A Reiterative Mode of DNA Synthesis Adopted by HIV-1 Reverse Transcriptase after a Misincorporation†

Miria Ricchetti* and Henri Buc*

Unité de Physicochimie des Macromolécules Biologiques (URA 1149 du CNRS), Institut Pasteur, 75724 Paris Cedex 15, France

Received May 29, 1996; Revised Manuscript Received September 3, 1996⊗

ABSTRACT: Amplification of oligonucleotide repeats is a major cause of variability and instability of genomes. This phenomenon is probably due to an aberration in the copying process of polymerases. We show here that in the presence of $MnCl_2$, mismatch formation commits HIV-1 reverse transcriptase to a new mode of DNA synthesis which generates repetitive products. This activity is distinct from terminal transferase since it requires specific DNA motifs in the template. This mechanism, which is processive, also works on homologous RNA templates where it generates reiterative products more than 150 nucleotides long. The corresponding mechanism, which involves extensive primer misalignment, is strikingly similar to that postulated for telomerases.

Reverse transcriptases are unusual DNA polymerases which convert the RNA plus strand of a retrovirus into a double-stranded DNA molecule competent for viral integration. In order to do so, this enzyme has to fulfill very different tasks, including recognition of the tRNA-primer binding site complex, escape from this elaborated tertiary structure, intra- and intermolecular strand transfer at terminal repeat sequences, and termination. We assume that the distinct requirement to accomplish *both* elongation and recognition steps, where the relation of the enzyme to its local nucleic acid counterpart is drastically different [see for example Gopalakrishnan et al. (1992), Götte et al. (1995), and Isel et al. (1996)], could significantly increase the overall inaccuracy on specific templates. We had previously shown that the pattern of simple misincorporations generated by various reverse transcriptases was extremely dependent on the sequence context and markedly enzyme-specific [Ricchetti & Buc, 1990; for other cases see, for example, Bebenek and Kunkel (1990) and Vartanian et al. (1991)]. We are specifically interested here by more complex phenomena—extensive insertions or deletions—which are an important source of variability in small genomes [for reviews, see Kunkel (1990) and Ripley (1990)] and which cannot be explained without invoking the formation of aberrant nucleic acid structures and of modifications in the enzyme's behavior. The elucidation of their mechanisms is also important in relation to mutational processes occurring in larger genomes, in particular the amplification of di- or trinucleotides, which is associated with a number of human heritable disorders and cancers and, generally, with genetic instability in eukaryotes (Caskey et al., 1992; Thibodeau et al., 1993). The mechanism which generates such sequence repeats is still unknown, due also to failure of the biochemical approach in reproducing extensive insertions. Nevertheless DNA polymerase slippage along the template, during the replication phase, is considered as a good candidate (Strand et al., 1993).

In order to reveal unusual replication patterns occurring *in vitro*, we used conditions known to enhance the error rate for simple misincorporations, *i.e.*, biased nucleotide pools and partial or total substitution of Mg^{2+} by Mn^{2+} (Beckman et al., 1985; Bebenek & Kunkel, 1990; Ricchetti & Buc, 1990) in simple run-off experiments performed on various single-stranded templates and using different RTs.[1] We specifically looked for the occurrence of long extensions which could be due to extensive slippage. Our objectives have been first to optimize the requirements for the appearance of such an activity and to distinguish it from a classical terminal transferase activity; second, to understand the mechanism for the generation of a given extension on a specific template; and finally, to compare the efficiency and the selectivity of this process on homologous DNA and RNA templates.

MATERIALS AND METHODS

Materials. HIV-1 RT was a gift of T. Unge (Uppsala, Sweden). HIV-1 RT RNase H^- (Glu478>Gln478) and the homologous wild type RT were a gift of S. Le Grice (Cleveland, OH). When necessary, enzymes were titrated according to Kati et al. (1992) and their concentrations were expressed in moles of active dimer. AMV RT, *Escherichia coli* DNA polymerase I and T7 DNA polymerase were purchased from Pharmacia, MoMLV RT (wild type and RNase H^-) from BRL, and T4 DNA polymerase and the Klenow fragment of DNA polymerase I were from Boehringer. The deoxynucleotide primers and templates, FPLC pure, were purchased from Genset. The deoxyribonucleotide templates were either synthesized on DNA templates [as described in Ricchetti and Buc (1993)] or purchased from Genset. Primers and templates of insufficient purity were purified by gel electrophoresis [as described in Ricchetti and

† This work was supported by the Agence Nationale pour la Recherche sur le SIDA.

* Tel: +33/1-4568 8504. Fax: +33/1-4061 3060.

⊗ Abstract published in *Advance ACS Abstracts,* November 1, 1996.

[1] Abbreviations: AMV, avian myeloblastosis virus; DTT, dithiothreitol; HFB, hybrid formation buffer; HIV-1, human immunodeficiency virus of type 1; HIV-2, human immunodeficiency virus of type 2; MoMLV, Moloney murine leukemia virus; nt, nucleotide; PAGE, polyacrylamide gel electrophoresis; PRB, polymerization reaction buffer; Pu, purine nucleotide; Py, pyrimidine nucleotide; RNase H, ribonuclease H; RT, reverse transcriptase.

S0006-2960(96)01274-3 CCC: $12.00

Table 1: Templates Used in This Study

```
                          20    25    30    35
3'ACCAUCCCGAUAUGUAAGAAUGAUAAAAUAAAUUGG 5'    R2-36
3'ACCATCCCGATAUGTAAGAATGATAAAATAAATTGG 5'    D2-36
3'.....................TAAAATAAAT 5'         D2-33
3'.....................TAAAAT 5'            D2-29
3'.....................TAAAAC 5'            29C
3'.....................TAAAA 5'             D2-28
3'.....................TAATAT 5'            26T
3'.....................TAAAAAAT 5'          29A6
3'.....................TGGGGT 5'            TG4T
3'.....................TGAGAT 5'            T(GA)2T
3'.....................TAGAGT 5'            T(AG)2T
3'.....................TAAAAAAAAAA 5'       D2-A
```

[a] Position 20 of the first RNA template corresponds to position 1154 of the HIV-1 genome (Van Beveren et al., 1985). Nucleotides critical for mismatch formation are underlined.

Table 2: Primers Used in This Study

```
Reference template :
                          20    25    30    35
3'ACCAUCCCGAUAUGUAAGAAUGAUAAAAUAAAUUGG 5'    R2-36

Primers :
5'TGGTAGGGCTATACATTCTTA 3'                   p21
5'TGGTAGGGCTATACATTCTTAC 3'                  p22
     5'GGGCTATACATTCTTACTA 3'                 p24
          5'ACATTCTTACTATTT 3'                p27
          5'ACATTCTTACTTTTT 3'                p27T
5'TGGTAGGGCTATACATTC 3'                  p18(G:C)
5'TGGTAGGGCTATACATTT 3'                  p18(G:T)
          5'ACATTCTTACTATTTTATTTAACC 3'       p36

Protruding primer :
          5'ACATTCTTACTTTTTTTTTTT 3'        p34pro
          5'ACATTCTTACTATTTTTTTTT 3'        p34Apro
```

[a] The primers are indicated according to their position with respect to the a reference template, R2-36. The residue at position 24 is underlined if it forms a mismatch when hybridized to any of the templates. Protruding primers (other than p34pro) containing T tails ranging from T_9 to T_{11} have also been used.

Buc (1993)]. dNTPs substrates, FPLC pure, were purchased from Pharmacia, and [γ-^{32}P]ATP was from Amersham. Templates and primers mentioned in this study are listed in Tables 1 and 2, respectively. They are numbered from the 5′ end of the primer or from the 3′ end of the template. The corresponding hybrids are listed in Table 3.

Polymerization Assay. Conditions were selected to allow the initial run-off phase of the reaction as well as the recycling to be followed. The standard reaction was performed at 37 °C by incubating 3 nM of the preannealed hybrid (prepared in the hybrid formation buffer, HFB, consisting of 27 mM Tris-HCl pH 7.6, 4 mM $MgCl_2$, 2 mM DTT, and 100 mM NaCl) in a buffer containing 50 mM Tris-HCl, 5 mM $MnCl_2$, 50 mM KCl, 5 mM DTT, pH 7.8 (called polymerization reaction buffer or PRB) with 6 nM of HIV-1 RT. When indicated in the text, the total divalent cation concentration ($MnCl_2 + MgCl_2$) was changed and their ratio varied. Dissociation constants of the hybrid-RT complexes vary from 4 to 10 nM for a DNA/DNA duplex matched at the initial position (corresponding to 2.4−1 nM of initial complex). They increase when mismatches are introduced in the duplex and when Mg^{2+} is substituted by Mn^{2+}. Unless otherwise stated, the reaction was started by addition of 1 mM dTTP for various times (or, alternatively, 100 μM of the four dNTPs for 5 min). It was stopped by addition of 50 mM EDTA in 90% formamide (1/3 vol/vol) and heating for 10 min at 90 °C. The DNA samples in formamide were run on 20% (or 10% when specified) polyacrylamide gels containing 8 M urea. The processivity assay was performed as described above, but after preincubation of the primer/template with the enzyme, polymerization was started by simultaneous addition of 200 μg/mL of yeast DNA and dTTP.

Autoradiography and Densitometric Analysis by Direct Radioactivity Measurements. Autoradiography of samples labeled with ^{32}P was performed by exposing gels to photostimulable storage phosphorimaging plates in a PhosphorImager instrument (Molecular Dynamics). Densitometric analysis was performed with ImageQuant Software V3.0 provided by Molecular Dynamics.

Permanganate Attack. Permanganate attack of the unpaired thymidines was performed as described in Sasse-Dwight and Gralla (1989), using a final concentration of $KMnO_4$ of either 10 or 50 or 100 mM for 30 s, in a total volume of 20 μL. Extended or regular synthesis (obtained in the presence of dTTP or dNTPs, respectively) was performed as described in the polymerization assay, and permanganate attack was generally started after 15 min of incubation. Double-stranded or partially single-stranded samples, which were used as controls, were preincubated with 6 nM HIV-1 RT in PRB in the absence of deoxyribonucleotide substrates.

RNase H Assay. The RNA template R2-36, labeled with ^{32}P at its 5′ end and hybridized to the indicated primer, was incubated at 37 °C in the presence of HIV-1 RT and of either dTTP or dNTPs (or buffer alone) at the same concentrations as for the polymerization assay. The reaction was stopped by addition of 50 mM EDTA in 90% formamide (1/3 vol/vol) and heating for 10 min at 50 °C. The RNA samples in formamide were run on a 20% polyacrylamide gel containing 8 M urea as indicated above.

