Deoxyribonucleotide-containing RNAs: a novel class of templates for HIV-1 reverse transcriptase

Severin O. Gudima, Ekaterina G. Kazantseva, Dmitry A. Kostyuk, Irina L. Shchaveleva, Olga I. Grishchenko, Lyudmila V. Memelova and Sergey N. Kochetkov*

Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, 117984 Moscow, Russia

Received July 10, 1997; Revised and Accepted September 15, 1997

ABSTRACT

Deoxyribonucleotide-containing RNA-like polynucleotides (dcRNAs) were synthesized by mutant T7 RNA polymerase and their structures confirmed by sequencing. dcRNAs annealed with a 20mer oligodeoxyribonucleotide primer were tested as templates/primers in the reverse transcription reaction catalyzed by HIV-1 reverse transcriptase (RT). All dcRNAs were shown to be efficient templates for both wild-type RT and RT mutants, containing 'AZTresistant' mutations. Differences in the patterns of the DNA products of RNA- and dcRNA-driven reverse transcription were demonstrated. The kinetic characteristics for dcRNAs utilization were compared with the corresponding parameters for RNA/DNA and DNA/DNA templates/primers. The respective K^m values for dcRNAs appear to be intermediate between those for RNA and DNA templates. A correlation equation connecting apparent Km value for template/ primer and the number of deoxyribonucleotide substitutions in RNA template is proposed.

INTRODUCTION

Reverse transcriptase (RT), a key enzyme of HIV metabolism, is able to carry out DNA synthesis using either an RNA or a DNA template (1–3). Some authors state that RT affinity for RNA and DNA templates $(K_m$ values) differ by about one to two orders of magnitude (4,5). The HIV isolates resistant to AZT and other nucleoside inhibitors contain RT bearing a number of mutations located around the template/primer binding site $(6-9)$. These phenomena suggest complex relationships between template/ primer and substrate/inhibitor binding. In this connection, studies of the principles of template binding and recognition by RT are of great interest. One of the approaches to study these aspects is template modification by chemical and enzymatic methods.

Recently we described a mutant form of T7 RNA polymerase (T7 RNAP) containing amino acid substitutions Y639F and S641A. The enzyme is capable of utilizing both rNTPs and dNTPs as nucleotide substrates, fully retaining its specificity for the phage T7 regulatory elements (i.e. promoters and terminators) (10). It presents a unique possibility of introducing deoxyribonucleotide residues into the RNA chain, resulting in formation of mixed ribo/deoxyribo

polynucleotides. These deoxyribonucleotide-containing RNA-like polymers (dcRNA) seem to be very promising models for both nucleic acid structural studies and studies of enzymes using nucleic acids as templates and substrates.

In this paper three main aims were achieved. First, large scale synthesis of dcRNAs was developed. Second, a new class of substrates was proposed as an instrument for studying RT template specificity. Third, the ability of RT to discriminate between dcRNAs and natural RNAs of identical base sequences was demonstrated. We hope that dcRNAs will also be useful in studying other enzymes of nucleic acid metabolism.

MATERIALS AND METHODS

Materials

Unlabelled ribonucleotides, deoxyribonucleotides, dideoxyribonucleotides, *Xba*I restriction enzyme and RNasin RNase inhibitor were purchased from Promega or Boehringer. Radioisotopes were obtained from Amersham. The sorbent for affinity chromatography, nickel *N*-tetraacetate–agarose, was purchased from Qiagen.

Oligonucleotide synthesis

The deoxyoligonucleotide template and primers were synthesized on an Applied Biosystems automatic synthesizer according to the instructions of the manufacturer.

RNA and dcRNA synthesis

The T7 RNAP mutant Y639F,S641 was obtained according to Kostyuk *et al*. (10). Both wild-type (wt) and mutant forms of T7 RNAP were purified as described previously (11). RNA and dcRNA templates were obtained by run-off transcription using plasmid pPV19. This plasmid is a derivative of plasmid pTZ19R (USB) containing an inserted *Sph*I–*Sal*I fragment of pBR322 (5).

