Rödel, G., Holl, J., Schmelzer, C., Schmidt, C., Schweyen, R. J., Weiss-Brummer, B., & Kaudewitz, F. (1983) in *Mitochondria 1983* (Schweyen, R., Wolf, K., & Kaudewitz, F., Eds.) pp 191-201, Walter de Gruyter & Co., Berlin.
Salser, W. (1977) *Cold Spring Harbor Symp. Quant. Biol. 42*, 985-1002.
SantaLucia, J., Jr., Kierzek, R., & Turner, D. H. (1990) *Biochemistry 29*, 8813-8819.
Steger, G., Hofmann, H., Förtsch, J., Gross, H. J., Randles, J. W., Sänger, H. L., & Riesner, D. (1984) *J. Biomol. Struct. Dyn. 2*, 543-571.
Stern, S., Weiser, B., & Noller, H. F. (1988) *J. Mol. Biol. 204*, 447-481.
Sutherland, J. C., & Griffin, K. P. (1983) *Biopolymers 22*, 1445-1448.
Sutherland, J. C., Lin, B., Mugavero, J., Trunk, J., Tomasz, M., Santella, R., Marky, L., & Breslauer, K. J. (1986) *Photochem. Photobiol. 44*, 295-301.
Tinoco, I., Jr., Uhlenbeck, O. C., & Levine, M. D. (1971) *Nature (London) 230*, 362-367.
Tinoco, I., Jr., Borer, P. N., Dengler, B., Levine, M. D., Uhlenbeck, O. C., Crothers, D. M., & Gralla, J. (1973) *Nature (London), New Biol. 246*, 40-41.
Tunis-Schneider, M. J. B., & Maestre, M. F. (1970) *J. Mol. Biol. 52*, 521-541.
Turner, D. H., Sugimoto, N., Jaeger, J. A., Longfellow, C. E., Freier, S. M., & Kierzek, R. (1987) *Cold Spring Harbor Symp. Quant. Biol. 52*, 123-133.
Turner, D. H., Sugimoto, N., & Freier, S. M. (1988) *Annu. Rev. Biophys. Biophys. Chem. 17*, 167-192.
Uhlenbeck, O. C., Borer, P. N., Dengler, B., & Tinoco, I., Jr. (1973) *J. Mol. Biol. 73*, 483-496.
Ulrich, E. L., John, E.-M. M., Gough, G. R., Brunden, M. J., Gilham, P. T., Westler, W. M., & Markley, J. L. (1983) *Biochemistry 22*, 4362-4365.
Varani, G., Wimberly, B., & Tinoco, I., Jr. (1989) *Biochemistry 28*, 7760-7772.
Weiss, M. A., Patel, D. J., Sauer, R. T., & Karplus, M. (1984) *Nucleic Acids Res. 12*. 4035-4047.
Williams, A., Jr., & Tinoco, I., Jr. (1986) *Nucleic Acids Res. 14*, 299-315.
Zimm, B. H., & Bragg, J. K. (1959) *J. Chem. Phys. 31*, 526-535.
Zuker, M. (1986) *Lect. Math. Life Sci. 17*, 87-124.
Zuker, M., Jaeger, J. A., & Turner, D. H. (1991) *Nucleic Acids Res.* (in press).

Photoaffinity Labeling of the Primer Binding Domain in Murine Leukemia Virus Reverse Transcriptase†

Radhakrishna S. Tirumalai and Mukund J. Modak*
Department of Biochemistry and Molecular Biology, University of Medicine and Dentistry of New Jersey, New Jersey Medical School, Newark, New Jersey 07103

Received December 3, 1990; Revised Manuscript Received April 10, 1991

ABSTRACT: We have labeled the primer binding domain of murine leukemia virus reverse transcriptase (MuLV RT) by covalently cross-linking 5′ end labeled $d(T)_8$ to MuLV RT, using ultraviolet light energy. The specificity and the functional significance of the primer cross-linking reaction were demonstrated by the fact that (i) other oligomeric primers, tRNAs, and also template-primers readily compete with radiolabeled $d(T)_8$ for the cross-linking reaction, (ii) under similar conditions, the competing primers and template-primer also inhibit the DNA polymerase activity of MuLV RT to a similar extent, (iii) substrate deoxynucleotides have no effect, and (iv) the reaction is sensitive to high ionic strength. In order to identify the primer binding domains/sites in MuLV RT; tryptic digests prepared from the covalently cross-linked MuLV RT and $[^{32}P]d(T)_8$ complexes were resolved on C-18 columns by reverse-phase HPLC. Three distinct radiolabeled peptides were found to contain the majority of the bound primer. Of these, peptide I contained approximately 65% radioactivity, while the remainder was associated with peptides II and III. Amino acid composition and sequence analyses of the individual peptides revealed that peptide I spans amino acid residues 72-80 in the primary amino acid sequence of MuLV RT and is located in the polymerase domain. The primer cross-linking site appears to be at or near Pro-76. Peptides II and III span amino acid residues 602-609 and 615-622, respectively, and are located in the RNase H domain. The probable cross-linking sites in peptides II and III are suggested to be at or near Leu-604 and Leu-618, respectively.

Moloney murine leukemia virus reverse transcriptase (MuLV RT)[1] is a single-subunit protein of M_r 80 000 and catalyzes both RNA- and DNA-directed DNA polymerase activities as well as a ribonuclease H (RNase H) activity (Dickson et al., 1982). The primary amino acid sequence of MuLV RT has been deduced from the nucleotide sequence of a noninfectious proviral DNA (Shinnick et al., 1981), and the exact location of the genome segment coding for RT within the Pol gene has been deciphered and confirmed by NH_2-

† This research was supported in part by a grant from the National Institute of Allergy and Infectious Diseases (AI-26652) and by a postdoctoral fellowship award (to R.S.T.) from the New Jersey Commission on Cancer Research.
* Address correspondence to this author.

[1] Abbreviations: MuLV RT, murine Moloney leukemia virus, reverse transcriptase; AMV RT, avian myeloblastosis virus reverse transcriptase; DNA pol I, *Escherichia coli* DNA polymerase I; dNTP, deoxynucleoside 5′-triphosphate; 8-azido-dATP, 8-azidodeoxyadenosine 5′-triphosphate; DTT, dithiothreitol; HPLC, high-performance liquid chromatography; TPCK, tosylphenylalanine chloromethyl ketone; HEPES, *N*-(2-hydroxyethyl)piperazine-*N*′-2-ethanesulfonic acid; SDS, sodium dodecyl sulfate; PTH, phenylthiohydantoin; TFA, trifluoroacetic acid; tRNA, transfer RNA.