Assessment of the Various Rate Constants. After a scan of the different lanes of a given gel (corresponding to various times of synthesis, see Figure 1a and 1b) the relative amounts of the initial substrate x_0 and of the consecutive products, x_i, were computed by integration of the corresponding bands y_i and normalization. An experiment is therefore characterized as a set of relative values of concentrations $x_0(t)$ ... $x_i(t)$. For a fully processive synthesis the system is defined by a set of rate constants k_i relating species $i-1$ to species i. An accurate determination of those constants can be performed according to Creighton and Goodman (1995). In general, a semiquantitative approach was used here. The transit times,

Table 3: Qualitative Efficiency of the Extended Synthesis

primer/template hybrid	overhanging templates (3′−5′)	substrate	efficiency of extended synthesis
p21/D2-36	GATAAAATAAATTGG	T	+++
p22/D2-36	ATAAAATAAATTGG	T	+++
p22/D2-33	ATAAAATAAAT	T	+++
p21/D2-29	GATAAAAT	T	+++
p22/D2-28	ATAAAA	T	±
p21/26T	GATAATAT	T	−
p22/D2-29	ATAAAAT	T	+++
p22/D2-29	ATAAAAT	T + A	−
p22/26T	ATAATAT	T	−
p22/29C	ATAAAAC	T	+++
p22/29C	ATAAAAC	T + G	−
p24/D2-29	AAAAT	T	−
p27/D2-33	ATAAAT	T	−
p27T/D2-33	* ATAAAT	T	+++
p22/29A6	ATAAAAAAT	T	++
p22/D2-A	ATAAAAAAAAA	T	±
p22/TG4T	ATGGGGT	C	+++
p22/T(GA)2T	ATGAGAT	T + C	+
p22/T(AG)2T	ATAGAGT	T + C	+
p22/TAAGGT	ATAAGGT	T + C	−
p22/TGGAAT	ATGGAAT	T + C	−
p18(G:C)/R2-36	AAUGAUAAAAUAAAUUGG	T	++
p18(G:T)/R2-36	*AAUGAUAAAAUAAAUUGG	T	+++++
p21/R2-36	GAUAAAAUAAAUUGG	T	+++
p24/R2-36	AAAAUAAAUUGG	T	+

	overhanging primers (5′−3′)	substrate	efficiency of extended synthesis
p34pro/D2-29	* *TTTTT	T	++
p34pro/D2-33	* * T	T	+++
p34pro/D2-33	* * T	A	−
p34Apro/D2-33	* T	T	−

[a] Column 1 lists the hybrids according to the nomenclature used in the two preceding tables. In column 2 the overhanging single-stranded region of the template (upper part) or of some primers (lower part) are aligned, preserving their homology in the double-stranded region. A star indicates a mismatch at the corresponding position. In column 3, the nature of the dXTP substrates used (at 1 mM concentration) is indicated by the symbol characterizing the corresponding base. Column 4 gives a rough estimate of the efficiency of the extended synthesis after 15 min of reaction: (+++++) the longest products exceed the template length by more than 100 nt; (+++) more than 60% of the products are longer than the template and the longest products exceed the template length by at least 15 nt; (++) around 25% of the products are longer than the template and the longest products exceed the template length by around 15 nt; (+) around 10% of the products are longer than the template and the longest products exceed the template length by <10−15 nt; (±) less than 5% of the products longer than the template are observed, the longest products exceed the template length by less than 6 nt; (−) undetectable extended synthesis. Overhanging primers p30pro to p33pro were also used after hybridization with D2-29. Primers p30pro, p31pro, p32pro, and p33pro are respectively 4, 3, 2, and 1 nucleotide shorter than p34pro at the 3′ end; these primers are not described in the tables.

which are the inverse of the rate constants k_i, are additive (Fersht, 1985). To obtain those values, all the products resulting from incorporation at step i were summed: $p_i(t) = \sum_{j=i}^{\infty} x_j(t)$. The average time required to incorporate on the average the ith nucleotide is $\sum_{j=1}^{j=i} k_j^{-1}$). It can be estimated by measuring the time interval τ_i required for $p_i(t)$ to reach a value equal to $1 - e^{-1}$ or 0.63. (This measure would be strictly accurate if product accumulation at step i could be fitted with a single exponential $1 - e^{-t/\tau_i}$). τ_i was then plotted against i to give the time course of incorporation, as in Figure 1c.

RESULTS

Anomalous Extension. On a DNA template corresponding to 36 nt of the HIV-1 p25 *gag* region, D2-36, elongation of a 5′ labeled DNA primer of 21 nt, p21, by HIV-1 RT in the presence of 5 mM $MnCl_2$ and of 1 mM dTTP results in the atypical pattern displayed in Figure 1a. An efficient synthesis takes place during the first minute of elongation for the first 14−15 nucleotide additions (lane 3). This was followed by a slow extension, beyond the last base of the template. This anomalous polymerization process, called here extended synthesis, generates in 45 min a homopolymeric extension up to 36 residues longer than the template (lanes 4−9). In contrast, addition of the four deoxynucleoside triphosphates yields the full-length product expected from the template sequence (lane 10).

On such gels, the intensity of a given band, $x_i(t)$, characterizes the degree of arrested synthesis at position i,

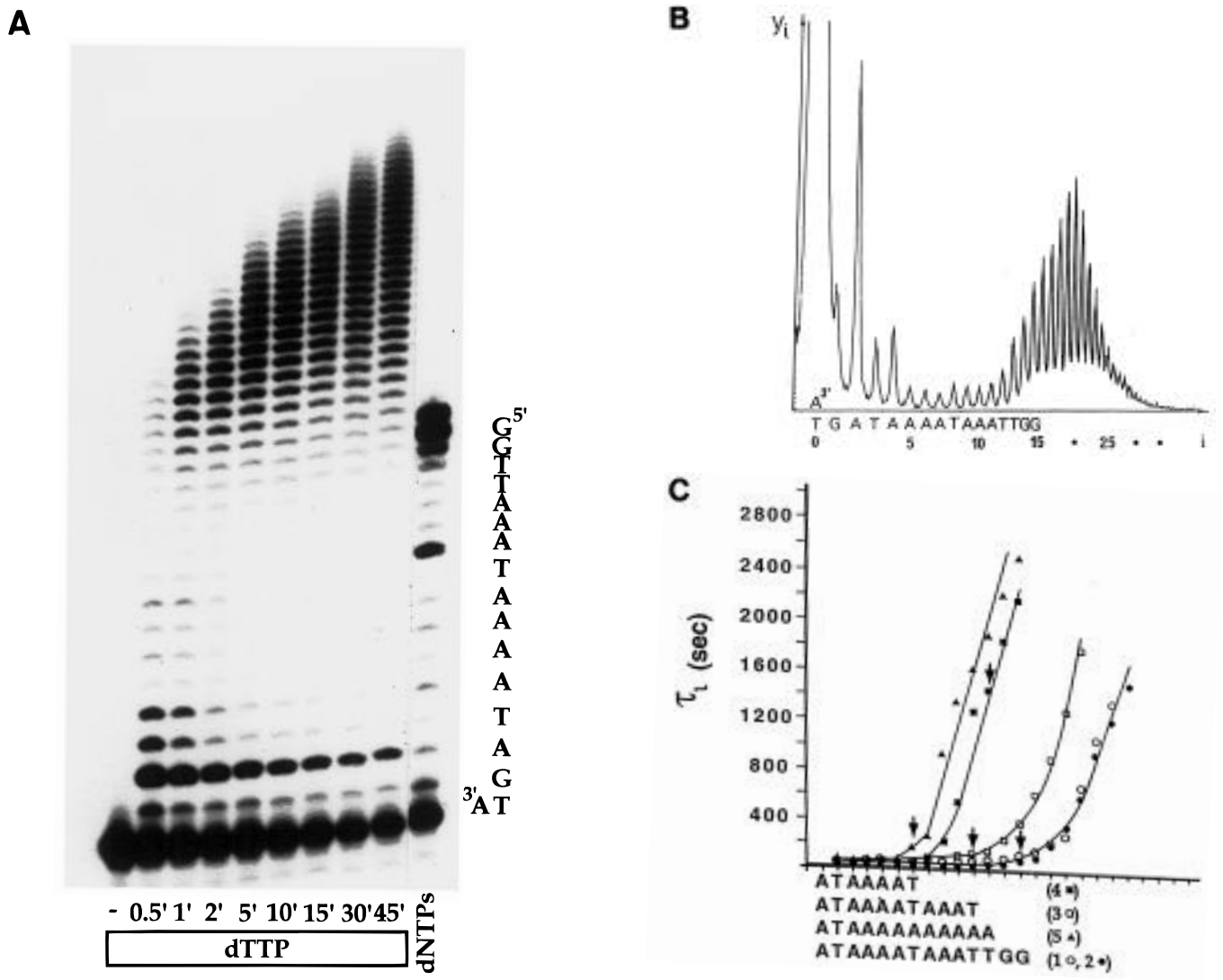

FIGURE 1: Visualization and analysis of the process of extended synthesis. (a) Time course of DNA synthesis visualized by PAGE on the preannealed hybrid p21/D2-36. Incubation of 1 mM dTTP for 30 s (lane 2) and 1, 2, 5, 10, 15, 30, and 45 min (lanes 3−9). In lanes 10 and 1 the mixture was incubated for 10 min, dTTP being replaced by 100 μM of dNTPs or omitted, respectively. On lane 10, the product one step longer than the template represents a blunt end addition [see Peliska and Benkovic (1992) and Patel and Preston (1994)]. Intermediate products are also visible. On the right side are indicated the 5′ template sequence and the last 3′ residue of the primer. (b) PhosphorImager scan of lane 5 of the preceding graph. The *i* consecutive steps take place at the expense of the hybrid substrate which by definition corresponds to $i = 0$. (c) Time course of dTTP incorporation at *i* consecutive positions for various hybrids. From scans analogous to the one displayed in 1b, the relative concentrations of all the species at various times of incubation, $x_0(t)$ ... $x_i(t)$, are obtained. The average time τ_i required to yield incorporation at position *i* is then computed for the samples as explained in Materials and Methods. τ_i was plotted as a function of *i* for five samples: (1) p21/D2-36, the original template, (2) p22/D2-36, the same template hybridized to a longer primer (these two independent determinations give the reproducibility of the experiments), (3) p22/D2-33, (4) p22/D2-29, and (5) p22/D2-A (see Tables 1−3 for nomenclature). The template sequences (from 3′ to 5′) are indicated at the bottom of the graph. Arrow indicates the end of the template.

at a given time *t* after substrate addition (for a representative example, see the scan displayed in Figure 1b). Quantitative evaluation of the average time, τ_i, required to incorporate the *i*th substrate was then possible, as explained in Materials and Methods. As revealed by the slow disappearance of the band corresponding to the incorporation of the first two dTMPs on the gel in Figure 1a, a rate-limiting step occurs during formation of a T:T mismatch at position 24 on the template. The corresponding Michaelis constant for dTTP incorporation was found to be unusually high ($K_m = 135$ μM). The subsequent elongation events during run-off synthesis correspond to very favorable kinetic parameters. No pause was seen in front of non-A residues on the template at positions 29 and 33−36 as if these bases, T and TTGG, respectively, were not read by the enzyme (see Figure 1a again). The rate of incorporation decreases when synthesis has passed the $_{3'}TA_4TA_3T_{5'}$ motif. The transit time required for each new addition can be estimated by the difference $\tau_i - \tau_{i-1}$. As shown in Figure 1c, curve 1, τ_i becomes quasi-linear with *i*, implying that the transit time then becomes approximately constant and equal to 3 min.