Preparative *in vitro* transcription was carried out according to Pokrovskaya *et al*. (12) with some modifications. The incubation mixture contained 200 mM HEPES–KOH, pH 7.5, 30 mM MgCl₂, 40 mM DTT, 2 mM spermidine, 100 mg/ml acetylated BSA, 200 U/ml RNasin, 150 mg/ml pPV19 linearized by *Xba*I, 500–2000 U/ml T7 RNAP or Y639F,S641 T7 RNAP and various concentrations of rNTPs and dNTPs: 4 mM rGTP and 2 mM each rATP, rCTP and rUTP (for synthesis of RNA); 4 mM rGTP, 8 mM

*To whom correspondence should be addressed. Tel: +7 095 135 0590; Fax: +7 095 135 1405; Email: kochet@genom-ii.eimb.rssi.ru

dCTP and 2 mM each rATP and rUTP (for dC-RNA); 4 mM rGTP, 8 mM dTTP and 2 mM each rATP and rCTP (for dT-RNA); 4 mM rGTP, 8 mM dCTP, 8 mM dTTP and 2 mM rATP (for dCdT-RNA). The reaction was carried out at 38C. After 4–5 h the solution was cleared by centrifugation from a magnesium pyrophosphate precipitate and subsequently extracted with phenol and chloroform. The nucleic acids were precipitated by phenor and emotion. The nucleus actual were precipitated by adding 7.5 M ammonium acetate, 100% ethanol to the upper phase followed by incubation at -70° C for 30 min and centrifugation at 14 000 r.p.m. for 10 min. The pellet was washed twice with cool 70% ethanol, vacuum dried and resuspended in 100 µl RNase-free sterile water. Finally, the transcripts were purified by preparative electrophoresis (13) and resuspended in TE buffer (10 mM Tris–HCl, pH 8.0, 1 mM EDTA). The 32P-labelled RNA and dcRNAs were synthesized as described earlier (14). The base sequences of RNAs and dcRNAs were determined by RT sequencing as described (15).

Reverse transcriptase purification and assay

Escherichia coli BL21(DE3)plysS cells transformed by plasmid pBRP-HR or its derivatives (16) were used as the sources of wt RT p66 homodimer or 'AZT-resistant' mutants containing amino acid substitutions: K219Q (RT219); D67N, K219Q (RT67,219); D67N, K70R, K219Q (RT67,70,219); D67N, K70R, T215F, K219Q (RT67,70,215,219). The cells were cultured in a volume of 250 ml and expression of the RT gene was induced with isopropyl-β-D-thiogalactoside as described earlier (17). All RT forms were purified by affinity chromatography using nickel *N*-tetraacetate–agarose resin as described earlier (16).

RNA (or dcRNA) and DNA heteropolynucleotides (109 nt) with identical base sequences were annealed (1:1) with 20mer deoxyoligonucleotide primer 5′-GCTCTCCCTTATGCGACT-CC-3′ and used as template/primer complexes in RT-catalyzed reactions. The RT assay mixture (20 µl) contained 50 mM Tris–HCl, pH 8.0, 50 mM MgCl₂, 75 mM KCl, 10 mM DTT, 0.05% NP-40, 50 μ M each dATP, dCTP, dGTP and dTTP, 4×10^6 c.p.m. $\left[\alpha^{-32}P\right]$ dATP, 200 nM wt RT or RT mutant and variable concentrations of template/primer. Samples were incubated for 50 min at 37 $^{\circ}$ C. The apparent $K_{\rm m}$ values for templates–primers were determined in conditions described by Reardon *et al*. (18) using the MICROCAL ORIGIN 3.5 program (MicroCal Software).

