*, respectively. The P site complex is formed only when the distance between the primer terminus and the Q258C is six base pairs, and the incoming dNTP is present (*lane 3*).

These RT(P) complexes with C2 and C3 linkers were crystallized, their structures solved, and preliminary results show no substantial differences at the polymerase active sites of the two complexes.[3]

*ATP-based Excision Reaction with P and N Complexes as Substrates*—We measured the enzymatic activities of purified, stable, cross-linked N and P complexes. We used the AZT-resistant enzyme (AZT-21-Δ428), because the AZT resistance mutations do not substantially affect the polymerase activity and increase the ATP-based excision activity by improving the binding of ATP (3). Starting with 27-Tem-$(P)_T(N)_A$/*20-Pri-C3 or 27-Tem-$(P)_A(N)_T$/*20-Pri-C3 and AZT-21-Δ428, we prepared covalent complexes of the enzyme with AZTMP-terminated primers positioned at the P site (AZT-21(P)·AZTMP) and at the N site (AZT-21(N)·AZTMP). After the excision reaction was performed unreacted ATP and any dinucleoside tetraphosphate was removed. The purified complexes were subjected to ATP-based excision reaction conditions. A control experiment showed that the primer was quantitatively chain-terminated by AZTMP in both the P and N complexes. The AZTMP-terminated N and P complexes were treated with 5 mM DTT to release the cross-linked nucleic acid, and the DTT-treated complexes were incubated with [$\alpha$-$^{32}$P]dTTP, as described under "Materials and Methods." None of the reduced complexes were able to incorporate [$^{32}$P]dTMP, suggesting that essentially all of the primers were terminated with AZTMP (not shown). Extension reactions showed that only the product derived from performing the excision reaction from the N complex was able to incorporate labeled dTMP (Fig. 6, *lane 2 versus lane 1*). The products derived from the P complex were not able to incorporate labeled dTMP, even when they were treated with DTT after the ATP-based excision reaction but before the extension reaction.

*Preparation and Purification of the Tetraphosphate Product of Excision*—The large excess of ATP was separated efficiently from the tetraphosphate product by the G-50 gel filtration step that retained most of the ATP. Any residual ATP was removed in an HPLC step where ATP was in the flow-through, and the dinucleoside tetraphosphate product (Ap4AZT) was bound to the column (retention time, 11.2 min). The retention time of a related Ap4A tetraphosphate (Sigma) was 10.4 min under these conditions. Additional compounds present were separated from the dinucleoside tetraphosphate product during the final purification step: a compound with retention time identical to that of standard AZTTP (13.1 min) and a compound that migrated with retention time 12 min (Fig. 7), which was present as an impurity in the preparation of ATP used for the excision reaction.

*Crystallization of N and P Complexes*—We used the cross-linking technique to prepare several N and P complexes of wild-type and drug-resistant RTs (WT-Δ428, AZT-21-Δ428, and M184V-Δ428) with template-primers terminated with AZTMP, PMPA, or 3TC. The yields of the cross-linking reactions (50–90%) allowed preparation of milligram amounts of complexes. The following complexes have been purified and crystallized: WT (P)-ddA (with C2 and C3 linkers), WT(N)-dCMP, WT(P)·AZTMP, WT(N)·AZTMP, WT(P)-3TCMP, WT(P)-PMPA, WT(N)-PMPA, AZT-21(P)·AZTMP, AZT-21(N)·AZTMP, M184V(P)·AZTMP, and M184V(N)·AZTMP. The crystal structures of the post-incorporation/pre-translocation (WT(N)·AZTMP) and post-translocation (WT(P)·AZTMP) complexes of

[3] S. Tuske, S. G. Sarafianos, C. J. Squire, A. D. Clark, Jr., J. Ding, P. Ilankumaran, A. R. Ramesha, H. Kroth, J. M. Sayer, D. M. Jerina, P. Boyer, S. H. Hughes, and E. Arnold, unpublished observations.