In order to know how frequently this anomalous process occurred, purine rich templates (see Table 1) and their complementary primers (Table 2) were synthetised. Hybrids discussed in the present text are given in Table 3, as well as the corresponding nucleotide substrates. Primers of various lengths were used to eliminate the occurrence of certain mismatches in the 3′ region of the template.

Requirements for Extended Synthesis. Addition of increasing concentrations of dATP to such a reaction mixture progressively inhibits the anomalous process (for an example, see Figure 2). An overall apparent inhibition constant can

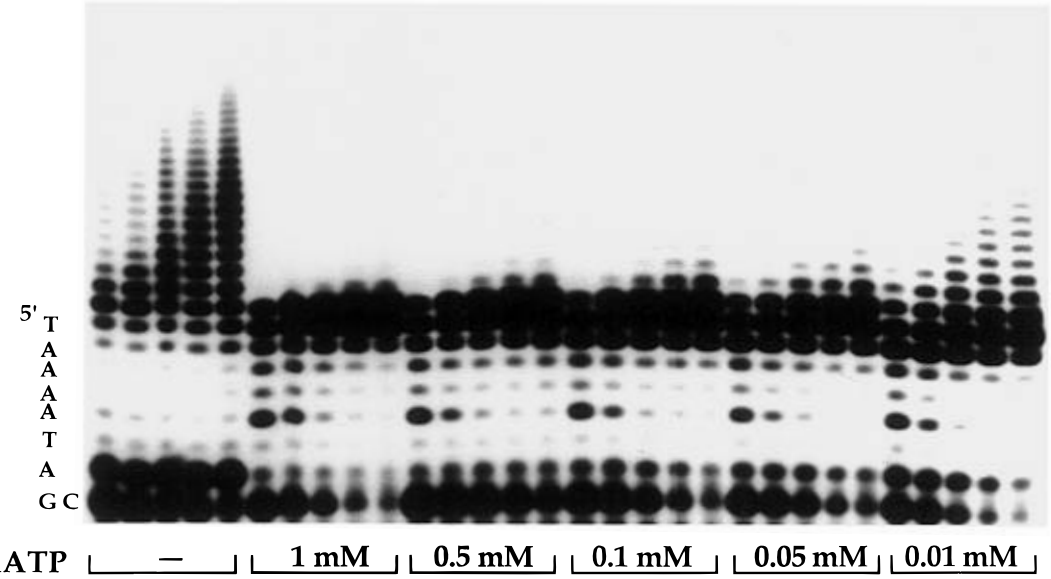

FIGURE 2: Inhibition of extended synthesis by dATP addition. Kinetics of extended synthesis on p22/D2-29 in the presence of dTTP (1 mM) and dATP (from 0 to 1 mM, other conditions as indicated at the bottom of the figure). For each dATP concentration the reaction was performed at 1, 2, 5, 10, and 15 min (from lane 1 to lane 5 of each group, respectively). On the left side are indicated the 5′ template sequence and the last (3′) complementary nucleotide of the primer. When extended synthesis is completely inhibited (as in the presence of 1 mM dATP), a product longer than the template is still present. It represents the blunt end addition which follows regular polymerization (see text and Figure 1a, lane 10).

be calculated. Its value, 2.4 μM, is similar to the Michaelis constant for incorporation of dATP at position 24 during normal synthesis, when dATP correctly pairs to the complementary T residue of the template. When a T:A match (T and A indicate the residues on the template and on the primer, respectively) was created at that position by using an adequate primer (as in hybrid p24/D2-29), extended synthesis in the presence of dTTP alone was completely suppressed (data not shown, but summarized in Table 3). For understanding the role of a mismatch at this critical position, a direct comparison was performed using the same template, and two primers which differ only by the presence or absence of a preformed T:T mismatch at position 24 (p27T/D2-33 versus p27/D2-33; see Table 3). Extended synthesis takes place again in the first case, while it is not visible in the second one (data not shown, but summarized in Table 3). As a control, the formation of a G:T mismatch at position 22 has no influence on the course and the extent of extended synthesis (compare the time course of DNA polymerization for p21/D2-36 and for p22/D2-36, samples 1 and 2 in Figure 1c). By these criteria, a specific misincorporation at position 24 appears to be required to trigger the appearance of extended synthesis.

The minimal motif required to generate extended products was assessed by modifications of the terminal sequence in the D2-36 template (in each case the extent and rate of extension were quantified as above; for a general summary see Table 3). As shown in Figure 1c, a single TA_4T motif can yield long extensions (sample 4). The rates (as well as the yields) are increased if the motif is repeated ($_{3'}TA_4TA_3T_{5'}$, sample 3) and if a random DNA tail is added ($_{3'}TA_4TA_3$-$TTGG_{5'}$, as in samples 1 or 2). Disruption of the core A_4 motif completely abolishes the anomalous process. When the template 26T (whose terminal sequence is $_{3'}TA_2TAT_{5'}$, and which differs from the efficient template D2-29 only by a A versus T substitution, see Table 3) was hybridized to primer p21 or p22, no extended synthesis was generated in the presence of dTTP (data not shown). Increasing the length of the homopolymeric TA_nT tract does not increase but rather decreases the efficiency of this anomalous synthesis: a TA_6T template (29A6) was less efficient than TA_4T (data not shown) and the TA_{10} template is very inefficient (sample 5 in Figure 1c). Deletion of the terminal 5′ pyrimidine in the TA_4T tract also leads to inefficient extensions (as with the template D2-28; data not shown). In marked contrast with slippage processes observed with other DNA polymerases (Ripley, 1990), the repeated A_n sequence must therefore be limited in length (optimal length, $n = 4$) and be flanked by two T residues.

Other sequences not related to the initial template were also able to generate extended products. We have systematically investigated sequences of the type $Py(Pu)_4Py$ using again in our assay Mn^{2+} instead of Mg^{2+} and mixtures of dCTP plus dTTP as substrates (see Table 3). TG_4T is efficient provided that dCTP is added. In the presence of dCTP alone, TA_4T is inefficient, as well as TG_4T, when dTTP is the sole substrate added, suggesting that the template for an anomalous synthesis is the polypurine sequence following the critical mismatch. This deduction is strengthened by the inspection of templates containing heteropolymeric sequences. Indeed $T(GA)_2T$ and $T(AG)_2T$ generate extended products if both dTTP and dCTP are added, while they both fail if only one of these substrates is present. However, extended synthesis on heteropolymeric templates is less efficient than on homopolymeric ones (data not shown).

Extended synthesis is observed if Mn^{2+} is replaced by another soft metal ion like Co^{2+} but not by Mg^{2+}, Cu^{2+}, or Zn^{2+} (data not shown). The concentration of Mn^{2+} could be lowered to 2 mM without affecting the occurrence of repetitive synthesis (data not shown). However a complete replacement of Mg^{2+} by Mn^{2+} is not required. When mixtures of Mg^{2+} and Mn^{2+} were used at fixed total divalent cation concentration (5 or 10 mM), half stimulation was observed at 1−2 mM Mn^{2+} (data not shown). Finally, Mn^{2+} can be first chelated on the primer/template hybrid at a

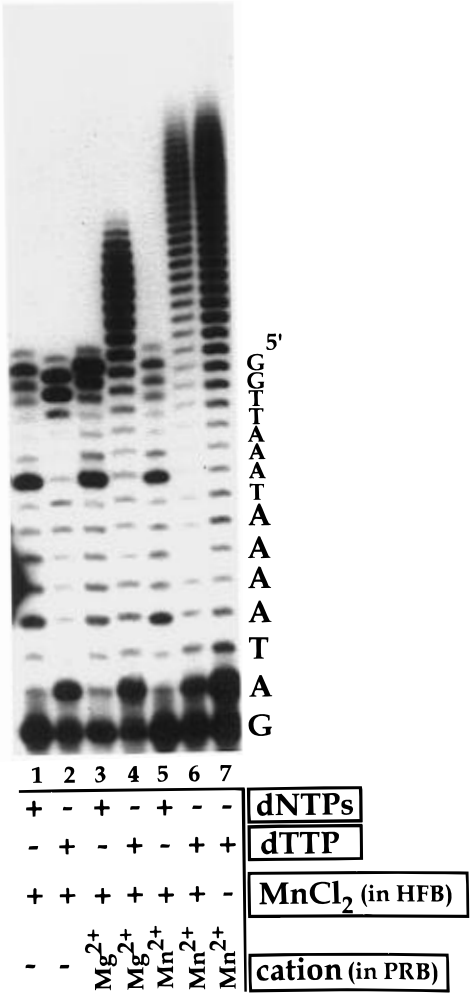

FIGURE 3: Effect of the hybrid formation conditions on the extent of extended synthesis. DNA synthesis performed on p22/D2-36 in the presence of 100 μM dNTPs for 5 min (lanes 1, 3, and 5) or in the presence of 1 mM dTTP for 15 min (lanes 2, 4, 6, and 7). The primer/template hybrid was formed in HFB (lane 7), or modified HFB where the $MgCl_2$ was replaced by 5 mM $MnCl_2$ (lanes 1−6), then 10-fold diluted in the hybrid storage buffer (10 mM Tris-HCl, pH 8, 100 mM NaCl) either in the absence (lane 7) or in the presence (lanes 1−6) of 5 mM $MnCl_2$. This hybrid was again diluted 10-fold in PRB which contained either no divalent cation (lanes 1 and 2) or 5 mM $MgCl_2$ (lanes 3 and 4) or 5 mM $MnCl_2$ (lanes 5−7). The reaction was then proceeded as described in Materials and Methods. The template sequence is indicated on the right.

concentration of 5 mM and then be diluted to 0.5 mM before addition of enzyme and 5 mM $MgCl_2$. The system is again fully committed toward extended synthesis. (Figure 3).

On the $Py(Pu)_nPy$ motives we have performed an extended synthesis assay with other DNA polymerases (*E. coli* DNA polymerase I, bacteriophage T4 and T7 DNA polymerases) and reverse transcriptases (AMV-RT, MoMLV-RT, or HIV-2 RT). With the exception of HIV-2 RT, all the other enzymes failed to produce extended synthesis (data not shown).