Gel electrophoresis

The agarose gel electrophoresis was performed as described (13). PAGE analysis was carried out in 10% polyacrylamide gel–8 M urea gels in TBE buffer (0.089 M Tris, 0.089 M borate, 2 mM EDTA, pH 8.0) for 2–3 h at 1500 V. The reactions were stopped by addition of sample buffer (30% glycerol, 10 mM Tris, pH 8.0, 1 mM EDTA, 0.2% bromophenol blue, 0.2% xylene cyanol, 8 M urea), heated for 2 min at 95° C and applied to the gels. After electrophoresis gels were dried and subjected to autoradiography.

RESULTS AND DISCUSSION

Synthesis of dcRNA templates for HIV-1 RT

The T7 RNAP mutant Y639F,S641A displays a unique ability to utilize dNTPs in T7 promoter-dependent transcription reactions

Table 1. Templates (total length 109 nt) used in this work

dNMP incorporated are in bold. The reverse transcription scheme is shown below.

(10). Thus run-off transcription in the presence of three rNTPs and one dNTP results in formation of a RNA-like single-stranded polynucleotide containing dNMP units in all the corresponding positions instead of the respective rNMP.

Figure 1 demonstrates electrophoretic patterns of single and double deoxypyrimidine nucleotide-containing dcRNAs obtained under non-denaturing (Fig. 1A) and denaturing conditions (Fig. 1B). Figure 1A illustrates effective synthesis of the fulllength RNA and RNA-like transcripts. The similar mobilities of the transcripts suggests a similarity between the secondary structure of RNAs and dcRNAs. The upper (minor) band on the agarose gel seems to be a conformer of the major product, as repeated electrophoresis of the latter eluted from the gel has demonstrated exactly the same pattern as the run-off transcript (data not shown). This conclusion is also supported by Figure 1B, demonstrating that the electrophoretic mobilities of single band full-size modified transcripts coincide with those of 'parental' RNA.

The primary structures of three dcRNAs obtained by polymerization using mutant T7 RNAP as well as those of two 'wild-type' RNAs synthesized by mutant and wt T7 RNAP respectively were confirmed by RT sequencing. The base sequences of all five full-size transcripts appeared to be identical, thus proving the same structure as predicted by the coding DNA chain of plasmid pPV19 (data non shown). Table 1 shows the total number of dNMP residues incorporated into dcRNAs as well as 5′-end sequences for all synthesized transcripts, with an indication of the respective deoxyribonucleotides. The templates obtained were annealed with the same deoxyribonucleotide primer and used in reverse transcription reactions as template/ primer duplexes.

dcRNA template utilization by RT

Analysis of reverse transcription using dcRNAs and 'parental' RNA as templates has demonstrated their efficient utilization by the wt and mutant RTs. Figure 2 shows the relative efficiencies of DNA synthesis at saturated concentrations of the substrates. It is clear that for all dcRNAs the products are formed in quantities comparable with that for the 'wild type' RNA.

Figure 1. RNA and dcRNA synthesis by mutant (lanes 1–4) and wt (lane 5) T7 RNAPs. (**A**) 1% agarose gel stained with ethidium bromide. (**B**) 10% polyacrylamide–8 M urea gel. Lanes 1 and 5, RNA; lane 2, dC-RNA; lane 3, dT-RNA; lane 4, dCdT-RNA. The details of *in vitro* transcription and electrophoresis procedures are described in Materials and Methods.

Figure 2. Efficiency of RT-catalyzed DNA synthesis with RNA and dcRNAs as the templates (template/primer concentration $1 \mu M$). The efficiency of RNA-directed reverse transcription was assumed to be 100%, independent of RT type used. 1, wt RT; 2, RT219; 3, RT67,219; 4, RT67,70,219; 5, RT67,70,215,219.

Figure 3 demonstrates that ddNTPs correctly terminate DNA chain synthesis independently of the complementary nucleotide sugar moiety (i.e. ribo- or deoxyribo-), thus proving the correctness of the primary structure of the DNA product.

