The EMBO Journal vol.12 no.2 pp.387-396, 1993

E.coli DNA polymerase I as a reverse transcriptase

Miria Ricchetti and Henri Buc

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

Communicated by H.Buc

The ability of *Escherichia coli* DNA polymerase I to retrotranscribe an RNA template was examined under steady-state conditions, using a primer extension assay which allows determination of kinetic constants on well-defined heterogeneous sequences. Equilibrium and rate constants for the initial binding step of the enzyme to two homologous DNA and RNA templates do not show striking differences. In both cases, under steady-state conditions, processivity limits the maximal velocity of the translocation process. The lower catalytic efficiency of the enzyme when it operates on RNA is then reflected by a 100-fold greater apparent average Michaelis constant for the deoxynucleotide substrates. We conclude that *E.coli* DNA polymerase I effectively transcribes both templates, its performances being limited in both cases by its intrinsically low processivity. Furthermore, DNA polymerase I is a strikingly accurate enzyme when operating on RNA. Magnesium has to be substituted by manganese so that a pattern of errors could be detected. This great accuracy results from a combination of factors. The 3′ to 5′ exonuclease activity is still operating but in a non-discriminative manner. Elongation of a mismatched primer terminus is markedly impaired. The forward polymerization rate of incorporation of an incorrect deoxynucleotide must be extremely low, when Mg^{2+} is present. In summary *E.coli* DNA polymerase I preserves its main characteristics when retrotranscribing RNA.

Key words: DNA polymerase I/fidelity/polymerase evolution/reverse transcriptase/template specificity

Introduction

Escherichia coli polymerase I is the most extensively studied DNA polymerase up to now. Its enzymatic mechanism as well as its 3′→5′ and 5′→3′ exonuclease activities have been investigated by a variety of kinetic and genetic approaches (for a recent review, see Carroll and Benkovic, 1991; cf. also Echols and Goodman, 1991). In addition structural data are available for the Klenow fragment, the large C-terminal domain of the protein which lacks the 5′→3′ exonuclease activity removed by proteolysis with the N-terminal domain (Ollis *et al.*, 1985; Beese and Steitz, 1989).

Following earlier qualitative reports (Lee-Huang and Cavalieri, 1964; Chamberlin, 1965; Goodman and Spiegelman, 1971; Karkas *et al.*, 1972; Wells *et al.*, 1972), Loeb *et al.* demonstrated in 1973 that Pol I can also copy synthetic as well as natural RNAs. In the original manuscript describing this observation, synthesis performed on 28S RNA was compared with that observed on calf thymus DNA, activated or not through nicking. The activity was of the order of 0.1−0.4% of that found on activated calf thymus DNA, and 6% of the activity measured on non-activated calf thymus DNA. When synthetic homopolymeric RNA templates were used, the rate of DNA synthesis varied from 20 to 120% of that observed on homologous DNA matrices.

Other DNA polymerases can also copy RNA into DNA *in vitro* with apparently more specific requirements for the sequence of the template. Nuclear DNA polymerase β (Pol β) involved in DNA repair accepts only poly(rA)·oligo(dT) (Chang, 1974). The mitochondrial DNA polymerase γ accepts the same template with a higher efficiency than Pol β while it fails to utilize poly(rC)·oligo(dG) and natural RNA templates (Gallo *et al.*, 1971). Polymerases α do not normally accept any RNA template with the exception of polyriboadenylic acids under specific reaction conditions and with a relatively low efficiency (Masaki and Yoshida, 1978; Yoshida *et al.*, 1981).

In vivo, copying RNA into DNA is a prerogative of reverse transcriptases, the enzymes specialized in the replication of retroviruses. Efficient bifunctionality of reverse transcriptases is thought to result from a decreased template selectivity of the single catalytic site at which both polymerization reactions occur. As they are also devoid of a 3′→5′ exonuclease activity, their fidelity is rather poor.