Extended Synthesis Does Not Proceed via Terminal Transferase Additions. The ability to catalyze consecutive additions of a given nucleotide onto the 3′ terminus of a DNA fragment without using the coding information present at the 5′ end of the template is reminiscent of the behavior of known terminal tranferases [see Kornberg and Baker (1992) and references therein]. Indeed addition of few extra nucleotides beyond the template termini has been documented for HIV-1 RT (one to four extra nucleotides) as well as for other DNA polymerases and reverse transcriptases (one extra nucleotide in general) (Patel et al., 1994; Clark, 1988). In these cases the reaction requires a duplex DNA but does not use the coding information of the template strand. In the present assay, the fact that extended synthesis requires a specific polypurine tract embedded within the template already indicated that our anomalous polymerization does not result from a classical terminal transferase activity present in HIV-1 RT. We specifically checked, however, that extended synthesis *does* need the presence of a template. When used alone, none of the tested primers (p21, p22, p36, or p34pro) is elongated by HIV-1 RT in the presence of either dTTP alone or of the four dNTPs. Furthermore, the reaction does not proceed *via* blunt end addition of a duplex DNA, even if this DNA contains a TA_4T sequence: the perfectly matched double-stranded hybrid p36/D2-36 fails to be elongated in the presence of any deoxynucleotide substrate (data not shown). For these reasons we exclude the possibility that extended products are generated by either a blunt end addition or by a terminal transferase activity.

Reorganization of the Hybrid During Extended Synthesis. Extended synthesis appears strikingly more processive than a classical elongation. Residence times of only 1−5 s per elongation step are normally observed when synthesis takes place with all four dNTPs. When DNA synthesis is performed on the p22/D2-36 hybrid in the presence of dTTP alone, and challenged by yeast DNA at the time of nucleotide addition, a very peculiar shift in the pattern of intensities of the various bands $x_i(t)$ is observed as displayed in Figure 4. In the initial phase of the reaction (2 min of incubation, Figure 4, left panel), the competitor drastically affects the concentrations of the first and the second bands of the pattern, indicating that reinitiation occurs at these steps in its absence. The overall shape of the pattern of extended synthesis ($i \geq 2$) is fully maintained when the competitor is present. An attenuation of each band by a similar factor (3.6 on average) is observed. The same 3.6 factor applies to the attenuation of the intensity of $x(1)$, the band preceding the formation of a T:T mismatch which engages the enzyme into the new mode of synthesis. Therefore, this slow step constitutes a "gate" after which the complex appears immune to the competitor. At longer times of incubation (for example, $t = 10$ min, Figure 4, right panel) the pattern of extended synthesis is again almost unaffected by the addition of the competitor, though one can notice a slight shift in the position of the maximum (from $i = 14$ to $i = 13$). The amount of material, which trails behind the wave of extended synthesis and which has passed the gate ($2 \leq i \leq 9$), is essentially the same in both experiments. It corresponds to a fraction of less than 9% of the binary complexes engaged in anomalous extension. It represents an upper limit for the amount of dissociated complexes at those steps. The evolution of such patterns with incubation times allows estimates of the average residence time of the RT−hybrid complex at each step of an extension. They range between 3 and 4 h.

HIV-1 RT must therefore display a tight affinity for the intermediates which accumulate during the anomalous DNA polymerization. To characterize such intermediates, a series of primers overhanging a template containing either one (D2-29) or two (D2-33) TA_nT motifs by one to five T motifs were synthesized (see Table 3 and its legend). After hybridization with those templates they were assayed for their

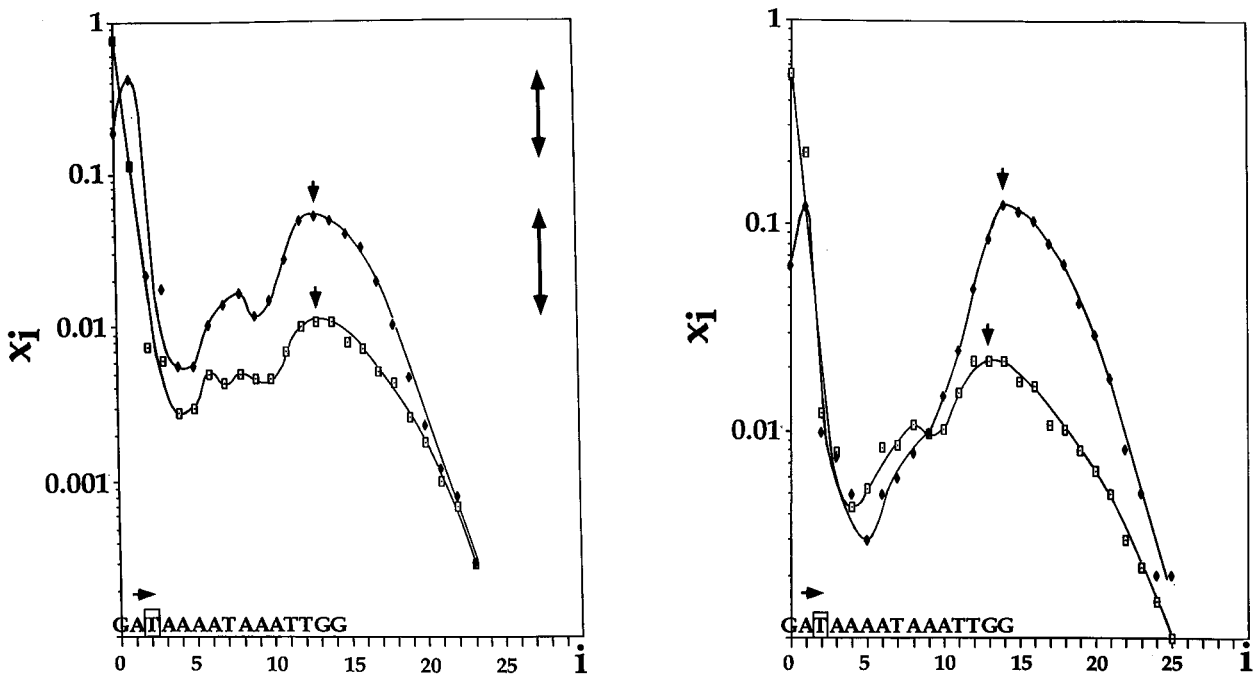

FIGURE 4: Processivity of extended synthesis. A DNA polymerization assay in the presence of dTTP alone was performed as described in Materials and Methods either in the absence or in the presence of competitor DNA, on the preannealed hybrid p22/D2-33. Incubation times ranged between 40 s and 15 min. Controls, designed to check the processivity of a synthesis performed in the presence of the four substrates, plus or minus competitor, were included. After a scan of the relevant lanes of the gels, the relative concentration of each extended species, x_i, was determined. The corresponding distribution is plotted here on a semilogarithmic scale as a function of i, the number of incorporated nucleotides for four representative scans. Left panel. Incubation time of 2 min. Distributions obtained in the absence (diamonds) or in the presence (rectangles) of competitor. Downward arrows indicate the common location of the maxima for the bands corresponding to an extended synthesis ($i = 13$). An attenuation factor can be defined as the ratio $x_i(+)/x_i(-)$, for each value of i, or the difference in ordinates of the two patterns in these semilogarithmic representations. It is represented here for $i = 1$ (just before formation of a T:T mismatch at position 24, boxed in the template sequence displayed in the abscissa) by the upper bar on the right, and at the common maximum (lower bar, on the right). Right panel. Incubation time of 10 min. Same representation as in the preceding case. The maximum intensities of the bands corresponding to extended synthesis occur now at $i = 14$ (in the absence of competitor) and at $i = 13$ (in its presence), as indicated by the two arrows.

ability to bind RT and to sustain reinitiation of a dTTP dependent elongation. Binding is assayed by the ability of the mixture to incorporate a nucleotide (here dATP), which correctly pairs with the 3′ end of the template motif (here a T). These overhanging primers are considered to be putative intermediates of the reaction, if they also incorporate several dTTP molecules, at a rate equivalent to the rate observed at the same step during steady-state synthesis (see scheme in Figure 5a). As shown in Figure 5b and 5c, incorporation of dATP occurs in *all* of the overhanging primers assayed in this series. When dATP is replaced by dTTP, not one but several thymidines are added sequentially, indicating that initiation of extended synthesis is also taking place (see Figure 5a and 5b again). In general however, addition occurs at a slower rate than the one observed during the corresponding forward synthesis. On a template terminated by a single TA_4T motif, the reinitiation rates vary from 0.022 to 0.03 min^{-1} per added nucleotide when the primer contained from five to one overhanging motifs (Figure 5c; other results not shown). The rate of reinitiation becomes equivalent to the forward rate of extended synthesis (0.3 min^{-1}) only when the D2-33 template, terminated by $_{3'}TA_4TA_3T_{5'}$, was hybridized with p34pro which contains a tail overhanging by a single T (Figure 5b). Reentry was then no longer rate limiting for the reestablishment of steady-state synthesis. Extensive primer overhang beyond the last base of the template is therefore prevented during extension of the primer. A constant repositioning of the primer strand with respect to the template (probably by some type of loop formation, as sketched in Figure 5a) should therefore occur.

As for regular extended synthesis, elongation of protruding primers hybridized to the template $_{3'}TA_4TA_3T_{5'}$ or TA_4T takes place only if $MnCl_2$ is present in the reaction buffer. Additionally, primer backward slippage and its subsequent elongation do not occur in the absence of the T:T mismatch at position 24: if p34pro/D2-33 was substituted with the hybrid p34Apro/D2-33 (which forms a T:A match at position 24, see Table 3), the protruding primer was not elongated. Primer backward slippage and extended synthesis are also both inhibited by the formation of a correct pairing downstream the A_4 motif. If the pyrimidine residue located at position 29 on the template ($_{3'}TA_4Py_{5'}$) is correctly paired, extended synthesis is abolished, even in the presence of a T:T mismatch at position 24. Inhibition of extended synthesis takes also place when the hybrid p22/29C (which has the critical sequence $_{3'}TA_4C_{5'}$) is elongated in the presence of both dTTP and dGTP, this last substrate being now correctly paired to the 5′ pyrimidine base (data not shown, but summarized in Table 3). On the contrary in the presence of dTTP alone (or of both dTTP and dCTP) extended products are normally formed. As efficient extension, repositioning of the primer strand and subsequent elongation are therefore dependent (i) on the presence of a misincorporation upstream of the poly-A tract and (ii) on the absence of a correct pairing downstream of the same homopurine tract.