0006-2960/91/0430-6436$02.50/0 © 1991 American Chemical Society

terminal and COOH-terminal sequencing of the enzyme (Copeland et al., 1985). The gene coding for MuLV RT has been cloned and overexpressed in *Escherichia coli* (Kotewicz et al., 1985; Roth et al., 1985). The availability of relatively large amounts of the recombinant enzyme and the knowledge of the primary amino acid sequence have made studies on the structure–function relationships in MuLV RT feasible. The knowledge gained with this system may be expected to serve as a model system for other retroviral reverse transcriptases.

Studies in our laboratory, over the past several years, have been aimed at defining the structure–function relationships in various DNA polymerases, including retroviral reverse transcriptases, using various active-site-directed reagents. For example, pyridoxal 5′-phosphate (PLP) was shown to be a substrate dNTP binding site directed reagent for a number of reverse transcriptases (Modak, 1976). A detailed characterization of PLP-mediated inactivation of RT was shown to be due to Schiff base formation at the Lys residues involved in substrate binding (Modak & Dumaswala, 1981). A subsequent structural analysis of PLP-modified MuLV RT revealed that Lys-103 and Lys-421 are the two PLP-reactive Lys residues present in the substrate dNTP binding pocket (Basu, A., et al., 1988). By use of 4-(oxoacetyl)phenoxyacetic acid (OAPA), Lys-329 has been identified as an essential residue in the template-primer binding function of MuLV RT (Nanduri & Modak, 1990). In our continued investigation to identify other important domains in this enzyme, we have taken advantage of the fact that DNA polymerases in general and reverse transcriptases in particular bind to oligomeric primers quite effectively. Wilson and co-workers have demonstrated that the DNA synthetic reaction catalyzed by reverse transcriptases follows an orderd mechanism (Majumdar et al., 1989), with the binding of the primer to the free enzyme as the first step. Binding of the template and the substrate dNTP then follows. Oligodeoxyribonucleotides as short as four to eight bases long have been shown to serve as primers for DNA synthesis by reverse transcriptases (Tamblyn & Wells, 1975; Modak & Marcus, 1977). The first successful labeling of a DNA polymerase with radiolabeled primer using ultraviolet energy was described for terminal deoxynucleotidyltransferase (Modak & Gillerman-Cox, 1982). This method was also used to assess the binding of template-primers to reverse transcriptases in specific inhibitor studies (Basu, A., et al., 1990; Nanduri & Modak, 1990). Therefore, labeling of the enzyme with oligomeric primers appeared to be an excellent method for the identification of the primer binding domain in MuLV RT. We have now used this method to prepare covalent complexes of MuLV RT with 5′-[^{32}P]d(T)$_8$. The labeled enzyme–primer complex was used for the isolation and identification of the peptides which constitute the primer binding domain. Details of this investigation are the subject matter of this paper.

Materials and Methods

Materials

All nucleoside triphosphates were obtained from Pharmacia. [γ-^{32}P]ATP was a product of New England Nuclear while oligodeoxyribonucleotides, d(A)$_8$, d(G)$_8$, and d(T)$_8$, and synthetic template-primers were obtained from Pharmacia. Sequenal-grade triethylamine and phenyl isothiocyanate were products of Pierce Chemicals. HPLC-grade acetonitrile and water were obtained from Fisher Scientific Co. Trypsin (TPCK) was from Worthington. T$_4$ polynucleotide kinase was from New England Biolabs. MuLV RT was purified from an *E. coli* clone generously provided by Dr. Stephen Goff of Columbia University. The purification protocol was adapted from the method of Roth et al. (1985). Briefly, cells containing the induced reverse transcriptase were disrupted by treatment with NP-40 and lysozyme. The soluble extracts were then sequentially chromatographed on DEAE-cellulose, phosphocellulose, and poly(ribouridylic acid)–agarose. The final enzyme preparation was >95% pure as judged by SDS–PAGE and had a specific activity of 16 000 units/mg of protein using r(A)$_n$·d(T)$_{12-18}$ as a template primer (Modak and Marcus, 1977).

Methods

Labeling of the 5′-OH of the Primer with [^{32}P]ATP Using T$_4$ Polynucleotide Kinase. d(T)$_8$, dephosphorylated at its 5′-terminus, was labeled with [γ-^{32}P]ATP using T$_4$ polynucleotide kinase, essentially as described by Maniatis et al. (1982). The reaction mixture consisting of 50 mM Tris-HCl, pH 7.6, 10 mM MgCl$_2$, 5 mM DTT, 0.5 mCi of [γ-^{32}P]ATP, 50 nmol of d(T)$_8$, and 800 units of T$_4$ polynucleotide kinase and was incubated at 37 °C for 1 h. The reaction was stopped by the addition of 2 μL of 0.5 M EDTA, and the 5′-[^{32}P]d(T)$_8$ was separated from unreacted [γ-^{32}P]ATP by electrophoresis on a 20% polyacrylamide–7 M urea gel. The labeled primer was eluted as described by Maxam and Gilbert (1980).

UV-Mediated Cross-Linking of [^{32}P]d(T)$_8$ to MuLV RT. MuLV RT (10 nmol) was mixed with [^{32}P]d(T)$_8$ (30 nmol) (the ratio of enzyme to primer was 1:3) in a reaction mixture containing 50 mM Hepes–KOH, pH 7.8, 10% glycerol, 1 mM DTT, and 1 mM MnCl$_2$ in a final volume of 3 mL. The mixture was incubated on ice for 30 min and irradiated with UV light (1080 ergs mm^{-2} s^{-1}) for 15 min at a distance of 10 cm. The irradiated mixture was subjected to heat denaturation (90 °C, 3 min) and quickly chilled in ice. Under these conditions, the enzyme protein along with the cross-linked primer precipitates, while most of the un-cross-linked primer remains in the supernatant. The precipitated enzyme–primer complex was recovered by centrifugation.

Determination of the Extent of Cross-Linking. An aliquot of the cross-linked protein was subjected to SDS–polyacrylamide gel electrophoresis on an 8% gel (Laemmli, 1970) followed by autoradiography. The radioactive band was excised and crushed, and the associated radioactivity was determined by liquid scintillation spectrometry.

Tryptic Digestion. The precipitated enzyme–primer complex (see UV-mediated cross-linking of [^{32}P]d(T)$_8$ to MuLV RT) recovered by centrifugation was washed extensively with 100 mM ammonium bicarbonate, pH 8.0, and suspended in 50 μL of 8 M urea in 800 mM ammonium bicarbonate. The resultant suspension was then diluted to 400 μL with water to reduce the urea concentration to 1 M and that of ammonium bicarbonate to 100 mM. TPCK-trypsin was then added at a protein:trypsin ratio of 50:1 and digested at 37 °C for 2 h. After 2 h, a second aliquot of trypsin (of similar size) was added, and digestion was continued overnight. The tryptic peptides were resolved by reverse-phase HPLC as described below.