It is therefore of interest to conduct parallel studies with well-characterized reverse transcriptases and with the least stringent of all DNA polymerases mentioned above, Pol I, in order to answer two questions. (i) To what extent is the ability to synthetize DNA from RNA comparable with the canonical DNA dependent DNA polymerase activity for *E.coli* DNA Pol I (in terms of template selectivity, catalytic activity and processivity)? (ii) Which features are lacking to make DNA polymerase I a really efficient reverse transcriptase (by comparing those constants with those obtained on similar templates with HIV-1 reverse transcriptase)? For this purpose we investigate here a number of kinetic parameters of Pol I using as templates short heteropolymeric RNAs using a classical methodology (cf. McClure and Jovin, 1975; Bambara *et al.*, 1976; Mizrahi *et al.*, 1986). The kinetic parameters were, whenever possible, compared with those obtained on homologous DNA templates. Our kinetic studies are intended to complement structural data obtained on these two enzymes (Ollis *et al.*, 1985; Arnold *et al.*, 1992; Kohlstaedt *et al.*, 1992). We hope they will help to specify the evolutionary relationship between those two modes of replication.

Results

Our basic assay consists of measuring the rate of elongation of a DNA primer, hybridized to the complementary sequence

© Oxford University Press

in the template, by DNA polymerase I or by its Klenow fragment.

The experiments have been performed on two homologous DNA and RNA templates (called D2 and R2 respectively) of identical sequence. Each template (either RNA or DNA) was hybridized with the appropriate DNA primer (called PG5); the nomenclature PG5/R2 therefore represents a hybrid formed between a DNA primer and an RNA template while PG5/D2 is a DNA−DNA hybrid. Two other primers which can hybridize up to position +2 of the preceding templates were synthetized terminated by a matched G_t:C (PG5TC) or a mismatched G_t:T (PG5TT) base pair (cf. Figure 1).

Polymerization was initiated by adding the enzyme to the hybrid, incubated for a given time at 25°C, and one or more deoxynucleotides. In the case of PG5/D2 or PG5/R2 the addition of the first complementary deoxynucleotide allows just one elongation step, while addition of three complementary deoxynucleotides was enough to yield full length DNA chains.

Run-off assays

Those matrices were first used in run-off experiments to ascertain average apparent Michaelis constants, $\bar{K}_m$, and average maximal velocities, $\bar{V}_{max}$, per elongation step as explained in Materials and methods. Enzyme (in excess) was preincubated with the RNA or the DNA hybrid for 30 s at 25°C. Two media of different ionic composition were selected to complement the studies on the corresponding rates of misincorporation (see below): KCl 50 mM, with either 6 mM $MgCl_2$ or 1 mM $MnCl_2$, in Tris−HCl 50 mM (pH 7.8).

```
5' GGTTAAATAAAATAGTAAGAATGTATAGCCCTACCA 3'  D2
        3' TACATATCGGGATGGT 5'  PG5
   3' CTTACATATCGGGATG 5'  PG5TC
   3' TTTACATATCGGGATG 5'  PG5TT
```

Fig. 1. Main DNA template (first line) and primers (lines 2−4) used in this study (cf. also Figure 6). The R2 template has an RNA sequence homologous to D2.

As shown in Figure 2, and in Table I, the main difference in these two types of experiment relies on the average apparent Michaelis constants for deoxynucleotide incorporation, a value clearly less favourable when using an RNA template. The average maximal rates are of the same order of magnitude and far lower than those measured for a single step on DNA (cf. Carroll and Benkovic, 1991) suggesting that enzyme recycling due to low processivity is often taking place. Values specific for the first step can also be obtained from the rate of depletion of the initial primer substrate as a function of deoxynucleotide concentration (cf. Materials and methods). Apparent Michaelis constants for dTTP of 25 nM (for DNA) and 1 μM (for RNA) were obtained for the same maximum velocity in 6 mM Mg^{2+}.