Repositioning of the primer strand was then probed on the same templates by permanganate attack of their unpaired

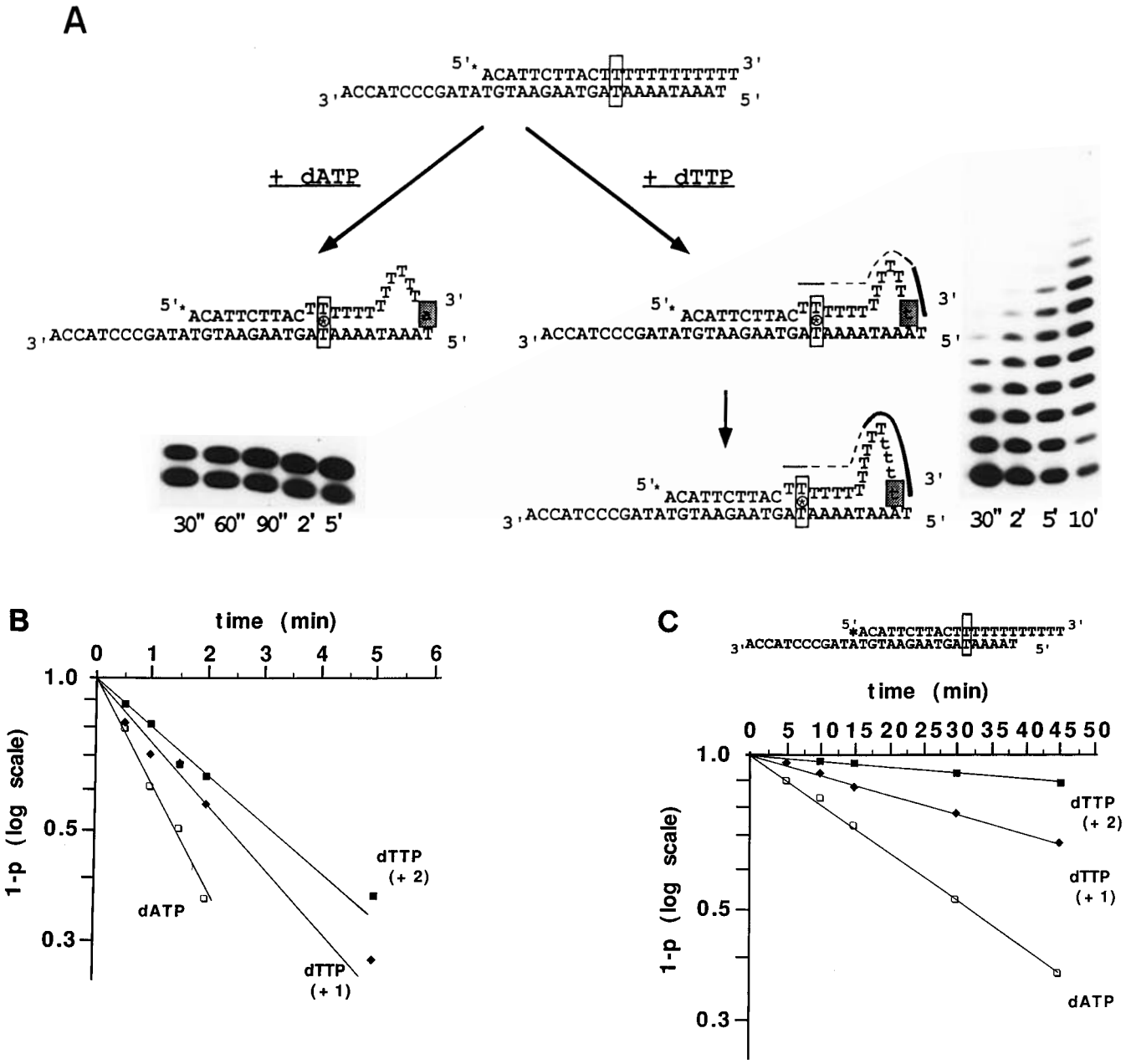

FIGURE 5: Reorganization of several hybrids during extended synthesis. (a) Hybrid p34pro/D2-33, a primer protruding by one base, is represented on the first line. The minimal reorganization of this primer as HIV-1 RT binds and incorporates one dATP molecule is represented on the left side (pairing of the incoming substrate with the terminal T residue on the template). On the right are schematically represented unpairings required for the first and the fourth incorporations when dTTP is added (in this case any A residue could probably be used as a nucleotide template). The original primer and template sequences are indicated in capital letters. Small letters represent the incorporated residues. The position where the primer is labeled is marked with a star. The critical mismatch at position 24 is boxed. The gels on the left and on the right sides show, at the times indicated, the actual elongations observed in the presence of dATP or dTTP, respectively. In both cases the concentration of the substrate is 500 μM. Significant differential reactivities of those structures with respect to $KMnO_4$ are displayed on the right. These reactivities were probed with the corresponding hybrids complexed with HIV-1 RT as dTTP was incorporated (see text). Full line, strong reactivity to $KMnO_4$; thin line, mild reactivity; dotted line, almost no reactivity. A circled star indicates the incorrect pairing at position 24. (b) Kinetics of elongation of p34pro/D2-33. A hybrid was formed between p34pro and D2-33 (see the upper part of panel a) and incubated with HIV-1 RT. After 15 min a single nucleotide monophosphate was added (700 μM). The disappearance of the primer, monitored by $x_0(t)$, and the appearance of consecutive products $p_i(t)$ were followed. They are displayed on a semilogarithmic plot, $\log x_0 = f(t)$, $\log(1-p_i) = g(t)$ allowing measurement of the corresponding rate constants $k_i'(t)$. Open square, addition of dATP, $k_1'(t) = 0.46$ min^{-1}; filled diamond, addition of dTTP, $k_1'(t) = 0.3$ min^{-1}, $k_2'(t) \approx 0.6$ min^{-1}. The latter rate constants are of the same order of magnitude as the ones measured during elongation of a recessed primer, for the formation of the same product: $k_{13}(t) = 0.32$ min^{-1}, $k_{14}(t) = 0.45$ min^{-1}. (c) The same experiments as in panel b, performed with the hybrid formed between D2-33 and p34pro (overhang of five nucleotides, see the upper part of the figure). Open square, addition of dATP, $k_1'(t) = 0.022$ min^{-1}; filled diamond, addition of dTTP, $k_1'(t) = 0.015$ min^{-1}, $k_2'(t) > 0.015$ min^{-1}. These rate constants are clearly smaller than the ones measured during steady-state elongation: $k_{11}(t) = 0.47$ min^{-1}, $k_{12}(t) = 0.30$ min^{-1}.

thymidines. Patterns of reactivities of hybrids alone and of hybrids complexed with the enzyme during normal or extended synthesis were compared, using the conditions described in Materials and Methods. Results obtained with hybrids p22/D2-29 and p22/D2-36 are self-consistent. They show three different levels of reactivity for the complexes: poly-dT tails expected to protrude from the template appear uniformly sensitive to $KMnO_4$, and the Ts located just

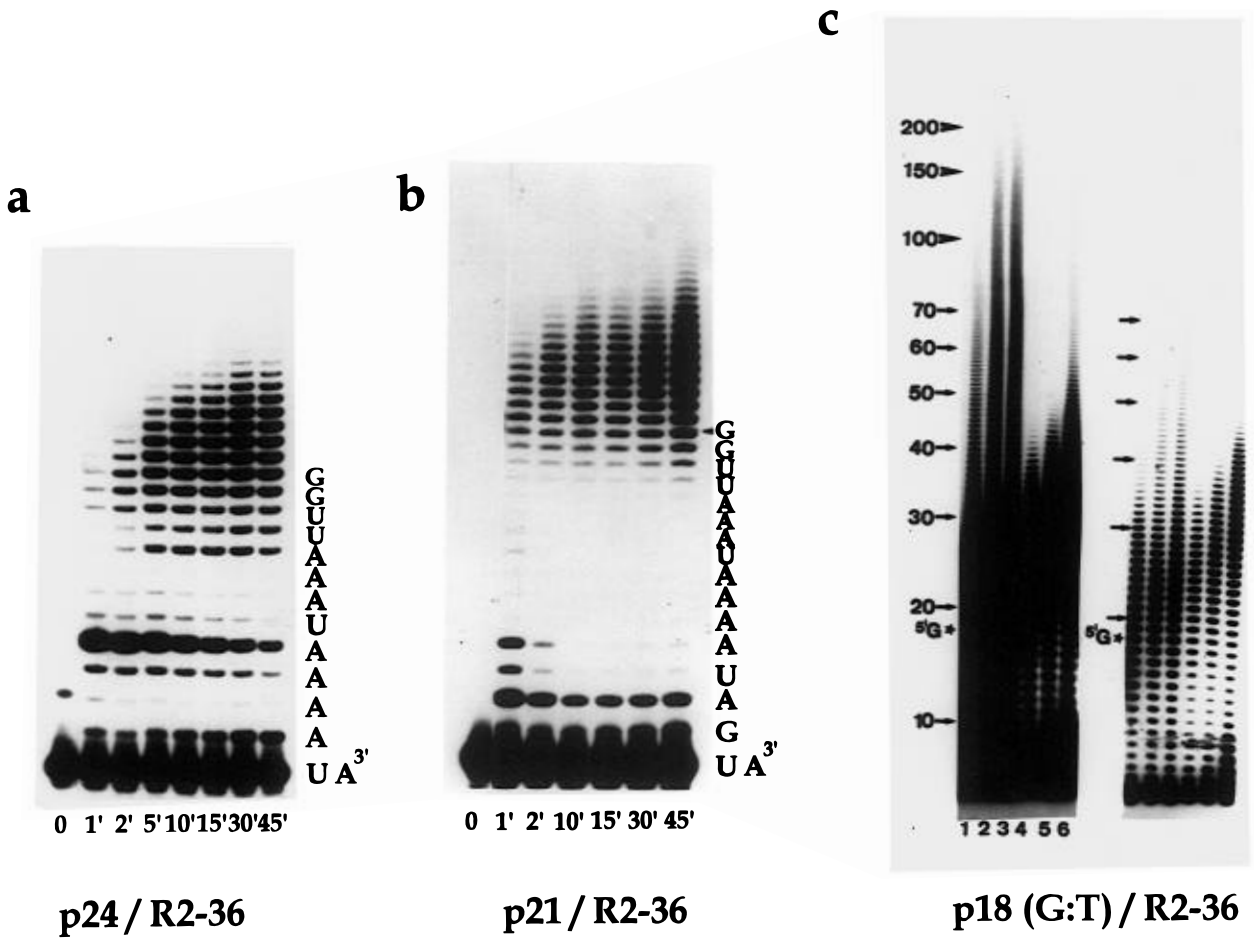

FIGURE 6: Progressive increase in the efficiency of extended synthesis on RNA templates as potential mismatches accumulate. (a) Extension of a primer matched with an RNA template up to the first U of the crucial motif UA_4U (p24/R2-36). Synthesis monitored at 1, 2, 5, 10, 15, 30, and 45 min (lanes 2−8) is rather inefficient. On the right, the template sequence is represented. The non-elongated primer is used as a marker in lane 1. (b) Extension on p21/R2-36 monitored as in a (the 5 min time point is missing); here the first uridine of the motif is not prematched. Extension is more efficient. The template sequence is represented on the right. (c) A new G:T mismatch is introduced at position 18 [p18(G:T)/R2-36]. On the left the G present at the 5′ end of the template is labeled with a star. The kinetics of extension are performed with a wild-type HIV-1 RT (lanes 1−3) and with an RNase H^- mutant (lanes 4−6). Left and right panels correspond to gel exposures of 40 and 5 h, respectively. The degree of polymerization of the consecutive products is given in nucleotides (arrows on the left side), the longer products (arrowheads) being estimated through a logarithmic extrapolation. In order to visualize bands corresponding to the longer products (>70 nt), which cannot be discriminated on a 20% PAGE, the samples have been separated on 10% PAGE.

downstream of the critical mismatch are uniformly protected while this mismatch and the base just upstream of it are both mildly exposed (data not shown). These results were used to propose the sketch displayed in Figure 5a, assuming that full protection results from a local double-stranded structure of the DNA.