Figure 4 shows the DNA products synthesized in RNA- and dcRNA-driven reverse transcription. The presence of deoxyribonucleotides in the templates results both in the appearance/ disappearance of major termination sites and changes in termination intensities compared with 'parental' RNA (see for example bands in positions +60 and +40). These distinctions are individual for each template used. As the base sequences of the 'parental' RNA and RNA-like polynucleotides are identical, one can suggest that RT is sensitive to fine changes caused by the sugar moieties of the introduced dNMPs. No marked distinctions in the patterns for wt RT and 'AZT-resistant' mutants were observed.

Figure 3. RT sequence analysis of RNA and dC-RNA synthesized by mutant T7 RNAP. The parts of the RNA and dC-RNA sequences between two DNA synthesis pausing sites at positions +54 and +65 are shown. Lanes 1 and 6, no ddNTP added; lanes 2 and 7, $+ 5 \mu M$ ddGTP; lanes 3 and 8, $+ 8 \mu M$ ddATP; lanes 4 and 9, $+8 \mu M$ ddTTP; lanes 5 and 10, $+8 \mu M$ ddCTP. The nucleotide sequences are shown on the right side of gel pictures. Base sequences of templates of all types (i.e. RNA and dcRNAs) were identical (data not shown).

Figure 4. DNA synthesis with RNA and dcRNAs as templates catalyzed by RT67,219. The synthesized DNA products were analyzed as described in Materials and Methods. Templates: lane 1, RNA; lane 2, dC-RNA; lane 3, dT-RNA; lane 4, dCdT-RNA. The locations of selected major DNA synthesis pausing sites (indicated as the primer 3′-end extension positions) are shown on the right.

Kinetic analysis

The efficient use of dcRNAs by RT allowed us to carry out a kinetic analysis of the reverse transcription reaction. The RT assay was performed at fixed concentrations of dNTP substrates and variable

concentrations of template/primer as described previously by Reardon *et al*. (18). Linearity of the double-reciprocal plots was observed independently of the template used, thus suggesting that dcRNA-directed reverse transcription obeys Michaelis–Menten kinetics (data not shown).

As the base sequences of the dcRNAs obtained were shown to be identical with that of the 'wild-type' RNA, a comparison of apparent *K*m values seems to be adequate for estimation of the relative affinities of RT for the respective template. Table 2 shows $K_{\rm m}$ ^{app} values for template–primer for wt RT and four 'AZT-resistant' mutants. As expected, maximal RT affinity was observed for DNA and minimal for RNA template. Corresponding K_m values for dcRNAs are intermediate. The introduction of even one deoxyribonucleotide species into the RNA chain resulted in a decrease in *K*m. This affirmation is justified both for the wt RT and all mutants tested. It seems that RT can discriminate between ribo- and deoxyribonucleotides in the template strand and this effect contributes to the K_m value.

The method introduced by Reardon *et al.* (18) provides K_m ^{app} values for template–primer binding, which approximate K_d . Besides, many authors have reported a close relationship between K_m and K_d for RT–template–primer binding (19,20). As RTcatalyzed DNA synthesis is a multistep process, then, according to the Briggs–Haldane relationship, the K_m ^{app} of the overall reaction will be equal to some superposition of all intermediate equilibrium steps reflecting the multiple binding/dissociation of enzyme–template/primer (if the reaction is purely distributive) or isomerization of enzyme–template/primer complexes (if it is processive). In any case, for synthesis of a DNA of *N* nucleotides length we can assume that

$$
K_{\rm m}^{\rm app} = \prod_{i=1}^{N} (K)_i
$$
 1

where $(K)i$ is an 'elementary' equilibrium constant for each reaction step.

Each step, independently of the mechanism involved (i.e. processive or distributive) consists of formation of a productive complex between the enzyme and template/primer with subsequent dNTP binding and phosphodiester bond synthesis. Correct dNTP binding is determined primarily by interactions between the incoming substrate and the appropriate nucleotide in the template sequence. As the affinity of RT for RNA and DNA templates differs considerably (4,5), one can propose that besides Watson–Crick pairing between bases, the types of sugars forming the reaction center for each step should contribute to the 'elementary' affinity constant. This means that the 'elementary' constant for the ribonucleotide–dNTP interaction (K_{r-d}) (in the case of the RNA template) should differ from that for deoxyribonucleotide–dNTP (K_{d-d}) .