As demonstrated by the gel assay, intermediates accumulate at some specific steps of the elongation process, notably at position +2 on the RNA template. The variability of the local constants on the present sequence is however, not very large. We have computed $\bar{V}_{max}$ and $\bar{K}_m$ for different incubation times (changing therefore the weight of early versus late steps) and have not observed changes larger than 20% in those average values. The value of the overall catalytic efficiency ($\bar{V}_{max}/\bar{K}_m$) should be conserved when recycling is not limiting. As shown in Table I, it is decreased by ~100 when DNA is replaced by RNA.

Association constants with DNA and RNA templates

Pol I was incubated at 17 nM with various concentrations of DNA−DNA or RNA−DNA hybrids. The reaction was started by addition of a single deoxynucleotide allowing

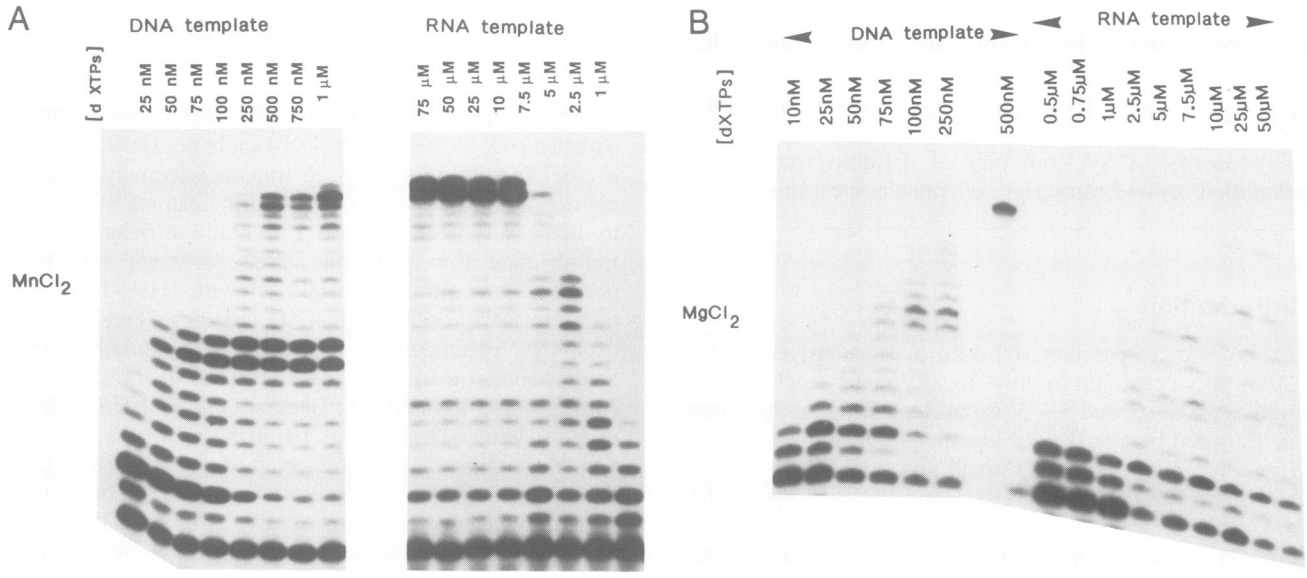

Fig. 2. Gel assay for the incorporation of consecutive nucleotides on PG5/D2 and on PG5/R2. 70 nM of Pol I were incubated with 3 nM of the indicated template−primer in the presence of increasing amounts of the four dNTPs. Reaction terminated after 15 s of incubation for PG5/D2 and after 30 s of incubation for PG5/R2. (**A**) reaction performed in the presence of 1 mM $MnCl_2$; (**B**) reaction performed in the presence of 6 mM $MgCl_2$. In this figure, degradation products (mainly corresponding to one excision step) are not shown. Their amount was always <15% of the total quantity of labelled primer.

elongation of the primer by one base only. Figure 3a shows that when the template is in excess typical biphasic curves are observed, extrapolating at zero time on the ordinate at a concentration which was taken as the concentration of the preformed enzyme-template. The time course of the reaction does not differ for two different matched DNA primers PG5/D2 and PG5TC/D2 when used at the same concentration. The same type of experiment was repeated in the presence of a DNA competitor (120 mg/ml); the initial burst could then be clearly defined (Figure 3b). Those conditions were therefore selected to define dose–response curves at increasing enzyme concentrations. From those curves, equilibrium dissociation constants respectively equal to 18 ± 3 nM (DNA template) and to 70 ± 8 nM (RNA template; average of two determinations) were obtained in 6 mM $MgCl_2$ (data not shown). These values were consistent with the amplitude of the initial burst observed in experiments similar to those shown in Figure 3a.