Extended Synthesis on RNA Templates. RNA sequences homologous to the previous DNA templates were then systematically assayed. The order of efficiency of the crucial motifs is essentially unchanged when the RNA sequences are substituted for DNA sequences, a $_{3'}UA_4UA_3U_{5'}$ motif being the most efficient. The role of various mismatches was investigated on this type of template (see Figure 6). Extension was again inefficient when the first U at position 24 in the $_{3'}UA_4UA_3U_{5'}$ motif is matched (p24/R2-36, Figure 6a). Synthesis increases when it is mismatched (p21/R2-36, Figure 6b). The introduction of a further G:T mismatch at position 18 markedly increases the rate of primer elongation, specifically when the enzyme carries an efficient RNase H activity [p18(G:T)/R2-36, Figure 6c; compare lanes 1−3 with lanes 4−6]. Very long products are observed with the wild type enzyme, probably because the steady degradation of the template by the RNase H activity carried by the enzyme resolves bulging problems during extended synthesis [reviewed in Le Grice (1993)].

Experiments performed with RNA templates allow positioning of the DNA polymerase site of HIV-1 RT during extended synthesis. To do so, we have first localized the RNase H site by the pattern of attack of a labeled template; we have then calculated the average distance (in nucleotides) between the two catalytic sites by the method described in the legend of Figure 7, and last, we have replotted the reactivity data on the sequence of the template after a corresponding translation. The pattern of degradation of a labeled RNA template was established either in the absence of any nucleotide, in the presence of all the dNTPs or of dTTP alone. In the absence of added substrate, the DNA polymerase site, P, is expected to sit at the 3′ end of the primer. In this case a distance of 15 bases separates the polymerization site, P, from the prevalent RNA cleavage site, R, when primers p21 or p24 are used (see Figure 7a,b, lanes b and c, and scheme in Figure 7e). An identical distance has been previously measured on a different hybrid by Peliska and Benkovic (1992). In agreement with these authors we also observed shorter 5′ labeled fragments

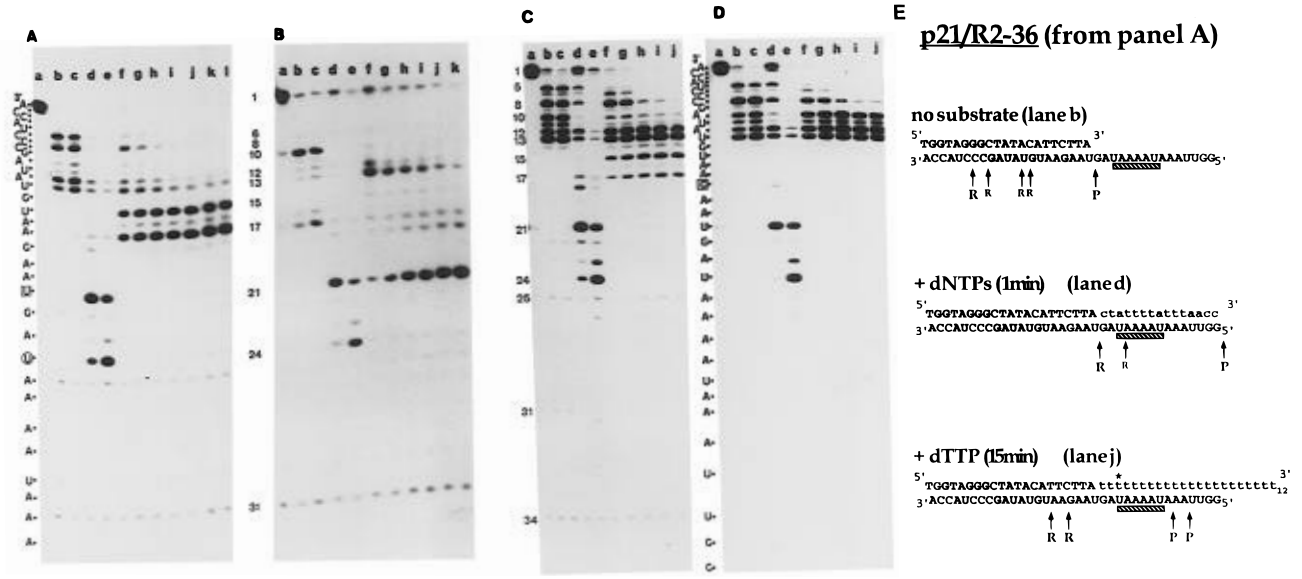

FIGURE 7: Nuclease degradation of the RNA template by the RNase H activity of HIV-1 RT. Scheme for positioning the catalytic sites on p21/R2-36. (a) p21/R2-36. (b) p24/R2-36. (c) p18(G:C)/R2-36. (d) p18(G:T)/R2-36. Lane a, unprocessed R2-36 RNA; lanes b and c, hybrid and HIV-1 RT incubated for 5 and 15 min, respectively; lanes d and e, hybrid, HIV-1 RT, and dNTPs incubated for 1 and 5 min, respectively; lanes f−l (panel a), hybrid, HIV-1 RT, and dTTP incubated for 1, 2, 5, 10, 15, and 30 and 45 min, respectively; lanes f−k (panel b), hybrid, HIV-1 RT, and dTTP incubated for 2, 5, 10, 15, 30, and 45 min, respectively; lanes f−j (panels c and d), hybrid, HIV-1 RT, and dTTP incubated for 7, 15, 30, 45, and 60 min, respectively. (e) As an example, a scheme for the major degradation patterns observed on the p21/R2-36 hybrid, panel a, is represented. The RNA template is labeled with ^{32}P at its 5′ end. The position of a prevalent RNase H degradation site (R), resulting from panel a, is shown by an arrow (minor degradation sites, appearing at longer incubation times, are indicated by an "R" of reduced size). The hypothetical position of the corresponding DNA polymerase site (P) is represented by an arrow located 15 residues more downstream. In the absence of substrate (lane b) or of dNTPs addition (lane d), the DNA polymerase site must be located at positions 22 and 36, respectively (see text). In both cases the predicted P site corresponds to the one calculated by adding 15 residues to the position of the R site. During extended synthesis in the presence of dTTP (lane j), two major RNase H degradation sites are present. The corresponding polymerization sites are therefore at positions 30 and 32 after translation by 15 residues on the sequence. The horizontal bar underlines the critical UA_4U motif. A star indicates the U:T mismatch at position 24. Primers and templates are indicated in capital letters. The incorporated nucleotides are represented by small letters.

appearing at longer reaction times. Fragments of 18−19 bases in length, the average distance between the RNase H and measured polymerase sites in the absence of substrate or during steady-state elongation (Kohlstaedt et al., 1992; Gopalakrishnan et al., 1992; Metzger et al., 1993), appear only for reaction times shorter than 30 s [see again Peliska and Benkovic (1992)]. In the presence of the four nucleotide substrates, full-length copies of the template were observed. Experiments monitoring RNase H degradation during regular synthesis indicate that the length between the end of the RNA/DNA duplex (polymerase site, P) and the prevalent RNA cut (RNase H site, R) is again 15 nucleotides (see Figure 7, lanes d and e in each panel, and scheme in Figure 7e). A fragment of 12 nt appears only for longer incubation times.

Patterns of progression of the RNase H activity were then monitored on different primer/template combinations as extended synthesis was taking place, in the presence of dTTP alone. According to the previous data, the position of an RNase H site at position *i* (lower part of the histogram, Figure 8) was used to locate the DNA polymerase site at *i*+15 (upper part of the histogram). On the less active hybrid (p24/R2-36), RNase H action degrades processively the whole template (Figure 7b, lanes f−k, and Figure 8a). A partial halt is observed when the mismatch is restored at position 24 (Figure 7a, lanes f−l, and Figure 7c, lanes f−j). After translation, the catalytic site is located mainly at A_{30} and A_{33} (p21/R2-36) or even at A_{27}-A_{28} [p18(G:C)/R2-36] (Figures 8b and 8c, respectively). On the mismatched hybrid p18(G:T)/R2-36, where the extended synthesis proceeds with maximal efficiency (see Figure 6c), the pattern of RNase H action remains quasi-stable (Figure 7d, lanes f−j). In this case the RNA sequence, onto which the catalytic site for retrotranscription appears to be positioned, is precisely the UA_4U motif (Figure 8d). Therefore, introduction of an additional mismatch at position 18 stops the processive RNase H activity. As a consequence, the most efficient extension is observed and the "consensus" RNA motif is maintained in proper register with the catalytic site.