High-Performance Liquid Chromatography. Reverse-phase HPLC was performed with a Varian Vista 5500 system using a Vydac C-18 column (4.6 mm i.d. × 25 cm, 5-μm particle size, 300-Å pore size). The column was equilibrated with 10 mM sodium phosphate, pH 6.8. The gradient employed was from 10 mM sodium phosphate, pH 6.8 (A), to 10 mM sodium phosphate, pH 6.8, containing 70% acetonitrile (B). The peptides were eluted with a 180-min gradient at a flow rate of 0.7 mL/min with the following schedule: 0–120 min, 0–40% B; 120–160 min, 40–70% B; 160–180 min, 70–100% B.

Absorbance was monitored simultaneously at 220 and 254 nm with a Varian Polychrom 9060 diode array detector. One-minute fractions were collected, and the radioactivity (Cerenkov) associated with them was determined. Peptide fractions containing radioactivity were dried in vacuo, redissolved in water, and further individually purified on a C-18 reverse-phase column with shallow gradients of acetonitrile in phosphate buffer. It is important to use a neutral pH solvent (sodium phosphate, pH 6.8) in the separation protocols of nucleotide cross-linked peptides since use of acidic pH (pH 2) such as 0.1% TFA results in a significant loss of the radiolabeled nucleotide moiety (Pandey & Modak, 1988a).

Amino Acid Composition Analyses. Purified peptides were individually desalted on C-18 cartridges (Sep-pak), and an aliquot of the desalted peptide was hydrolyzed with 6 N HCl containing 0.2% phenol for 16 h at 115 °C. The resulting amino acids were converted to their phenylthiocarbamoyl (PTC) derivatives and analyzed by HPLC as described by Stone and Williams (1986).

Amino Acid Sequence Analyses. Amino acid sequence analyses of the purified, desalted peptides were carried out at the Protein Chemistry Facility of Yale University headed by Dr. Ken Williams. An Applied Biosystems Model 470A gas-phase sequenator (Williams et al., 1986) was employed in these analyses.

RESULTS

Characterization and Properties of UV-Mediated Cross-Linking of Labeled Primer to MuLV RT. A typical time course of cross-linking of 5'-[^{32}P]d(T)$_8$ to MuLV RT is shown in Figure 1 (panel A). In this experiment, a constant enzyme to primer ratio (1:3, mol/mol) was used. There was a linear increase in the covalent cross-linking of primer to the enzyme up to 15 min of exposure to UV light. Beyond 15 min, there was a nonlinear increase accompanied by some breakdown of the labeled protein (data not shown). Therefore, for all further experiments, a 15-min exposure to UV light was used. The cross-linking reaction was also found to be proportional to the concentration of the enzyme (Figure 1, panel B). Furthermore, optimal cross-linking of the enzyme to the primer was found to occur when the molar ratio of the enzyme to the primer was 1:3 to 1:5 (Figure 1., panel C). Therefore, all further work was carried out by using an enzyme to primer ratio of 1:3. The binding/cross-linking of the primer to the enzyme does not appear to be dependent on the presence of divalent cations (Mg^{2+}/Mn^2), since, in the presence of EDTA, as much as 80% of the observed cross-linking of the enzyme to the primer was still evident (data not shown). Under optimal conditions, the efficiency of cross-linking of the enzyme to the primer was approximately 8–10%.

Specificity of Cross-Linking of the Primer to MuLV RT. In order to demonstrate that primer cross-linking represents a true binding reaction, the effect of increasing ionic strength on the efficiency of cross-linking of the primer was examined. As seen from Table I, it is clear that the extent of cross-linking of d(T)$_8$ to MuLV RT decreases as the concentration of NaCl in the reaction mixture is increased. NaCl at a concentration of 200 mM appears to completely abolish the binding of the primer, and therefore the cross-linking reaction.

Identification of the primer binding domain by the primer cross-linking reaction, that we had sought, requires that the primer cross-linking that we detect is functionally equivalent to primer binding in a catalytically active enzyme. The reduction in the cross-linking of 5'-[^{32}P]d(T)$_8$ in the presence of other oligodeoxynucleotides, such as d(A)$_8$ or d(G)$_8$, and template-primers, such as r(A)$_n$·d(T)$_{15}$ or activated calf thymus