Residence times

Complexes of DNA Pol I and PG5/R2 were equilibrated in the same buffer as above. Dissociation of the complexes was monitored in the presence of competing calf thymus DNA (120 μg/μl) which traps dissociating enzyme thereby preventing its reassociation with the template. Preformed complexes were exposed to calf thymus DNA for times ranging between 5 s and 5 min before the polymerization reaction was started. At the standard salt concentration the complex decays exponentially. The measured residence time, 22 ± 3 s, depends neither on the concentration of competing DNA, nor on the dTTP concentration used to monitor the amount of residual binary complex. This value has to be compared with a residence time of 25 ± 3 s found in a parallel assay with PG5/D2 (Figure 4).

The steady-state rate of elongation of the primer (second phase in Figure 3a) can be characterized by a time constant τ_c, corresponding to the duration of one cycle (dissociation

Table I. Elongation of the PG5/D2 and of the PG5/R2 hybrids: average kinetic parameters

Template	Divalent ion	$\bar{K}_m$ (μM)	$\bar{V}_{max}$ (nt/s)	$\left[\frac{\bar{V}_{max}}{\bar{K}_m}\right]$ nt/μM.s	$<i>$
DNA	$MgCl_2$	0.66	1.43	2.2	4.0
RNA	$MgCl_2$	25	0.71	2.8×10^{-2}	76 2.2
DNA	$MnCl_2$	0.1	0.74	7.4	4.2
RNA	$MnCl_2$	28.6	1.25	4.4×10^{-2}	120 3

The nature of the template is given in the first column. The nature of the divalent ion (respectively 6 mM for Mg^{2+} and 1 mM for Mn^{2+}) in the second one. Apparent average Michaelis constants determined through run-off assays, $\bar{K}_m$, and corresponding maximal velocities (in nucleotides per second), $\bar{V}_{max}$, are listed in columns 3 and 4; the catalytic efficiency $\bar{V}_{max}\bar{K}_m$ and the ratio of catalytic efficiencies (DNA versus RNA) in column 5. $<i>$ given in column 6 is the average number of steps observed before dissociation and is an index of processivity. Standard deviations for values given in columns 3, 4 and 6 are of the order of 20%.