DISCUSSION

The present experiments describe conditions where DNA polymerization by HIV-1 reverse transcriptase departs from its classical mode of elongation and "stutters" on DNA as well on RNA templates. This aberrant mode of synthesis prevails when a misincorporation is introduced and when millimolar concentrations of $MnCl_2$ are used. Divalent cations are known to deeply affect the accuracy of DNA polymerases (Beckman et al., 1985). In the present case manganese can be first chelated to the primer-template before enzyme and magnesium addition, suggesting that the primary role of this cation is to induce or to stabilize a particular geometry of the primer-template. Mn^{2+}, being a soft ion, binds preferentially soft ligands, like the heterocyclic nitrogen atoms of the bases, while inner sphere coordination of Mg^{2+}, a hard ligand, is more readily accomplished by the oxygen ligands of the ribose phosphate backbone. In particular, mismatches which expose heterocyclic nitrogen atoms to the

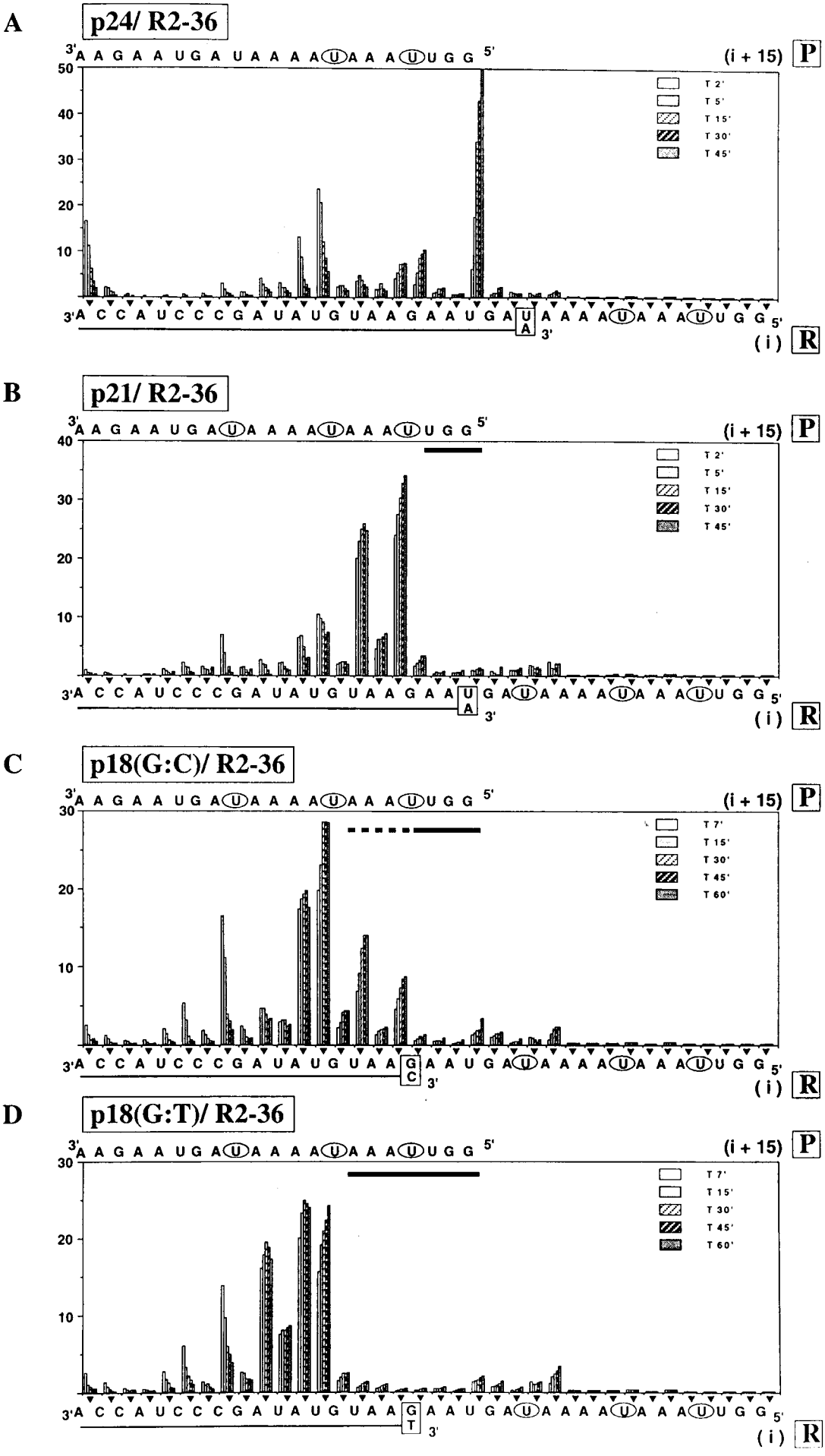
A
p24/ R2-36
B
p21/ R2-36
C
p18(G:C)/ R2-36
D
p18(G:T)/ R2-36
(i + 15)
P
(i)
R
T 2'
T 5'
T 15'
T 30'
T 45'
T 7'
T 60'

FIGURE 8: Positioning of the catalytic sites during extended synthesis. The pattern of RNase H activity of the wild-type HIV-1 RT on different hybrids is followed during extended synthesis (in the presence of dTTP) at the times indicated. Histograms display the intensity of the corresponding bands as shown in Figure 7. The degradation pattern, and the corresponding position of the catalytic site, is then compared to the respective elongation profile shown in Figure 6. The bottom sequence (positions *i* and boxed R for RNase H site) positions the actual RNase H degradation sites quantified from Figure 7. The upper sequence (positions *i*+15 and boxed P for polymerase site) positions the corresponding DNA polymerase sites after a translation by 15 nt. On both sequences the U residues located in the critical motif are circled. (a) RNase H degradation pattern observed on p24/R2-36 at the times indicated. The pattern is positioned with respect to the template sequence, written in the lower part of the histogram, and is read with respect to the upper sequence to locate the corresponding position of the catalytic site, P, involved in reverse transcription as extension proceeds. After this translation the highest cutting frequency takes place after the twenty-first residue, a uridine. It positions the catalytic site for polymerization at the end of the template (after the thirty-sixth nucleotide), as expected for run-off synthesis. (b) Processive RNase H degradation on p21/R2-36 monitored at the times indicated. RNase H action is halted three nucleotides upstream of the limit position observed in a. Accordingly, the catalytic site for polymerization is within the last UA_3U motif (upper sequence). (c) p18(G:C)/R2-36. The primer has been shortened by three nucleotides, and a thymidine residue, T, has been introduced by HIV-1 RT in front of position 21 (previously occupied by an adenine residue, A, in a and b). Extended synthesis occurs at a pace slightly faster than in b (results not shown). RNase H degradation is slowed down. During the 60 min of incubation, the catalytic site for reverse transcription appears to drift from UA_4U to UA_3U. (d) p18(G:T)/R2-36. RNase H degradation is halted at position 13 on the template sequence (lower part of the histogram). The prevalent cuts locate the DNA polymerization site on the upper sequence within the UA_4U motif during the 60 min of incubation (the horizontal bar below the upper sequence delimits the RNA region which is not yet accessible to reverse transcription according to our positioning of the catalytic site).

solvent are more likely stabilized by Mn^{2+} rather than by Mg^{2+} (Pan et al., 1993). Furthermore the nucleotide pools were highly biased, this being a requirement for observing this process under the present *in vitro* conditions. Indeed some *in vitro* assays performed in the absence of Mn^{2+} lead from time to time to sequence insertions which could also be due to the phenomenon analyzed here (Martinez et al., 1995).

This efficient mode of synthesis of polypyrimidine tracts is not due to blunt end or terminal nucleotidyl transferase addition. In sharp contrast with these processes, the present activity requires and copies a specific template sequence immediately following a critical site of mismatch formation. The sharp specificity of the reaction is also illustrated by the fact that HIV-1 and HIV-2 RTs are the only polymerases able to perform efficiently this aberrant mode of synthesis on the $Py(Pu)_nPy$ motives of the template. Most certainly, synthesis of products exceeding the template size have been previously reported using HIV-1 RT (Buiser et al, 1991; Andreola et al., 1993) and other DNA polymerases (Guieysse et al., 1995), but the analyses of the corresponding experiments have suggested a mechanism involving enzyme dissociation followed by a switch to a new template. Here, in contrast, addition of competitors only mildly affects the process by complexing only those RT molecules which have dissociated before entering the critical template sequence.

There is no significant difference between the process observed on DNA and on equivalent RNA templates except that reiterative synthesis can be more efficient in this second case and that the enzyme positioning is better defined.

If we pool together the observations made on those various templates, the data support an overall mechanism where the enzyme operates in a mode analogous to telomerases [Shippen-Lentz & Blackburn, 1990; Blackburn, 1993; and in more detail, Gilley and Blackburn (1996)]: a *commitment* step requiring Mn^{2+} takes place during the slow formation of a T:T (or U:T) mismatch and leads to a state of high processivity where misalignment between the two strands of the duplex is made easier (Figure 9, step 1). This is followed by an *elongation cycle* (step 2). Significant extension past the last base of the core polypurine motif cannot take place without strand slippage which in turn requires some DNA unwinding (step 3). Requirements for efficient extension, and for efficient resumption of polypyrimidine synthesis, where the enzyme has to rebind to a protruding primer are similar: presence of Mn^{2+} and presence of a limited core of purines flanked by two pyrimidines. They can be explained by the necessity of favoring unwinding during reshuffling of the primer. Un-

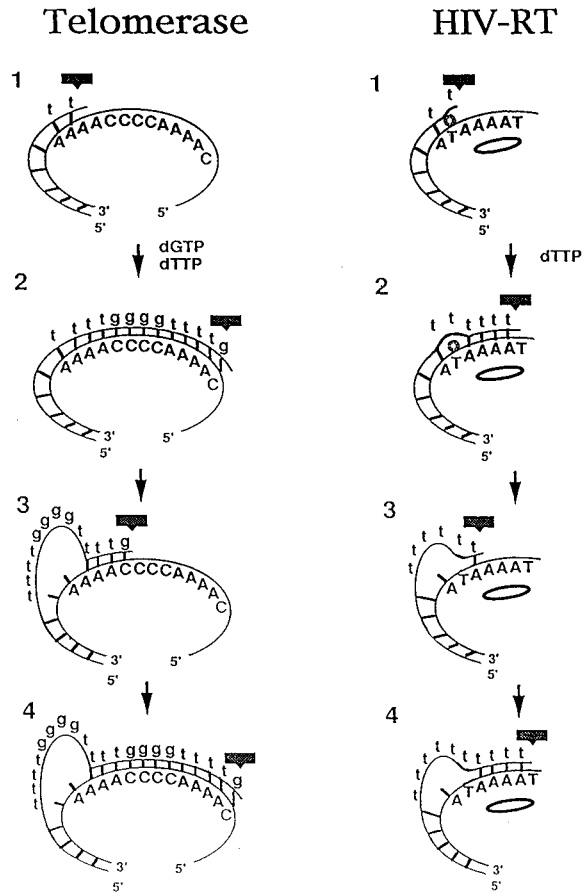

FIGURE 9: HIV-1 RT as a telomerase. Left panel. Mechanism postulated by Shippen-Lentz and Blackburn (1990) for *Euplotes* telomerase. Right panel. An equivalent mechanism proposed for extended synthesis by HIV-1 RT on a DNA template. Step 1, the first T:T mismatch commits the enzyme to a new mode (the star indicates a non-canonical primer/template pairing); step 2, copy of an A_n stretch and encounter with the second thymine residue of the template; initiation of backward strand slippage; step 3, repositioning of the catalytic site on an A residue of the template; step 4, new cycle of polymerization (analogous to step 2). Note that, as it extends the primer, the catalytic site (labeled by a box with arrowhead) passes back and forth on the enzyme domain (indicated by an oval) which holds the crucial template motif.

winding is easier if the length of the correctly paired template purine tract is limited, expecially when followed by a $_{3'}$Py-Pu$_{5'}$ step where the last nucleotide is mispaired. The potassium permanganate footprinting experiments indicate no permanent bulge upstream the crucial mismatch. They also fail to detect any new hot spot when the overhanging primer slides back. Therefore, the constant repositioning of the catalytic site likely occurs at random on any A motif of the consensus DNA sequence (see step 3, Figure 9). This interpretation is also supported by the unimodal Gaussian distribution of the extended products (see for example Figure 1b), a pattern obtained when all of the forward transit times are of the same order of magnitude (Fersht, 1985).