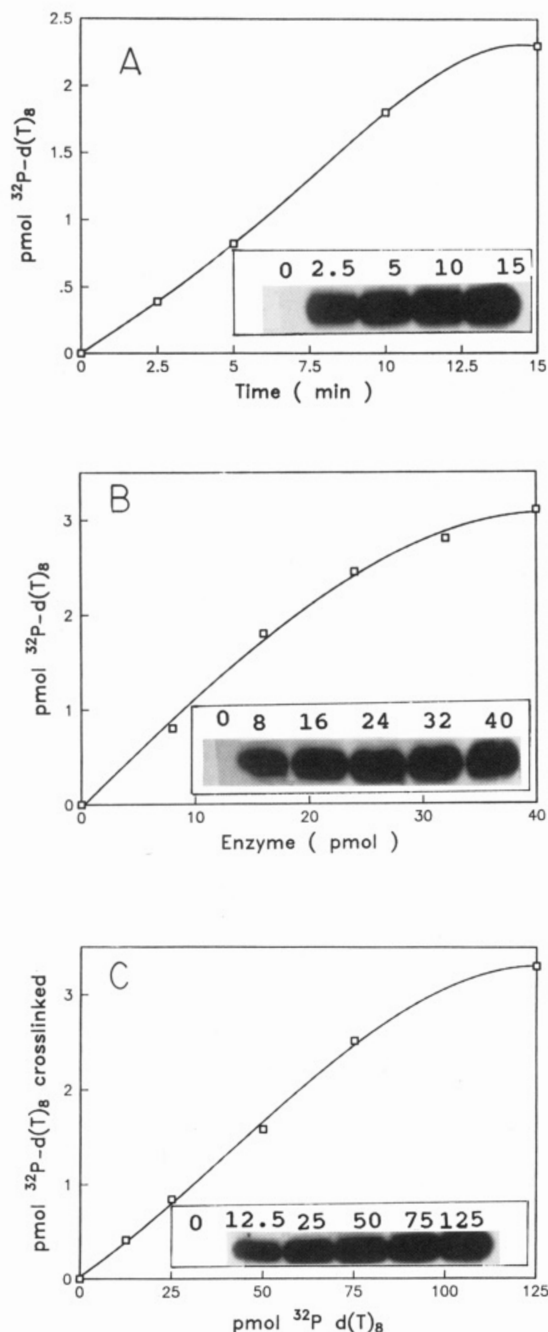

FIGURE 1: Properties of the UV-mediated cross-linking of primer to MuLV RT. MuLV RT was cross-linked to 5'-[^{32}P]d(T)$_8$ in a reaction mixture containing 50 mM Hepes–KOH, pH 7.8, 1 mM $MnCl_2$, 10% (v/v) glycerol, and 1 mM DTT under a variety of conditions as follows. Panel A represents a typical time course of cross-linking as a function of time of exposure to UV light. In this experiment, the enzyme (25 pmol) to primer (75 pmol) ratio (1:3) was kept constant. Panel B shows the linearity of the cross-linking reaction with the concentration of the enzyme (8–40 pmol) at a constant time of exposure to UV light (15 min). Panel C shows the effect of increasing primer concentration (12.5–125 pmol) on cross-linking to a fixed concentration of MuLV RT (25 pmol). In all the above experiments, after cross-linking, the enzyme–primer complexes were subjected to SDS–PAGE on a 8% gel followed by autoradiography. The radioactivity associated with them was determined by liquid scintillation spectrometry. The insets to panels A–C show autoradiograms of the UV-mediated cross-linking of MuLV RT to 5'-[^{32}P]d(T)$_8$ under the respective conditions.

DNA (Table I), indicated competitive effects. Similar reduction in the cross-linking reaction was not observed in the presence of substrate dNTPs or rNTPs (Table I).

The cross-linking of d(T)$_8$ to MuLV RT and its correlation/equivalence to functional binding of the primer are

Table I: Properties of the Primer Cross-Linking Reaction[a]

additions	pmol of $[^{32}P]d(T)_8$ cross-linked
(1) none	2.46
(2) 75 pmol of poly$[d(A)_8]$	1.37
(3) 75 pmol of poly$[d(G)_8]$	1.2
(4) 500 μM dATP	2.3
(5) 500 μM ATP	2.3
(6) 25 mM NaCl	2.32
(7) 50 mM NaCl	1.7
(8) 100 mM NaCl	0.51
(9) 200 mM NaCl	<0.1
(10) 150 pmol of $r(A)_n \cdot d(T)_{15}$	<0.1
(11) 5 μg of calf thymus activated DNA	<0.1

[a] MuLV RT (25 pmol) was cross-linked to 5'-$[^{32}P]d(T)_8$ (75 pmol) in the presence of different oligo- and polynucleotides, salt, and nucleotides, at the indicated concentrations, for 15 min in the presence of UV light. The reaction mixture consisted of 50 mM Hepes–KOH, pH 7.8, 10% glycerol, 1 mM DTT, and 1 mM $MnCl_2$. After cross-linking, the enzyme–primer complexes were subjected to SDS–PAGE on an 8% gel followed by autoradiography. The radioactive bands were excised, and the associated radioactivity was determined by liquid scintillation spectrometry.

suggested by the following observations: (i) The cross-linking and hence the binding of $d(T)_8$ to the enzyme are inhibited by the native primer tRNA (*E. coli*) (Figure 2, panel A). Under similar conditions, tRNA also inhibits, to a similar degree, the DNA polymerase activity of MuLV RT on an $r(A)_n \cdot d(T)_8$ template-primer (Figure 2, panel A). This correlation strongly suggests that $d(T)_8$ binds to the functionally significant primer binding site. (ii) $r(C)_n \cdot d(G)_8$ inhibits the cross-linking of $d(T)_8$ to MuLV RT, and the degree of inhibition appears to correlate well with the inhibition of the DNA polymerase activity on $r(A)_n \cdot d(T)_8$ template-primer (Figure 2, panel B). Furthermore, $d(G)_8$ alone, at the concentrations used in the template-primer $r(C)_n \cdot d(G)_8$, also inhibits the cross-linking of $d(T)_8$ to MuLV RT (Figure 2, panel C), suggesting the equivalence of the free primer binding site to the functionally significant primer binding site. (iii) The equivalence of the free primer binding site to the significant primer binding site, when present as a template-primer, is further strengthened by the fact that the labeling pattern of the tryptic peptides obtained from the enzyme cross-linked to 5'-$[^{32}P]d(T)_8$ and that obtained from the enzyme cross-linked to $r(A)_n \cdot 5'$-$[^{32}P]d(T)_8$ are similar (data not shown). These results are consistent with the observation that the primer or template-primer combines with the same form of the free enzyme (Majumdar et al., 1988, 1989). However, the extent of labeling with primer-template is significantly lower than that observed with primer alone. Therefore, we have used primer alone for labeling of the enzyme in further studies. Further indication of the specificity of the binding and the cross-linking of the primer to a few defined sites comes from the fact that only 3 tryptic peptides out of a total of 65 peptides seem to be cross-linked to the primer (see below). These observations clearly suggest that the cross-linking of the primer to the enzyme represents a true primer binding reaction in MuLV RT.

Detection of Primer Cross-Linking Site(s) in MuLV RT. In order to determine the domain/peptide involved in the binding/cross-linking to the primer, enzyme–primer complexes were prepared by using 10 nmol of enzyme and 30 nmol of primer, irradiated with UV light (see Methods), and subjected to tryptic digestion. The tryptic peptides were resolved by reverse-phase HPLC on a C-18 column using a sodium phosphate–acetonitrile gradient at neutral pH (Figure 3). It