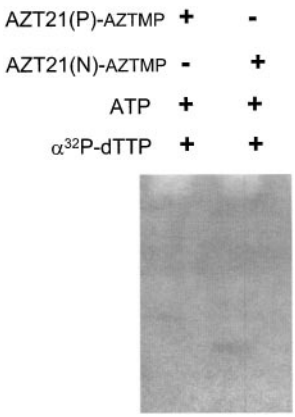



FIG. 6. **Excision of the primer terminus when the end of the primer is bound at the N, but not at the P site.** Covalent complexes of AZT-resistant RT (AZT-21) with AZTMP-terminated primers positioned at the P (AZT-21(P)·AZTMP) or at the N site (AZT-21(N)·AZTMP) were prepared as described under "Materials and Methods." The complexes were purified by IMAC and heparin chromatography and subjected to ATP-based excision reaction conditions. After completion of the excision reaction the excess ATP was removed by size exclusion chromatography, and the RT·DNA complexes were tested for their ability to extend their primers and incorporate radiolabeled [$^{32}$P]dTMP.

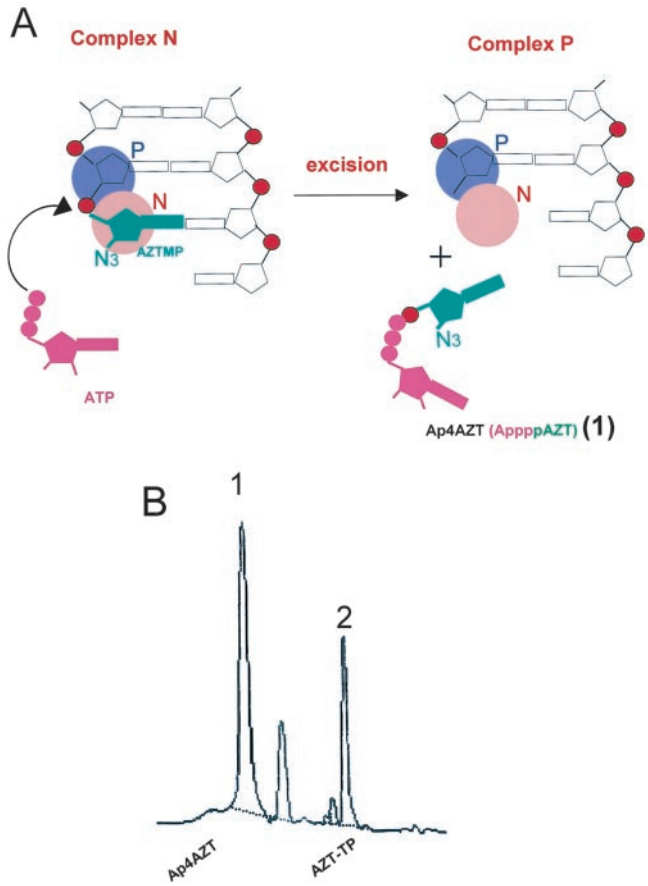



FIG. 7. *A*, reaction scheme of ATP-based excision of AZTMP from RT(N)·AZTMP. The reaction involves a nucleophilic attack of the gamma phosphate of ATP on the AZTMP phosphate at the 3′ primer terminus. The ATP and AZTMP are linked through a 5′-5′ bond to form the tetraphosphate product. *B*, HPLC purification of the tetraphosphate product (Ap4AZT) of ATP-based excision. RT was removed by passing the reaction mixture through a nickel-charged IMAC column. The protein-free flow-through fraction was then passed through a G-50 gel filtration column to remove the bulk of unreacted ATP. Final purification was on an HPLC Higgins DNA column (10 × 250 mm) as described under "Materials and Methods." Retention times are as follows: ATP, 0.5 min (not shown); Ap4AZT, 11.2 min (*peak 1*); AZTTP, 13.1 min (*peak 2*). An unknown contaminant of ATP migrates with a retention time of 12 min (*peak* between *Ap4AZT* and *AZTTP*).

HIV-1 RT cross-linked to AZTMP-terminated primer-templates have been solved and described elsewhere (24). Representative portions of the WT(N)·AZTMP complex electron density maps at 3.0-Å resolution are shown in Fig. 8*A*. The geometry and conformation of the tether explains why both C2 and C3 linkers can cross-link between RT and DNA without straining other protein-nucleic acid interactions. Cross-linking of DNA to RT allowed trapping of the AZTMP-terminated primer terminus at the N site in an elusive pre-translocated conformation that has been so far invisible by spectroscopic and biochemical methods (Fig. 8*B*).