Then, in the case of a dcRNA template of *N* nucleotides with *n* ribonucleotide links, K_m ^{app} should be equal to:

$$
K_{\rm m}^{\rm app} = (K_{\rm r-d})^n \times (K_{\rm d-d})^{N-n}
$$
 2

where K_{r-d} and K_{d-d} are the 'elementary' constants for the reaction steps as indicated above.

Then

$$
lnK_m^{app} = [n \times ln(K_{r-d}/K_{d-d})] + (N \times lnK_{d-d})
$$
 3

and the plot of $\ln K_m$ ^{app} versus *n* should be linear, with the intercept and slope equal to MnK_{d-d} and $ln(K_{r-d}/K_{d-d})$ respectively. Figure 5 clearly demonstrates the correlation between $K_{\rm m}^{\rm app}$ values for wt RT and the number of ribonucleotides in the template. The theoretical K_m ^{app} values calculated from Equation **3** (Table 2, lines 2–6) are in close agreement with the experimental ones. Thus the data obtained confirm the correctness of the assumptions made.

Theoretical values for $K_{\rm m}$ ^{app} calculated from Equation **3** are in parentheses; differences between experimental and theoretical values (%) are in bold.

Figure 5. Dependence of K_m ^{app} for the template/primer on the ribo/deoxyribo composition of the template. ln*K*m values for wt RT (Table 2) were plotted against *n* (number of rNMPs in the template) according to Equation 3 (transcribed template region $N = 78$). $K_{r-d} = 1.095$ nM; $K_{d-d} = 1.038$ nM; $r =$ 0.98429.

Recent data (21) demonstrate that the primer 3′-terminal nucleotide and particularly the corresponding sugar moiety influence the affinity for RT. Thus the K_i value for an AZTterminated oligo(dT)·poly(rA) complex is about three times lower than the K_d of the non-terminated duplex analogue (21) . The data presented here suggest that the sugar moiety of the template nucleotide to be paired with the incoming dNTP immediately downstream of the 3′-end of the primer also contributes to the affinity of RT for the template/primer complex.

For wt RT and all RT mutants tested similar values of *K*r–d as well as of K_{d-d} were obtained (Table 2). This result indicates that the 'AZT-resistant' mutations, known to be located around the template binding cleft of RT (6,8,9), are apparently insensitive to the fine interactions between the incoming dNTP and the template nucleotide to be paired with it. This fact alone is not surprising, since mutated amino acid residues (67, 70 and 215) are rather far from the putative nucleotide substrate binding centre and thus are unlikely to interact directly with the pair to be formed between the dNTP and the complementary nucleotide of the template. Only the sidechain of residue 219 is located close to the dNTP binding site, but its direct interference in recognition and utilization of incoming dNTP is only hypothetical. (6).

The results presented above illustrate the advantages of dcRNAs as tools in studying nucleic acid–protein interactions. The kinetic approach employed here is believed to be applicable to other nucleotide polymerases. The only seeming limitation to the use of Equation **3** to other enzymes is the case when the latter is strictly processive and the equilibrium constant of the initial stage

(i.e. template/primer binding to the enzyme) differs greatly (at least one order of magnitude) from all subsequent isomerization steps. In this case an additional term, $\ln K_{\text{init}}$ should be included in Equation **3**. This alteration results in the plot intercept value and calculation of K_{r-d} and K_{d-d} requiring determination of K_{init} in an independent experiment. The data obtained here (Table 2) demonstrate that, at least for RT, this is not the case and the method developed can be fully applied.

ACKNOWLEDGEMENTS

This work was supported by the Russian Fundamental Science Foundation (grants 96-04-48543 and 97-04-49376) and the Howard Hughes Medical Institute (grant 75195-545003).