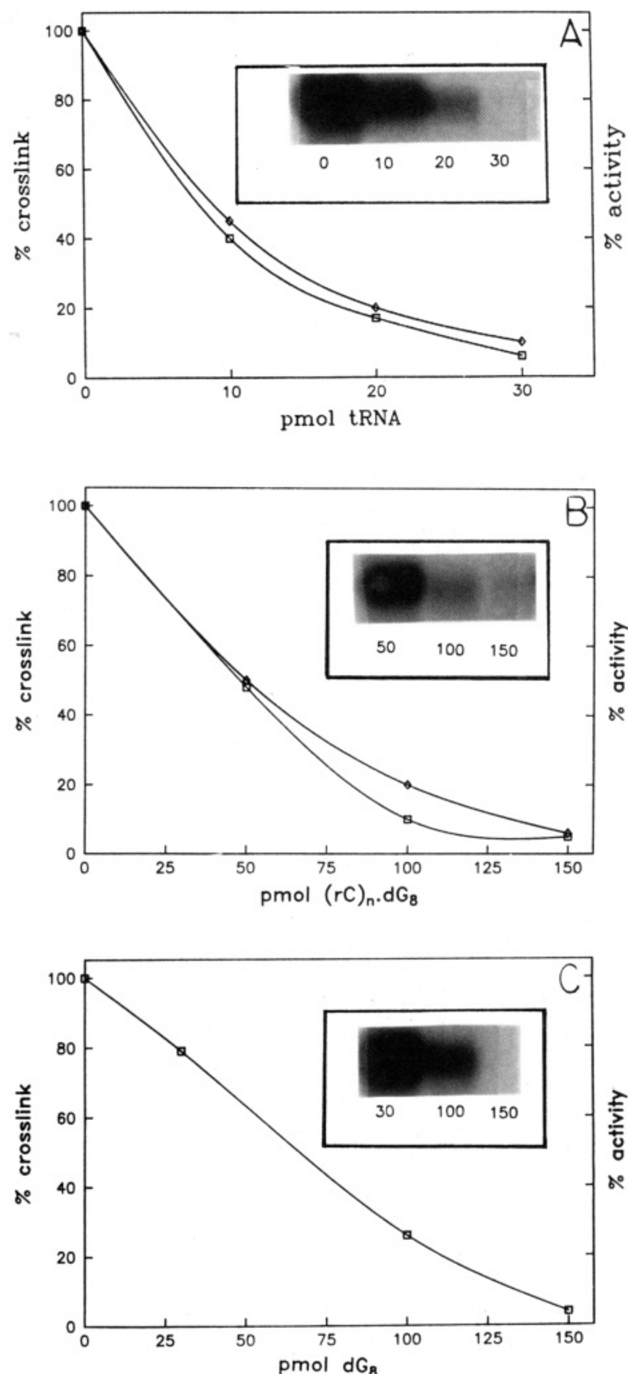

FIGURE 2: Effect of primers and template-primer on (i) UV-mediated cross-linking of $d(T)_8$ to MuLV RT and (ii) DNA polymerase activity of MuLV RT. RNA-directed DNA polymerase activity of MuLV RT was measured by using $r(A)_n \cdot d(T)_8$ template-primer (10:3 template-primer:nucleotide ratio, mol/mol) at 25 °C for 5 min with $[^3H]$TTP as the substrate. The ratio of enzyme to 3'-OH of primer (mol/mol) in the reaction mixture was 1:3. The $[^3H]$TMP incorporated was determined by liquid scintillation spectrometry. The activity (◇) was measured in the presence of different concentrations of tRNA (Panel A) and $(rC)_n \cdot d(G)_8$. In a separate experiment, MuLV RT (25 pmol) was cross-linked (□) to 5'-$[^{32}P]d(T)_8$ (75 pmol) in the presence of the indicated concentrations of tRNA (panel A), $(rC)_n \cdot d(G)_8$ (panel B), and $d(G)_8$ (panel C), in the presence of UV light for 15 min as detailed in the legend to Figure 1. The insets in panels A–C show autoradiograms of MuLV RT cross-linked to 5'-$[^{32}P]d(T)_8$ under the respective conditions.

should be pointed out that no detectable change in the overall peptide pattern of primer cross-linked and control enzymes was discernible (data not shown), indicating that exposure of enzyme protein to ultraviolet light has not caused any damage

Table II: Amino Acid Composition of Primer Cross-Linked Peptides I, II, and III

	residues/mol, of peptide					
	I		II		III	
amino acid residue	found	expected from sequence 72–80[a]	found	expected from sequence 602–609[a]	found	expected from sequence 615–622[a]
Asp	0	0	0.1	0	0.9	1
Glu	0.9[b]	0	1.4	1	0.8	1
Ser	0.1	0	1.4	1	0.2	0
Gly	1.1	1	2.4	2	0.2	0
His	0.8	1	0.1	0	0.1	0
Arg	0.9	1	0	0	0	0
Thr	0.3	0	1.4	1	0	0
Ala	0.1	0	0.2	0	1.4	1
Pro	0.8	1	0.2	0	0.1	0
Tyr	0.2	0	0.1	0	0	0
Val	0.1	0	0	0	0.1	0
Met	0.1	0	0.1	0	0.1	0
Ile	1.8	2	0.2	0	0.9	1
Leu	0.7	1	2.5	2	3.4	3
Phe	0.1	0	0.1	0	0.2	0
Lys	1.1	1	1.4	1	1.5	1

[a] Numbers represent residues in the primary amino acid sequence of MuLV RT (Shinnick et al., 1981). [b] In peptide I, Gln is detected as Glu, during composition analysis.

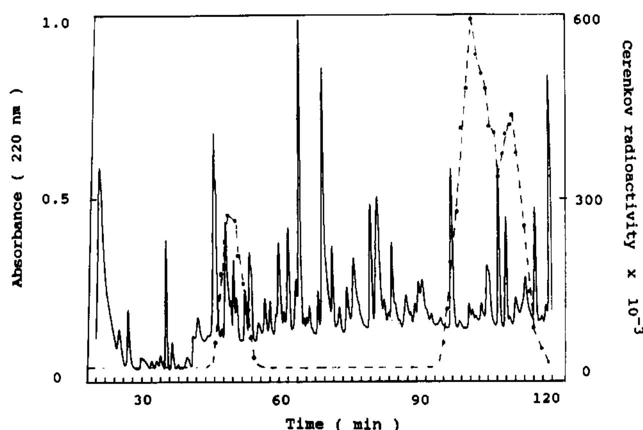

FIGURE 3: Reverse-phase HPLC separation of tryptic peptides of MuLV RT cross-linked to the primer. Ten nanomoles of MuLV RT was cross-linked to 30 nmol of 5′-[^{32}P]d(T)$_8$ as described under Methods. The cross-linked enzyme–primer complex was subjected to tryptic digestion. The tryptic peptides were resolved by reverse-phase HPLC on a Vydac C-18 column, equilibrated with 10 mM sodium phosphate, pH 6.8. Peptides were eluted by increasing the concentration of acetonitrile (0–70%) in 10 mM sodium phosphate, pH 6.8, as described under Methods. One-minute fractions were collected. The 220-nm absorbance profile (solid line) and the associated Cerenkov radioactivity (dashed line) are depicted in the figure.

to the native protein per se. However, when the radioactivity associated with the individual fractions was determined, three rather broad peaks of radioactivity were clearly noted. Of these, the peak eluting around 50 min represented free or dissociated d(T)$_8$ while the two peaks eluting between 98 and 113 min represented peptide-associated radioactivity. The peptides present in all 16 fractions (98–113 min) were further individually resolved by reverse-phase HPLC on a C-18 column with a shallow gradient of acetonitrile. A typical pattern of peptide separation obtained from 3 of the 16 fractions is shown in Figure 4. For further analyses, only peptides which were associated with radioactivity were used. Thus, 3 distinct, ^{32}P-labeled peptide species were identified from the reruns of the 16 fractions. They were designated as peptides I, II, and III in order of their retention time. By an examination of the radioactivity distribution among the three peaks, it was evident that peptide I contained nearly 65% of the total radioactivity while peptides II and III were associated with nearly equal amounts. Thus, peptide I appears to contain the major site of cross-linking.