## DISCUSSION

We describe a method to trap intermediates of the polymerization reaction in a stable form that can be used for biochemical, biophysical, and crystallographic studies. The intermediates are pre- and post-translocation complexes of HIV-1 reverse transcriptase with a DNA oligonucleotide that can be terminated with AZTMP or other NRTIs.

The methodology used was similar to the protocol developed by the Verdine and Harrison laboratories to prepare the closed RT(ter)-ddNMP·dTTP complex (1). By changing the composition of the reactants we were able to incorporate an additional nucleotide, generating the N complex. We modified the protocol so that the cross-link is with the primer rather than with the template strand. Cross-linking to the primer strand allows the use of easily available un-modified templates that can be changed to determine the type of the incoming dNTP or dNTP analog. The template sequence can be changed readily at the active site to allow experiments with AZT, 3TC, PMPA, any dNTP or NRTI using a single modified oligonucleotide (the primer).

To optimize the cross-linking protocol we synthesized several templates and primers that contained a dG residue with a disulfide linker at the exocyclic $N^2$ amino group. These were synthesized from suitably protected, fluorinated oligonucleotides by reaction with bis(3-aminopropyl)disulfide. We have developed a procedure for preparing this precursor in excellent yield (90%) and high purity. This method utilizes benzyltriethylammonium tetrathiomolybdate (9, 10) to introduce the disulfide linkage and a *tert*-butyloxycarbonyl group (8) to protect the amine. The *tert*-butyloxycarbonyl group is easily removed to give a product that requires no further purification. Previous methodology (13) afforded this material only in lower yield (~20%). Furthermore, the earlier procedure (25) was undesirable, because it involved the use of hydrazine to remove a phthalimido-protecting group. Potential hydrazine contamination of the bis(3-aminopropyl)disulfide product is difficult both to detect and to remove and likely competes favorably with the primary amine for substitution on the fluorinated purine, resulting in low yields of the desired substituted oligonucleotide. Use of bis(3-aminopropyl)disulfide dihydrochloride prepared by the present method gave substituted oligonucleotides with a superior yield and purity as compared with previous preparations. Complete purification was easily achieved in a single HPLC step.

Cross-linking of DNA to an RT that has the Q258C mutation in both subunits decreased the stability of the complex and resulted in dissociation of the p51 subunit from the cross-linked p66·DNA complex. This is probably caused by the disruption of the interactions of the Gln-258 side chain on the p51 subunit that normally resides in a hydrophobic pocket at the interface between the RNase H subdomain of p66 and the thumb subdomain of p51 (van der Waals interactions with Val-536, Tyr-441, Val-496, of p66 and with Val-254, Leu-283, and Leu-289 of