REFERENCES

- 1 Peliska,J.A. and Benkovic,S.J. (1992) *Curr. Biol*., **2**, 521–523.
- 2 Peliska,J.A. and Benkovic,S.J. (1992) *Science*, **258**, 1112–1118.
- 3 Jacobo-Molina,A. and Arnold,E. (1991) *Biochemistry*, **30**, 6351–6361.
- 4 Parker,B.W., White,E.L., Shaddix,S.C., Ross,L.J., Buckheit,R.W.,Jr, Germany,J.M., Secrist,J.A., Vince,R. and Shannon,W.M. (1991) *J. Biol*. *Chem*., **266**, 1754–1762.
- 5 Pokholok,D.K., Gudima,S.O., Yesipov,D.S., Dobrynin,V.N., Rechinsky,V.O. and Kochetkov,S.N. (1993) *FEBS Lett*., **325**, 237–241.
- 6 Tantillo,C., Ding,J., Jacobo-Molina,A., Nanni,R.G., Boyer,P.L., Hughes,S.H., Pauwels,R., Andries,K., Janssen,P.A.J. and Arnold,E. (1994) *J. Mol. Biol*., **243**, 369–387.
- 7 DeClercq,E. (1995) *J. Med. Chem*., **38**, 2491–2517.
- 8 Kohlstaedt,L.A., Wang,J., Freidman,J.M., Rice,P.A. and Steitz,T.A. (1992) *Science*, **256**, 1783–1790.
- 9 Jacobo-Molina,A., Ding,J., Nanni,R.G., Clark,A.D., 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. USA*, **90**, 6320–6324.
- Kostyuk,D.A., Dragan,S.M., Lyakhov,D.L., Rechinsky, V.O., Tunitskaya,V.L., Chernov,B.K. and Kochetkov,S.N. (1995) *FEBS Lett*., **369**, 165–168.
- Rechinsky, V.O., Dragan, S.M., Kostyuk, D.A., Lyakhov, D.L., Kochetkov, S.N. (1991) *Mol. Biol. (Moscow)*, **25**, 1588–1593.
- 12 Pokrovskaya,I.D. and Gurevich,V.V. (1994) *Anal. Biochem*., **220**, 420–423.
- 13 Maniatis,T., Fritsch,E.F. and Sambrook,J. (1982) *Molecular Cloning: A Laboratory Manual*. Cold Spring Harbor Laboratory Press, Cold Spring Harbor, NY.
- 14 Tunitskaya,V.L., Memelova,L.V., Kostyuk,D.A., Gudima,S.O. and Kochetkov,S.N. (1997) *Mol. Biol. (Moscow)*, **31**, 353–358.
- 15 Bordier,B., Helene,C., Barr,P.J., Litvak,S. and Sarih-Cottin,L. (1992) *Nucleic Acids Res*., **20**, 5999–6006.
- 16 Pokholok,D.K., Gudima,S.O., Memelova,L.V.,Jr, Yesipov,D.S., Rechinsky,V.O. and Kochetkov,S.N. (1996) *Biochemistry (Moscow)*, **61**, 107–113.
- 17 Pokholok,D.K., Gudima,S.O., Yesipov,D.S., Dobrynin,V.N., Memelova,L.V., Rechinsky,V.O. and Kochetkov,S.N. (1994) *Biochemistry (Moscow)*, **59**, 547–553.
- 18 Reardon,J.F., Furfine,E.S. and Cheng,N. (1991) *J. Biol. Chem*., **266**, 14128–14134.
- 19 Kedar,P.S., Abbots,J., Kovacs,T., Lesiak,K., Torrence,P. and Wilson,S.H. (1990) *Biochemistry*, **29**, 3603–3611.
- 20 Martin-Hernandez,A.M., Domingo,E. and Menendez-Arias,L. (1996) *EMBO J*., **15**, 4434–4442.
- 21 Heidenreich,O., Kruhoffer,M., Grosse,F. and Ekstein,F. (1990) *Eur. J. Biochem*., **192**, 621–625.