Table III: Amino Acid Sequence Analyses of Cross-Linked Peptides

cycle	peptide I residue	peptide II residue	peptide III residue
1	Leu	Gly, Thr[a]	Asp, Ala[a]
2	Gly	Leu, Ser	Glu, Leu
3	Ile	Leu, Glu	Ile, Leu
4	Lys	Thr, Gly	Leu, Lys
5		Ser, Lys	Ala
6		Glu	Leu
7		Gly	Leu
8		Lys	Lys

[a] Peptides II and III had secondary sequences which are fragments of the same primary sequence.

Amino Acid Composition and Sequence Analyses of the Primer Cross-Linked Peptides. An aliquot of each of the peptides was subjected to amino acid composition analysis as described under Methods. The results (Table II) indicated that the amino acid composition of peptide I closely matched that of the peptide encompassing amino acid residues 72–80 in the primary amino acid sequence of MuLV RT. Peptides II and III appeared to represent amino acid residues 602–609 and 615–622, respectively, in the primary amino acid sequence of MuLV RT. In order to further confirm the amino acid composition results and identify the probable site of cross-link to the primer, an aliquot of each of the three peptides was subjected to amino acid sequence analyses (Table III). The amino acid sequence of peptide I, which by amino acid composition analysis was thought to represent amino acid residues 72–80 in the primary amino acid sequence of MuLV RT, appeared to match well with the N-terminal four residues (72–75) beyond which the sequencing reaction ends. The most likely reason for such a stop is the possibility that Pro-76 could be the site of cross-linking. Stoppage of the sequencing reaction at the site of the nucleotide cross-link has been reported (Rush & Konigsberg, 1990). In so far as peptides II and III are concerned, the sequence analyses indicated the presence of jagged N-termini in each of these peptides. In peptide II, a primary amino acid sequence representing amino acid residues 602–609 could be deduced along with a secondary sequence representing amino acid residues 605–609. Similarly,

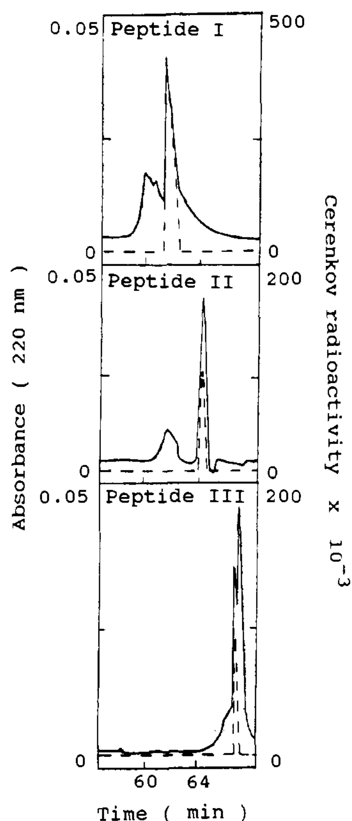

FIGURE 4: Purification of primer cross-linked peptides I–III. Tryptic peptides of MuLV RT cross-linked to the primer were initially separated on a C-18 HPLC column as described in the legend to Figure 2. Peptide fractions eluting between 98 and 114 min which were associated with radioactivity were individually collected, lyophilized, and redissolved in water. They were further individually resolved on a Vydac C-18 HPLC column equilibrated with 10 mM sodium phosphate, pH 6.8 (solvent A). The figure depicts the purification of fractions eluting at 102, 106, and 112 min (Figure 2). The peptides were eluted at a flow rate of 0.7 mL/min with increasing concentrations of solvent B (70% acetonitrile in 10 mM sodium phosphate, pH 6.8) as follows: 0–25% B (0–10 min), 25–40% (10–100 min). One-minute fractions were collected. The 220-min absorbance profile of the peptides (solid line) and the associated Cerenkov radioactivity are depicted in the figure.

in peptide III, a primary sequence corresponding to residues 615–622 was clearly evident along with a secondary sequence representing residues 619–622. Thus, Leu-604 in peptide II and Leu-618 in peptide III may have been the sites of primer cross-linking and probably yielded to cleavage at these sites during sequencing reactions.

Discussion

Ultraviolet light induced photochemical cross-linking of proteins to nucleic acids has been used as an attractive approach to identify specific amino acids at the interfaces of protein–nucleic acid complexes (Shetlar, 1980; Merrill et al., 1988). Irradiation with UV light brings about the formation of a "Zero length" covalent bond between the protein and the nucleic acid at the binding points. In principle, by the input of UV light energy, any amino acid residue which is in contact with any nucleotide residue of DNA/RNA can be induced to form a covalent cross-link (Hockensmith et al., 1991). Such a cross-linking reaction may therefore be expected to represent a true binding reaction. This laboratory has previously reported conditions for the covalent cross-linking of nucleotide substrates to *E. coli* DNA pol I and terminal deoxynucleotidyltransferase by exposing enzyme–substrate complexes to UV irradiation. (Modak & Gillerman-Cox, 1982; Abraham & Modak 1984; Pandey et al., 1987; Pandey & Modak, 1988b, 1989). Similarly, exposure of enzyme–template-primer complexes to UV irradiation has been shown to result in the covalent binding of the primer strand to *E. coli* DNA pol I (Basu, S., et al., 1988) as well as MuLV RT (Basu, A., et al., 1990). In the enzyme reaction scheme of reverse transcriptases, primer binding was reported to be the first step (Majumdar et al., 1989). We therefore expected that enzyme–primer complexes could be covalently cross-linked by exposing these complexes to ultraviolet light energy.