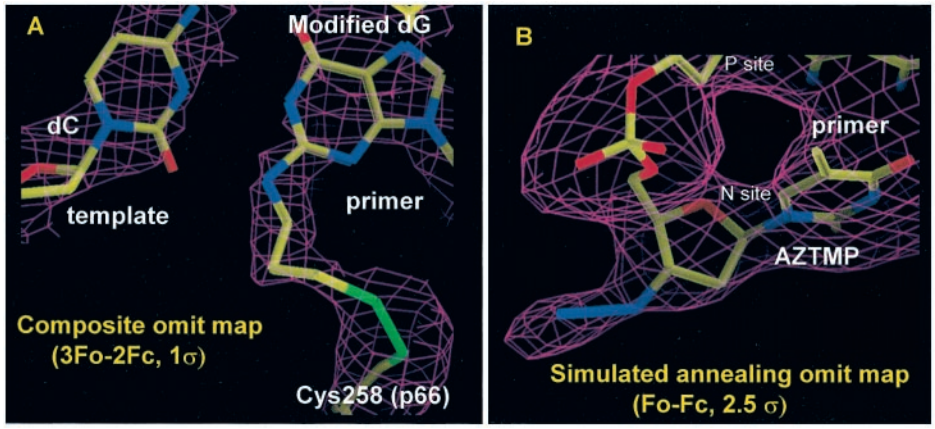



FIG. 8. **Electron density maps at 3.0-Å resolution from the crystal structure of the pre-translocated (N) complex of HIV-1 RT·DNA.** The primer strand has been terminated at the 3′-end with AZTMP and cross-linked to HIV-1 RT as described under "Materials and Methods." Two regions of the structure are shown: *A*, the region surrounding the tethered CG* base pair in the template-primer and residue Cys-258 in helix H of RT. The coordinates for the tether were not included in the calculation of the composite omit $3F_o - 2F_c$ map (contoured at 1 $\sigma$). *B*, the dNTP-binding site (N site) is occupied by the 3′-end of the primer, which has been elongated by AZTMP, but not translocated to the P site because of the covalent cross-link between primer and RT. The coordinates of AZTMP and of residues in a 6-Å radius around AZTMP were not included in the calculation of the simulated annealing $F_o - F_c$ omit map (contoured at 2.5 $\sigma$). The structural details of the N complex are described elsewhere (PDB code 1N6Q) (24).

p51). Cross-linking of an RT carrying the Q258C mutation only in the p66 subunit to modified DNA produced stable complexes and was used in all of our studies.

The cross-linking efficiency of RT to DNA was unaffected by changes in the length of the linker. RT cross-links to nucleic acids with either C2 or C3 tethers with the same efficiency suggesting that there are no substantial changes in the binding of the nucleic acid. This is consistent with preliminary structural results from the crystal structures of RT(P)ddA (primer linked to RT with a C2 tether) and of R(T(P)ddA (primer linked to RT with a C3 tether)[2] that show no changes in the binding of the nucleic acid. These results suggest that the tether does not constrain the binding of nucleic acid in a biologically irrelevant conformation. Instead, the cross-linking reaction is likely to be catalytically relevant, because it takes place *in situ*, whereas the enzyme is polymerizing, and the nucleic acid is bound to the enzyme in a productive mode.

We determined the optimal distance in nucleotides between the cross-linking site and the dNTP-binding site by placing the modified G at the fourth, fifth, or sixth base pair from the P site and assessing the efficiency of cross-linking. Cross-linking is optimal when the distance is six nucleotides, suggesting a requirement for precise positioning of the nucleic acid. This is in agreement with the results of the previous cross-linking experiments (1) with a similar system. We also show that cross-linking is equally efficient when the modified G is in the template or in the primer oligonucleotide at the sixth base pair from the P site. When the N site is occupied the cross-linking yield was substantially improved. This suggests that some minor adjustments in the position of the p66 thumb may accompany formation of a productive catalytic complex. Additional factors that we did not examine in this study, such as sequence, type, and structure of nucleic acid, may also affect the cross-linking efficiency. This is an alternative mechanism that would explain the requirement for a dNTP: In the absence of a dNTP, the oligonucleotide can equilibrate between orientations with the 3′ primer terminus at either the N or the P site, whereas in the presence of a dNTP the 3′ terminus is blocked from the N site, restricting the oligonucleotide to a single orientation that is optimal for cross-linking. This may provide a high local concentration of the disulfide linkage in the vicinity of Cys-258 and enhance the cross-linking reaction.

We predicted that excision occurs if, and only if, the 3′-OH primer end is positioned at the N site (3). In the present study we provide experimental evidence to support this proposal. Using stable N site and P site complexes we showed that the excision reaction can take place when the 3′-end of the primer is bound at the N, but not at the P site. The stable N site complex can be used to study other aspects of the excision mechanism. The main problems in studying the ATP-based excision reaction are: 1) the primary substrate for the excision reaction, the N complex, is not stable because of translocation of the nucleic acid and 2) the secondary substrate for the excision reaction, ATP, can bind either as a substrate for the excision reaction in a pocket formed by amino acid residues involved in resistance to AZT, or ATP can bind at the N site as a substrate for the polymerization reaction. The $K_m$ for incorporation of rNTP by a related RT (murine leukemia virus) is ~0.4 mM (26). HIV-1 RT has also been shown to incorporate rNTPs under similar conditions (27). Because the excision reaction requires ~3 mM ATP, a high concentration of rNTP can compete with the normal dNTPs for binding at the N site (3). These factors complicate the study of the mechanism of ATP-based excision. However, use of covalently trapped N complex prevents translocation of the nucleic acid and blocks the binding of ATP at the N site, which should facilitate studies of the mechanism of excision-based NRTI resistance.