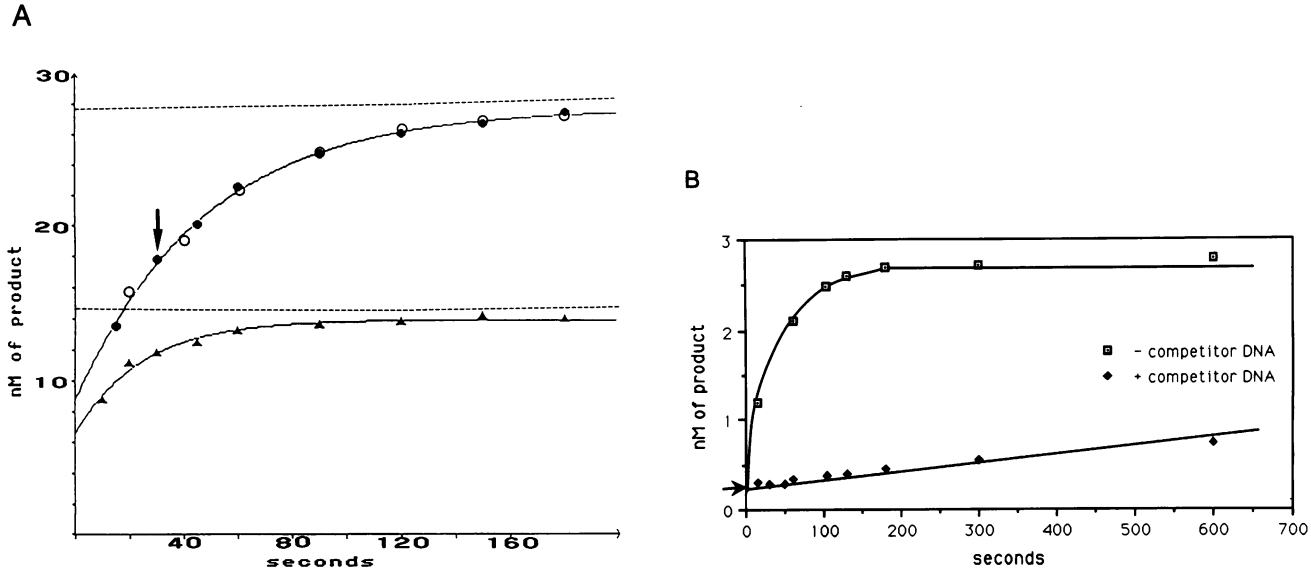

Fig. 3. (**A**) Time course of elongation of a DNA primer by one nucleoside monophosphate. Experimental conditions are as described in Materials and methods with the following specifications: ●, PG5/D2 30 nM; ○, PG5TC/D2 30 nM; ▲, PG5/D2 15 nM. Incubation time (with 70 nM enzyme): 60 s in the magnesium containing buffer. dTTP is added at the final concentration of 10 μM. Extrapolation of the time course of the steady-state behaviour to zero time (via an exponential fit) leads to an intercept with the ordinate indicative of a rapid burst of synthesis. It corresponds to the amount of preformed enzyme–hybrid complex. The arrow indicates the time required to perform one cycle of synthesis during the steady state. (**B**) Time course of the incorporation of a single deoxynucleotide on PG5/R2 in the presence or absence of competitor calf thymus DNA (final concentration: 120 μg/ml; in buffer containing 6 mM $MgCl_2$) of 3 nM of primer hybridized with the template with 8 nM of active enzyme. Reaction started with 10 μM of dTTP, with or without the competitor.

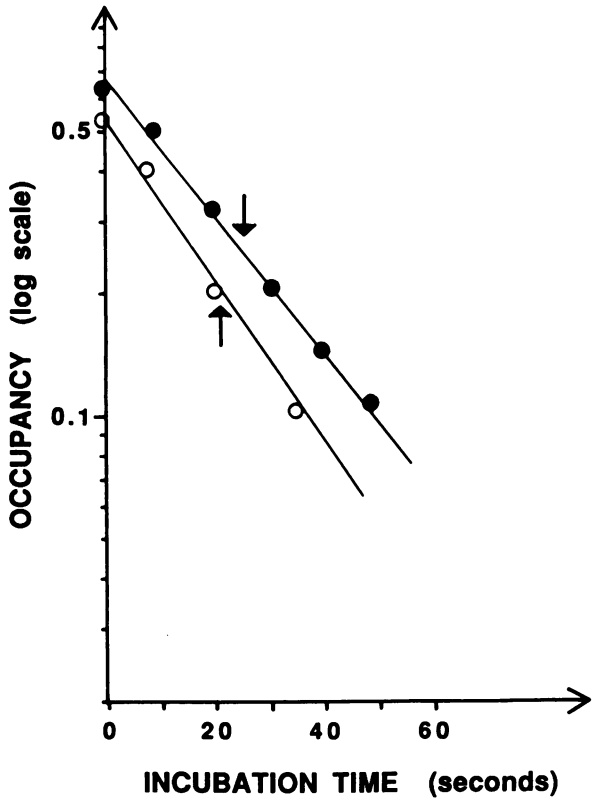

Fig. 4. Residence times of Pol I on PG5/D2 (●) or on PG5/R2 (○). Enzyme (70 nM) was incubated with 3 nM of the hybrid in the same buffer as in Figure 3. At zero time, 120 μg/ml of competitor calf thymus DNA were added and incubated for a time *t* (given on the abscissa), before addition of the dTTP substrate (120 nM for DNA, 10 μM for RNA) for 15 s. The amount of synthesis was taken as proportional to the amount of enzyme present on the hybrid at the site to be elongated at time *t*. A semi-log plot of this occupancy (given here in arbitrary units) against incubation time yields the residence times for the two complexes (indicated by two arrows).