On RNA templates, the RNase H footprinting studies strongly suggest that an RT domain firmly grips the template consensus motif UA$_4$U onto which another domain, which carries the catalytic site and the primer, performs reiterative elongation. This implies a *dislocation* between two protein domains which are maintained in register during normal elongation. Such a dislocation is reminiscent of the occurrence of two different modes of dissociation at pause or termination (leading to free enzyme and hybrid) and during strand transfer, when RTs maintain a firm grip on their primer, and weaken their interaction with the 5′ end of the original template to accept the incoming terminal repeat sequence (Peliska & Benkovic, 1992), a reaction equivalent to a dissociation between primer/enzyme on the one hand and template on the other (a similar suggestion was made earlier by Greider in 1991). One can finally notice that the action of RNase H has isolated a PyPu$_n$Py sequence which bears similarity with consensus RNA motifs occurring at the active site of telomerases (Blackburn, 1993).

An extended synthesis lasting for more than 1 h and leading to products more than 150 nt in length implies an alternative mode of synthesis having a very long lifetime. Such alternate modes of polymerization have been previously described for other polymerases. Their existence has been documented for *E. coli* DNA polymerase I and for wheat germ RNA polymerase II; it has also been shown that these modes are sequence-specific (Papanicolau et al., 1986; de Mercoryol et al., 1990). The present observations are even more similar to those of Erie et al. (1993). These authors could not explain the formation of several mismatches in a row by *E. coli* RNA polymerase without invoking the existence of a branched pathway leading to a partition at each elongation step between a conformation able to elongate a mismatched RNA transcript and an unactivated state.

In general, we would like to suggest that alternate modes of synthesis which appear on specific sequences during elongation by polymerases (and the concommitant pattern of errors that they generate) are indicative of enzyme−nucleic acid conformations which are crucial at another stage of the polymerization cycle: for example, conformational changes in *E. coli* RNA polymerase have been documented during escape from abortive cycling (Carpousis & Gralla, 1980; Xiong & Reznikoff, 1993), to overcome pausing during elongation (Chamberlin, 1995; Nudler et al., 1995a), or as the enzyme encounters an oligo T tract which signals termination (Macdonald et al., 1993; Nudler et al. 1995b). The enzyme then moves forward by alternate smooth and inchworm-like movements (Chamberlin, 1995; Nudler et al., 1995a,b; Macdonald et al., 1993). In those cases also, these alternate states are invariably associated with a relaxed versus a strained conformation of the primer (here the RNA) or of the single-strand DNA template.

The easy expansion of repeat regions in several genes involved in human diseases could also be generated by DNA polymerases performing during elongation an operation expected to take place at another stage. Sequence expansions do not seem to be exceptionnally rare events: rather, *in vivo* data indicate that these anomalies have a high tendency to appear but that they are eliminated by a functional repair system (Strand et al., 1993; Johnson et al., 1995). Without exceptions, the "excess" number of repetitions which generates a pathological state arises on a sequence background which is already highly repeated [for review, see Ashley and Warren (1995)]. Current explanations of the formations of these repeats invoke extensive primer slippage during DNA replication. According to this model, repositioning of the primer and subsequent DNA synthesis takes place on a repeat unit located more upstream (Strand et al., 1993). This still hypothetical mechanism, however, does not account for the formation of the initial "core" unit of repetitions. We propose that, under specific polymerization conditions, eukaryotic DNA polymerases can also perform reiterative synthesis on specific template sequences. The number of units of this initial "core" can be further increased by the same mechanism or by slippage-mediated repositioning errors. By analogy with HIV-1 RT, where the "stuttering" behavior is activated by mismatch formation, it is probable that replication handicaps induce eukaryotic DNA polymerases to undergo these alternative modes of synthesis.

ACKNOWLEDGMENT

We thank Andrew Travers, Margaret Buckingham, and Peter von Hippel for scientific comments on a former version of the manuscript and Shahragim Tajbakhsh for critical reading.

REFERENCES

Andreola, M.-L., Dufour, E., Tarrago-Litvak, L., Jamkovoy, V. I., Levina, A. S., Barr, P. J., Litvak, S., & Nevinsky, G. A. (1993) *Biochim. Biophys. Acta 1173*, 147−155.

Ashley, C. T., & Warren, S. T. (1995) *Annu. Rev. Genet. 29*, 703−728.

Bebenek, K., & Kunkel, T. A (1990) *Proc. Natl. Acad. Sci. U.S.A. 87*, 4946−4950.

Beckman, R. A., Mildvan, A. S., & Loeb, L. L. (1985) *Biochemistry 24*, 5810−5817.

Blackburn, E. H. (1993) in *Reverse Transcriptase* (Skalka, A. M., & Goff, S. P., Eds) pp 411−424, Cold Spring Harbor Laboratory Press, Plainview, NY.

Buiser, R. G, DeStefano, J. J., Mallaber, L. M., Fay, P. J., & Bambara, R. A. (1991) *J. Biol. Chem. 266*, 13103−13109.

Carpousis, A. J., & Gralla, J. D. (1980) *Biochemistry 19*, 3245−3253.

Caskey, C. T., Pizzuti, A., Fu, Y.-H., Fenwick, R. G., Jr., & Nelson, D. L. (1992) *Science 256*, 784−789.

Chamberlin, M. J. (1995) in *The Harvey Lectures*, Vol. 88, pp 1−21, Wiley-Liss, New York.

Clark, J. M. (1988) *Nucleic Acids Res. 20*, 9677−9685.

Creighton, S., & Goodman, M. F. (1995) *J. Biol. Chem. 270*, 4759−4774.

de Mercoyrol, L., Soulié, J. M., Job, C., Job, D., Dussert, C., Palmari, J., Rasigni, M., & Rasigni, G. (1990) *Biochem. J. 269*, 651−658.

Erie, D. A., Hajiseyedjavadi, O., Young, M. C., & von Hippel, P. H. (1993) *Science 262*, 867−873.

Fersht, A. (1985) in *Enzyme Structure and Mechanism*, 2nd ed., pp 47−50, 377−378, W. H. Freeman and Company, New York.

Gilley, D., & Blackburn, E. H. (1996) *Mol. Cell. Biol. 16*, 66−75
Gopalakrishnan, V., Peliska, J. A., & Benkovic, S. J. (1992) *Proc. Natl. Acad. Sci. U.S.A. 89*, 10763−10767.
Götte, M., Fackler, S., Hermann, T., Perola, E., Cellai, L., Gross, H. J., Le Grice, S. F. J., & Heumann, H. (1995) *EMBO J. 14*, 833−841.
Greider, C. W. (1991) *Mol. Cell. Biol. 11*, 4572−4580.
Guieysse, A. L., Praseuth, D., Grigoriev, M., Harel, A., & Helene, C. (1995) *Biochemistry 34*, 9193−9199.
Isel, C., Lanchy, J. M., Le Grice, S. F., Ehresmann, C., Ehresmann, B., & Marquet, R.(1996) *EMBO J. 15*, 917−924.
Johnson, R. E., Kovvali, G. K., Prakash, L., & Prahash, S. (1995) *Science 269*, 238−240.
Kati, W. M., Johnson, K. A., Jerva, L. F., & Anderson, K. S. (1992) *J. Biol. Chem. 267*, 24607−24613.
Kohlstaedt, L. A., Wang, J., Friedman, J. M., Rice, P. A., & Steitz, T. A. (1992) *Science 256*, 1783−1790.
Kornberg, A., & Baker, T. A. (1992) in *DNA Replication*, pp 223−225, W. H. Freeman and Company, New York.
Kunkel, T. A. (1990) *Biochemistry 29*, 8003−8011.
Le Grice, S. F. (1993) in *Reverse Transcriptase* (Skalka, A. M., & Goff, S. P., Eds.) pp 163−191, Cold Spring Harbor Laboratory Press, Plainview, NY.
Macdonald, L. E., Zhou, Y., & McAllister, W. T. (1993) *J. Mol. Biol. 232*, 1030−1047.
Martinez, M. A., Sala, M., Vartanian, J.-P., & Wain-Hobson, S. (1995) *Nucleic Acids Res. 23*, 2573−2578.
Metzger, W., Hermann, T., Schatz, O., Le Grice, S. F., & Heumann, H. (1993) *Proc. Natl. Acad. Sci. U.S.A. 90*, 5909−5913.
Nudler, E., Goldfarb, A., & Kashlev, M. (1995a) *Science 265*, 793−796.
Nudler, E., Kashlev, M., Nikiforov, V., & Goldfarb, A. (1995b) *Cell 81*, 351−357.
Pan, T., Long, D. M., & Uhlenbeck, O. C. (1993) in *The RNA World* (Gesteland, R. F., & Atkins, J. F., Eds.), pp 271−302, Cold Spring Harbor Laboratory Press, Plainview, NY.
Papanicolau, C., Lecomte, P., & Ninio, J. (1986) *J. Mol. Biol. 189*, 435−448.
Patel, P. H., & Preston, B. D. (1994) *Proc. Natl. Acad. Sci. U.S.A. 91*, 549−553.
Peliska, J. A., & Benkovic, S. J. (1992) *Science 258*, 1112−1118.
Perrino, F. W, Preston, B. D., Sandell, L. L, & Loeb, L. A. (1989) *Proc. Natl. Acad. Sci. U.S.A. 86*, 8343−8347.
Ricchetti, M., & Buc, H. (1990) *EMBO J. 9*, 1583−1593.
Ricchetti, M., & Buc, H. (1993) *EMBO J. 12*, 387−396.
Ripley, L. S. (1990) *Annu. Rev. Genet. 24*, 189−213.
Sasse-Dwight, S., & Gralla, J. D. (1989) *J. Biol. Chem. 264*, 8074−8081.
Shippen-Lentz, D., & Blackburn, E. H. (1990) *Science 247*, 546−552.
Strand, M., Prolla, T. A., Liskay, R. M., & Petes, T. D. (1993) *Nature 365*, 274−277.
Thibodeau, S. N., Bren, G., & Schaid, D. (1993) *Science 260*, 816−819.
Van Beveren, C., Coffin, J., & Hughes, S. (1985) in *RNA Tumor Viruses*, Vol. 2, (Weiss, R., Teich, N., Varmus, H., & Coffin J., Eds.) p 1110, Cold Spring Harbor Laboratory Press, Plainview, NY.
Vartanian, J.-P., Meyerhans, A., Asjö, B., & Wain-Hobson, S. (1991) *J. Virol. 65*, 1779−1788.
Xiong, X. F., & Reznikoff, W. S. (1993) *J. Mol. Biol. 231*, 569−580.

BI961274V