The cross-linked N site complexes can be used for the enzymatic synthesis of pure dinucleoside tetraphosphate products. Reaction of ATP with RT and AZTMP-terminated nucleic acid yields not only Ap4AZT but also a mixture of other dinucleoside tetraphosphate products (Ap4dT, Ap4dA, Ap4dC, Ap4dG) that result from sequential excision steps at the 3′ primer terminus. It is difficult to purify Ap4AZT from other tetraphosphates with similar chromatographic properties. Using cross-linked complex N for the excision reaction yields only one dinucleoside tetraphosphate because the template-primer cannot translocate after excision of the nucleotide from the 3′-end of the primer. We used this method to prepare Ap4AZT free of ATP, AZTTP, and other dinucleoside tetraphosphate. Other tetraphosphate products consisting of combinations of various NTPs and various NRTIs are currently being prepared using the same primer, cross-linked to RT, and appropriate templates that allow incorporation of the respective NRTIs.

The covalently cross-linked polymerization intermediates are also useful in probing the mechanism of translocation of the

enzyme on the nucleic acid. Kinetic, biochemical, and structural studies on translocation have been hampered in the past, because the unstable pre-translocation intermediate could not be isolated. Preparation of stable pre- and post-translocation complexes has allowed structural studies with these intermediates, the results of which will be published elsewhere (24). Comparison of the structures of pre- and post-translocated intermediates provides the first glimpses into the molecular basis of translocation. The same general strategy can also be employed to probe the molecular mechanism of the excision reaction by studying the structure of RT in complex with NRTI-terminated inhibitor at the N site and incoming ATP. It should be possible to prepare reaction intermediates of other DNA or RNA polymerases and with related enzymes that perform phosphotransferase reactions, such as integrases, endonucleases, exonucleases, ligases, recombinases, topoisomerases, and others.

In conclusion, we adapted a method of cross-linking of a polymerase to a nucleic acid substrate to isolate a novel intermediate of the polymerization reaction in a stable form amenable to structural and biochemical studies. This method allows us to address several aspects of the mechanism of DNA polymerization, translocation, and excision-based resistance of RT to NRTIs.

*Acknowledgments*—We thank other Arnold laboratory members for helpful discussions and Pat Clark for continuing assistance in purification of recombinant HIV-1 RT.