of the enzyme from position +1 followed by reassociation at position 0, cf. Materials and methods). When enzyme is in excess over the equilibrium constant, this step is limited by the dissociation occurring at +1. A value of 30 ± 2 s was found for the upper curves of Figure 3a, showing an overall consistency of the data.

Processivity

Run-off assays were conducted in parallel in the presence and in the absence of the competitor used for measuring the residence time. The competitor was added at the same time as the deoxynucleotide substrates and the average number of elongation steps performed by the enzyme $<i>$ was computed in both cases (cf. Materials and methods); weak processivity was observed for those intermediates which accumulate during the steady-state assay performed in the absence of competitor. After 10 s of synthesis, the pattern observed in the presence of competitor is stable; the average processivity which is computed is of the same order of magnitude for both templates (cf. Table I).

Fidelity of incorporation by Pol I in the presence of Mg^{2+}

When replicating DNA, Pol I can introduce incorrect deoxynucleotides into the product chain with frequencies varying from 10^{-3} to 5×10^{-5} (Kuchta *et al.*, 1988; Eger *et al.*, 1991). When the 3′→5′ exonuclease activity is operating, this values drops by a factor averaging 30 (Eger *et al.*, 1991). Using the same assay, we asked whether the enzyme was able to incorporate incorrect deoxynucleotides on RNA templates, at what rate and whether the exonuclease activity was operating. Eighteen different positions were investigated along the RNA template for the incorporation of any non-complementary nucleotide. Reactions were performed as described in the legend to Figure 5 in the presence of the incorrect deoxynucleotide but in the absence of the correct one. After preincubation of the enzyme with the hybrid, the reaction was started by addition of the incorrect deoxynucleotide at high concentrations (10 μM, 100 μM and 1 mM) and was stopped at different times (15 s, 30 s, 1 min, 2 min and 5 min). A single mismatch (probably G_t:T) occurring with a very low error rate ($\sim10^{-6}$) was found among the 53 combinations tested (cf. lane 5 Figure 5a). The control experiment showed the expected insertion of the complementary deoxynucleotide.

The unexpected fidelity of Pol I on RNA templates was confirmed when the experiment was performed with the Klenow fragment. No misincorporation was visible throughout the 27 combinations tested (data not shown). Poly(rA)·oligo(dT) was also used as a template in the presence of either dGTP or dCTP or dATP under the conditions described above for the heteropolymeric template. In this case again and in agreement with a previous analysis performed by Travaglini *et al.* (1975), no misincorporation was detected.

The order of magnitude for the threshold of the detectable error rate on RNA templates, f*, can be computed knowing the sensitivity of our detection assay for the incorporation of an incorrect deoxynucleotide (the maximum concentration of the substrate used was 1 mM for an incubation time of 5 min) and the value of the ratio $\bar{V}_{max}/\bar{K}_m$ for the corresponding correct deoxynucleotide (cf. Ricchetti and Buc, 1990). This threshold is equal to 5×10^{-7}. The detection of a single mismatch above this value in our assay stresses the great accuracy of the polymerization process on our RNA templates.

An equivalent set of positions was tested in parallel on DNA templates of identical sequence (compare first and third boxes in Figure 6). Here 28 misincorporations, among 47 combinations tested, were clearly visible. The deoxynucleotide concentration was, in certain cases, lowered by a factor of 100 to adopt a threshold value f* equivalent to the one valid for RNA. The five misincorporations tested under these conditions were still clearly detectable.