Reverse transcriptases have been shown to efficiently use oligodeoxyribonucleotides as a primer in the synthetic reaction directed by RNA or DNA templates (Tamblyn & Wells, 1975; Modak & Marcus, 1977). We, therefore, used $d(T)_8$ as a representative primer, since MuLV RT has been shown to utilize this primer with poly[r(A)] as a template quite efficiently (Modak & Marcus, 1977). The choice of $d(T)_8$ over longer oligomers was based on the fact that the template-primer binding cleft/groove in MuLV RT covers 10–12 base pairs as judged by DNase I footprinting experiments (S. Basu and M. J. Modak, unpublished observations). When MuLV RT–$[^{32}P]d(T)_8$ complex is exposed to UV energy, covalent cross-linking occurs between the primer and the enzyme as judged by SDS–PAGE followed by autoradiography. The characteristics of the cross-linking reaction indicated a specificity of the reaction, since other oligodeoxyribonucleotide primers as well as bonafide template-primers readily compete for binding to the enzyme and thereby reduce the cross-linking of $[^{32}P]d(T)_8$ (Table I; Figure 2). To ascertain the functional significance of the primer binding site judged by the cross-linking reaction, the effects of native primers and template-primers on the cross-linking and catalytic reactions were compared. Retroviral reverse transcriptases have been known to use a specific tRNA for their RNA-directed DNA synthesis (Taylor, 1977; Gilboa et al., 1979). For example, avian myeloblastosis RT has been known to be associated with and utilize $tRNA^{Trp}$ as a primer. AMV RT specifically binds to $tRNA^{Trp}$ in the absence of template AMV RNA (Hazeltine et al., 1977; Panet & Berliner, 1978). In the case of MuLV RT, $tRNA^{Pro}$ is used as a preferred primer, although other tRNAs can substitute $tRNA^{Pro}$ (Hazeltine et al., 1977; Panet & Berliner, 1978). We therefore compared the effects of (a) a mixture of tRNAs and (b) free $d(G)_8$ alone as well as in the form of $r(C)_n{\cdot}dG_8$, on the cross-linking of $d(T)_8$ to MuLV RT and its catalytic activity. The results shown in Figure 2 clearly suggest that the decrease in cross-linking of the primer in the presence of various competitors agrees well with the decrease in the enzyme activity. Finally, the similar pattern of labeling of tryptic peptides obtained from MuLV RT labeled with primer alone or primer-template (data not shown) strongly suggests the functional significance of the free primer $[d(T)_8]$ binding site in MuLV RT. Thus, the primer cross-linking reaction appears to represent a true primer binding reaction in MuLV RT. Since the cross-linking of the radiolabeled primer to the enzyme results in the formation of a covalent bond between the primer nucleotides and the amino acid residues present at the binding site, it is possible to identify the binding site by analyses of the tryptic peptides obtained from primer cross-linked enzyme species.

The cross-linked enzyme–$[^{32}P]d(T)_8$ complex was subjected to digestion by trypsin. The resulting tryptic peptides were resolved by reverse-phase HPLC as discussed in the section under Results. By repeated HPLC runs, three distinct labeled peptides, I–III, were identified, purified, and sequenced. Peptide I spans amino acid residues 72–80, and peptide II

Table IV: Amino Acid Sequences of Peptides I, II, and III

peptide I		peptide II		peptide III	
number[a]	residue	number	residue	number	residue
72	Leu	602	Gly	615	Asp
73	Gly	603	Leu	616	Glu
74	Ile	604	Leu[b]	617	Ile
75	Lys	605	Thr	618	Leu[b]
76	Pro[b]	606	Ser	619	Ala
77	His	607	Glu	620	Leu
78	Ile	608	Gly	621	Leu
79	Gln	609	Lys	622	Lys
80	Arg				

[a] Numbers correspond to the primary amino acid sequence of MuLV RT (Shinnick et al., 1981). [b] Probable site of cross-link.

spans amino acid residues 602–609, whereas peptide III is constituted by amino acid residues 615–622 in the primary amino acid sequence of MuLV RT. While peptide I is located in the N-terminal region of MuLV RT which harbors the DNA polymerase activity, peptides II and III are in the C-terminal region of MuLV RT wherein the RNase H activity has been shown to reside.

Peptide I constitutes the principal domain/peptide involved in the binding of the primer, since nearly 65% of the radiolabeled primer is associated with it. Pro-76 in peptide I is the probable site of the primer–nucleotide cross-link. This assignment is based on the amino acid composition and sequence analyses. Amino acid composition analysis (Table II) showed an acceptable match to the primary amino acid sequence of MuLV RT between residues 72 and 80 which reads as Leu-Gly-Ile-Lys-Pro-His-Ile-Gln-Arg. Amino acid sequence analysis (Table III) of this peptide, however, revealed the presence of the following sequence: Leu-Gly-Ile-Lys, which is only the N-terminal half of peptide I. In fact, a sequence represented by the N-terminal half of peptide I itself constitutes a qualified typtic peptide with a trypsin cleavage site at Lys-75. However, Lys-75 is followed by Pro-76 (Table IV), and the inability of trypsin to cleave the Lys–Pro bond is well-known (Keal, 1982). Furthermore, amino acid composition analysis (Table II) had indicated the presence of all nine residues in this peptide. Therefore, the observed partial sequencing of the peptide through the first four amino acids probably results from a stop of the sequencing reaction after Lys-75. The failure to sequence beyond Lys-75 may be justified if the cross-link to the primer were to occur at the carbonyl of Pro-76. Such a condition would result in an interference with the chemistry of the Edman reaction in a way that prevents further sequencing beyond Lys-75. A similar scenario has been observed in the sequence analyses of the substrate dNTP cross-linked peptide of the Klenow fragment of *E. coli* DNA pol I, wherein 8-azido-dATP was used as a photoprobe (Rush & Konigsberg, 1990). Since the sequencing reaction stopped at Lys-75, it is not clear if any other residue beyond Pro-76, for example, His-77, is also involved in the cross-linking reaction. His-77 could provide a reactive surface for the cross-link, as shown in *E. coli* DNA pol I, where His-881 has been reported to be the site of photochemical cross-link to the substrate TTP (Pandey et al., 1987).

It is relevant to recall that this very same peptide, constituted by amino acid residues 72–80, was found to comigrate with the peptide that harbors the putative dNTP binding site (Basu, A., et al., 1988). Thus, the resistance of the Lys-75–Pro-76 bond to cleavage by trypsin was already documented. We therefore conclude that peptide I encompasses amino acid residues 72–80 (Table IV) with Pro-76 as the probable site of cross-link.

It is also important to point out that the C-terminal half of peptide I is part of a unique secondary structure composed of two polypeptide double helices containing an antiparallel sequential $\beta\beta$ dimer (Chou & Fasman, 1978). Gerard et al. (1986) have suggested that amino acid residues 77–92 and 97–115, which constitute such structures in MuLV RT, could be involved in the binding of nucleic acids.

Peptides II and III represent relatively minor sites of cross-linking with nearly equal amounts of radioactivity associated with them. Amino acid sequence analyses of these peptides indicate the presence of secondary sequences (Table III) which represent partial sequences of their primary sequence; i.e., they have jagged N-termini, which suggests a cleavage/break of these peptides at Leu residues. Alternatively, cleavage at Leu could also arise from chymotryptic activity of the trypsin used for proteolysis. However, such a probability seems highly unlikely, because (a) the trypsin used for generating the tryptic peptides was treated with TPCK, a strong inhibitor of chymotryptic activity, and (b) other tryptic peptides generated with the same stock of TPCK-trypsin show proper tryptic cleavage termini (data not shown). At the present time, we like to suggest that the residues at the break points are the probable sites of cross-link to the primer.

A computer-aided comparison of amino acid sequences of different reverse transcriptases and that of *E. coli* RNase H led to the proposal that the N-terminal domain of reverse transcriptase should contain the DNA polymerase activity and the C-terminal domain should contain RNase H activity. This concept was later independently confirmed by Tanese and Goff (1988) and Kotewicz et al. (1988), by cloning the two domains separately and showing the association of the appropriate activities with the individual clones. These studies indicated that the two activities must utilize distinct and nonoverlapping active sites for nucleic acid binding and catalysis.