REFERENCES

1. Huang, H., Chopra, R., Verdine, G. L., and Harrison, S. C. (1998) *Science* **282,** 1669–1675
2. Arion, D., Kaushik, N., McCormick, S., Borkow, G., and Parniak, M. A. (1998) *Biochemistry* **37,** 15908–15917
3. Boyer, P. L., Sarafianos, S. G., Arnold, E., and Hughes, S. H. (2001) *J. Virol.* **75,** 4832–4842
4. Boyer, P. L., Sarafianos, S. G., Arnold, E., and Hughes, S. H. (2002) *J. Virol.* **76,** 9143–9151
5. Meyer, P. R., Matsuura, S. E., So, A. G., and Scott, W. A. (1998) *Proc. Natl. Acad. Sci. U. S. A.* **95,** 13471–13476
6. Meyer, P. R., Matsuura, S. E., Tolun, A. A., Pfeifer, I., So, A. G., Mellors, J. W., and Scott, W. A. (2002) *Antimicrob. Agents Chemother.* **46,** 1540–1545
7. Boyer, P. L., Sarafianos, S. G., Arnold, E., and Hughes, S. H. (2002) *J. Virol.* **76,** 3248–3256
8. Mattingly, P. G. (1990) *Synthesis*, 366–368
9. Ramesha, A. R., and Chandrasekaran, S. (1992) *Synth. Commun.* **22,** 3277–3284
10. Ramesha, A. R., and Chandrasekaran, S. (1994) *J. Org. Chem.* **59,** 1354–1357
11. Doi, J. T., and Musker, W. K. (1985) *J. Org. Chem.* **50,** 1–4
12. Evans, B., Doi, J. T., and Musker, W. K. (1990) *J. Org. Chem.* **55,** 5280–5286
13. Peletskaya, E. N., Boyer, P. L., Kogon, A. A., Clark, P., Kroth, H., Sayer, J. M., Jerina, D. M., and Hughes, S. H. (2001) *J. Virol.* **75,** 9435–9445
14. Erlanson, D. A., Chen, L., and Verdine, G. L. (1993) *J. Am. Chem. Soc.* **115,** 12583–12584
15. DeCorte, B. L., Tsarouhtsis, D., Kuchimanchi, S., Cooper, M. D., Horton, P., Harris, C. M., and Harris, T. M. (1996) *Chem. Res. Toxicol.* **9,** 630–637
16. Jacobo-Molina, A., Clark, A. D., Jr., Williams, R. L., Nanni, R. G., Clark, P., Ferris, A. L., Hughes, S. H., and Arnold, E. (1991) *Proc. Natl. Acad. Sci. U. S. A.* **88,** 10895–10899
17. Boyer, P. L., Ferris, A. L., and Hughes, S. H. (1992) *J. Virol.* **66,** 1031–1039
18. Clark, A. D. J., Jacobo-Molina, A., Clark, P., Hughes, S. H., and Arnold, E. (1995) *Methods Enzymol.* **262,** 171–185
19. Boyer, P. L., Ferris, A. L., Clark, P., Whitmer, J., Frank, P., Tantillo, C., Arnold, E., and Hughes, S. H. (1994) *J. Mol. Biol.* **243,** 472–483
20. Ding, J., Das, K., Hsiou, Y., Sarafianos, S. G., Clark, A. D., Jr., Jacobo-Molina, A., Tantillo, C., Hughes, S. H., and Arnold, E. (1998) *J. Mol. Biol.* **284,** 1095–1111
21. Sarafianos, S. G., Das, K., Clark, A. D., Jr., Ding, J., Boyer, P. L., Hughes, S. H., and Arnold, E. (1999) *Proc. Natl. Acad. Sci. U. S. A.* **96,** 10027–10032
22. Sarafianos, S. G., Das, K., Tantillo, C., Clark, A. D., Jr., Ding, J., Whitcomb, J. M., Boyer, P. L., Hughes, S. H., and Arnold, E. (2001) *EMBO J.* **20,** 1449–1461
23. Jacobo-Molina, A., Ding, J., Nanni, R. G., Clark, A. D., Jr., Lu, X., Tantillo, C., Williams, R. L., Kamer, G., Ferris, A. L., Clark, P., Hizi, A., Hughes, S. H., and Arnold, E. (1993) *Proc. Natl. Acad. Sci. U. S. A.* **90,** 6320–6324
24. Sarafianos, S. G., Clark, A. D., Jr., Das, K., Tuske, S., Birktoft, J. J., Ilankumaran, P., Ramesha, A. R., Sayer, J. M., Jerina, D. M., Boyer, P. L., Hughes, S. H., and Arnold, E. (2002) *EMBO J.* **21,** 6614–6624
25. Dirscherl, W., and Weingarten, F. W. (1951) *Justus Liebigs Ann. Chem.* **574,** 131–139
26. Gao, G., Orlova, M., Georgiadis, M. M., Hendrickson, W. A., and Goff, S. P. (1997) *Proc. Natl. Acad. Sci. U. S. A.* **94,** 407–411
27. Boyer, P. L., Sarafianos, S. G., Arnold, E., and Hughes, S. H. (2000) *Proc. Natl. Acad. Sci. U. S. A.* **97,** 3056–3061

**Trapping HIV-1 Reverse Transcriptase Before and After Translocation on DNA**

Stefan G. Sarafianos, Arthur D. Clark, Jr., Steve Tuske, Christopher J. Squire, Kalyan Das, Dequan Sheng, Palanichamy Ilankumaran, Andagar R. Ramesha, Heiko Kroth, Jane M. Sayer, Donald M. Jerina, Paul L. Boyer, Stephen H. Hughes and Eddy Arnold

*J. Biol. Chem. 2003, 278:16280-16288.*
*doi: 10.1074/jbc.M212911200 originally published online January 28, 2003*

---