The results presented in this paper demonstrate that the DNA primer analogue $d(T)^*_8$, binds in the polymerase domain, which seems justified. The minor sites of binding are in the RNase H domain, wherein the RNA of the RNA·DNA hybrid is cleaved. Thus, the identification of a primer binding site in the RNase H domain is also justified. Pro-76 in the polymerase domain is in the vicinity of Lys-103, implicated in the binding of the substrate dNTP (Basu, A., et al., 1988, 1990). Proximity of the dNTP and primer binding domains is favorable for phosphodiester bond formation. Leu-604 in the RNase H domain is in the vicinity of Thr-582 and Asp-583 implicated as important residues for the RNase H activity of MuLV RT (Repaske et al., 1989). Insofar as peptide III is concerned, it lies in a region (609–671) which by deletion has been shown to be nonessential for the enzymic activity (Gerard et al., 1986). Thus, the significance of a primer binding site in this region is not clear at the present time.

In summary, our studies have identified two distinct regions for primer binding in MuLV RT, and these regions appear to represent catalytic domains responsible for the expression of DNA polymerase and RNase H activities.

Acknowledgments

We thank Dr. Stephen Goff of Columbia University for the generous gift of the MuLV RT clone used in this study. We also thank Valerie Brooks for excellent secretarial assistance.

References

Abraham, K. I., & Modak, M. J. (1984) *Biochemistry 23*, 1176–1182.

Basu, A., Nanduri, V. B., Gerard, G. F., & Modak, M. J. (1988) *J. Biol. Chem. 263*, 1648–1653.

Basu, A., Basu, S., & Modak, M. J. (1990) *J. Biol. Chem. 265*, 17162–17166.

Basu, S., Basu, A., & Modak, M. J. (1988) *Biochemistry 27*, 6710–6716.

Chou, P. Y., & Fasman, G. D. (1978) *Annu. Rev. Biochem. 47*, 251–276.

Copeland, T. D., Gerard, G. F., Hixson, C. W., & Oroszlan, S. (1985) *Virology 143*, 676–679.

Dickson, C., Eisenmann, H., Fan, E., Hunter, E., & Teich, N. (1982) in *Molecular Biology of Tumor Viruses: RNA Tumor Viruses* (Weiss, R. A., Teich, N. M., Varmus, H. E., & Goffin, J. M., Eds.) pp 513–648, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Gerard, G. F., D'Alessio, J. M., Kotewicz, M. L., & Noon, M. C. (1986) *DNA 5*, 271–279.

Gilboa, E., Mitra, S. W., Goff, S., & Baltimore, D. (1979) *Cell 18*, 93–100.

Haseltine, W. A., Pent, A., Smoler, D., Baltimore, D., Petus, G., Harada, F., & Dahlberg, J. E. (1977) *Biochemistry 16*, 3625–3632.

Hockensmith, J. W., Kubasek, W. L., Vorachekc, W. R., Evertsz, E. M. & von Hippel, P. H. (1991) *Methods Enzymol.* (in press).

Keal, B. (1982) *Methods in Protein Sequence Analysis* (Elzinga, M. Ed.) pp 291–304, Humana Press, Clifton, NJ.

Kotewicz, M. L., D'Alessio, J. M., Driftmier, K. M., Blodgett, K. P., & Gerard, G. F. (1985) *Gene 35*, 249–258.

Kotewicz, M. L., Sampson, C. M., D'Alessio, J. M., & Gerard, G. F. (1988) *Nucleic Acids Res. 16*, 265–277.

Laemmli, U. K. (1970) *Nature 227*, 680–685.

Majumdar, C., Abbotts, J., Broder, S., & Wilson, S. H. (1988) *J. Biol. Chem. 263*, 15657–15665.

Majumdar, C., Stein, C. A., Cohen, J. S., Broder, S., & Wilson, S. H. (1989) *Biochemistry 28*, 1340–1346.

Maniatis, T., Fritsch, E. F., & Sambrook, J. (1982) *Molecular Cloning: A Laboratory Manual*, Cold Spring Harbor Laboratory, Cold Spring Harbor, NY.

Maxam, A. M., & Gilbert, W. (1980) *Methods Enzymol. 65*, 499–560.

Merrill, B. M., Stone, K. L., Cobianchi, F., Wilson, S. H., & Williams, K. R. (1988) *J. Biol. Chem. 263*, 3307–3313.

Modak, M. J. (1976) *Biochemistry 15*, 3620–3626.

Modak, M. J., & Marcus, S. L. (1977) *J. Biol. Chem. 252*, 11–19.

Modak, M. J., & Dumaswala, U. J. (1981) *Biochim. Biophys. Acta 654*, 227–235.

Modak, M. J., & Gillerman-Cox, E. (1982) *J. Biol. Chem. 257*, 15105–15109.

Nanduri, V. B., & Modak, M. J. (1990) *Biochemistry 29*, 5258–5264.

Pandey, V. N., & Modak, M. J. (1988a) *J. Chromatogr. 444*, 133–139.

Pandey, V. N., & Modak, M. J. (1988b) *J. Biol. Chem. 263*, 3744–3751.

Pandey, V. N., & Modak, M. J. (1989) *J. Biol. Chem. 264*, 867–871.

Pandey, V. N., Stone, K. L., Williams, K. R., & Modak, M. J. (1987) *Biochemistry 26*, 2033–2038.

Panet, A., & Berliner, H. (1978) *J. Virol. 36*, 692–700.

Repaske, R., Hartley, J. W., Kavlick, M. F., O'Neill, R. R., & Austin, J. B. (1989) *J. Virol. 63*, 1460–1464.

Roth, M. J., Tanese, N., & Goff, S. P. (1985) *J. Biol. Chem. 260*, 9326–9335.

Rush, J., & Konigsberg, W. H. (1990) *J. Biol. Chem. 265*, 4821–4827.

Shinnick, T. M., Lerner, R. A., & Sutcliff, J. G. (1981) *Nature 293*, 543–548.

Stone, K. L., & Williams, K. R. (1986) *J. Chromatogr. 359*, 203–212.

Tamblyn, T. M., & Wells, R. D. (1975) *Biochemistry 14*, 1412–1425.

Tanese, N., & Goff, S. P. (1988) *Proc. Natl. Acad. Sci. U.S.A. 85*, 1777–1781.

Taylor, J. M. (1977) *Biochim. Biophys. Acta 473*, 57–71.

Williams, K. R., Hemmings, H. C., Jr., Lopresti, M. B., Konigsberg, W. H., & Greengard, P. (1986) *J. Biol. Chem. 261*, 1890–1903.