Misincorporation ability of Pol I in the presence of Mn^{2+}

The metal ion Mn^{2+} is known to increase the error rate of Pol I on homopolymeric DNA templates, mainly affecting the $\bar{V}_{max}$ value for the incorporation of the incorrect deoxynucleotide (El-Deiry *et al.*, 1988). We substituted 6 mM $MgCl_2$ by 1 mM $MnCl_2$, the optimal concentration of this salt for polymerization according to a previous screening (data not shown). We repeated our assay on two RNA templates poly(rA)·oligo(dT) and on PG5/R2 for some of the positions tested above. As shown in Figures 5 and 6, the misincorporation ability of Pol I was dramatically increased. As we had previously checked in the run-off assay

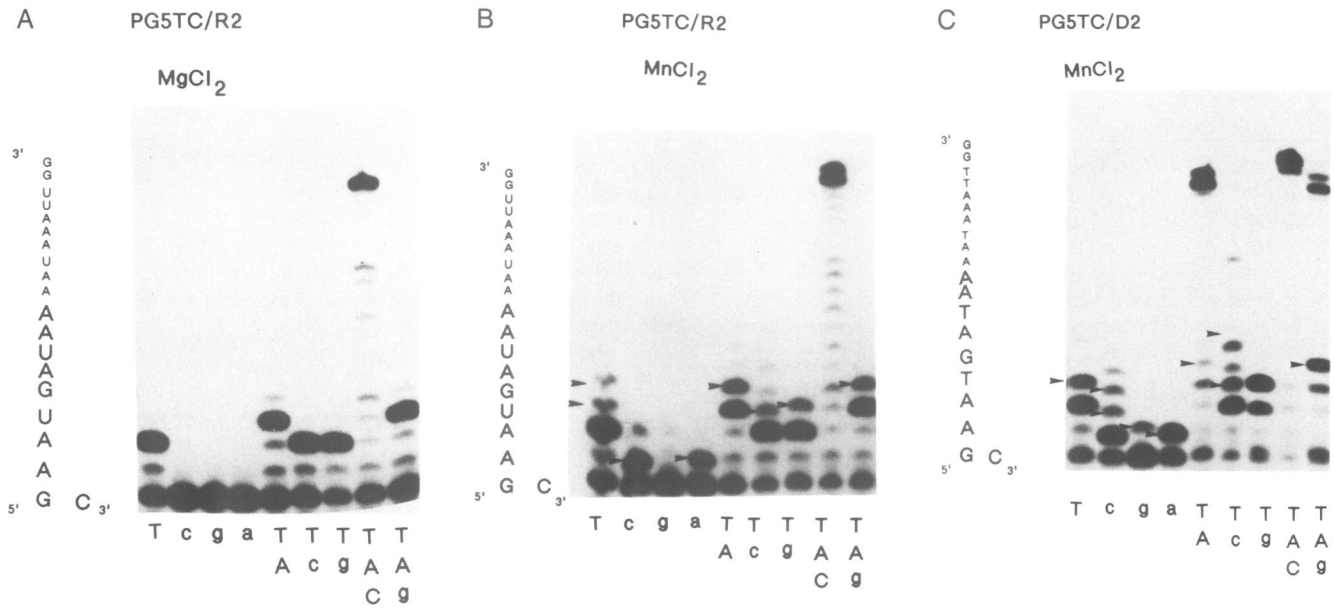

Fig. 5. Gel assay for deoxynucleoside incorporation. 35 nM of Pol I were incubated with 3 nM of the DNA−DNA or RNA−DNA hybrids in the presence of indicated deoxyribonucleotide for 5 min at 25°C. The template sequence and the nucleoside at the 3′ end of the primer are shown on the left. The added deoxynucleotides (1 mM) are shown at the bottom: capital letters represent nucleotides which correctly pair with the template while lower case letters represent nucleotides which lead to misincorporation. Arrow shows a position where misincorporation takes place. (**A**) PG5/R2 in 6 mM $MgCl_2$ containing buffer; (**B**) PG5/R2 in 1 mM $MnCl_2$ containing buffer; (**C**) PG5/D2 in 1 mM $MnCl_2$ containing buffer.

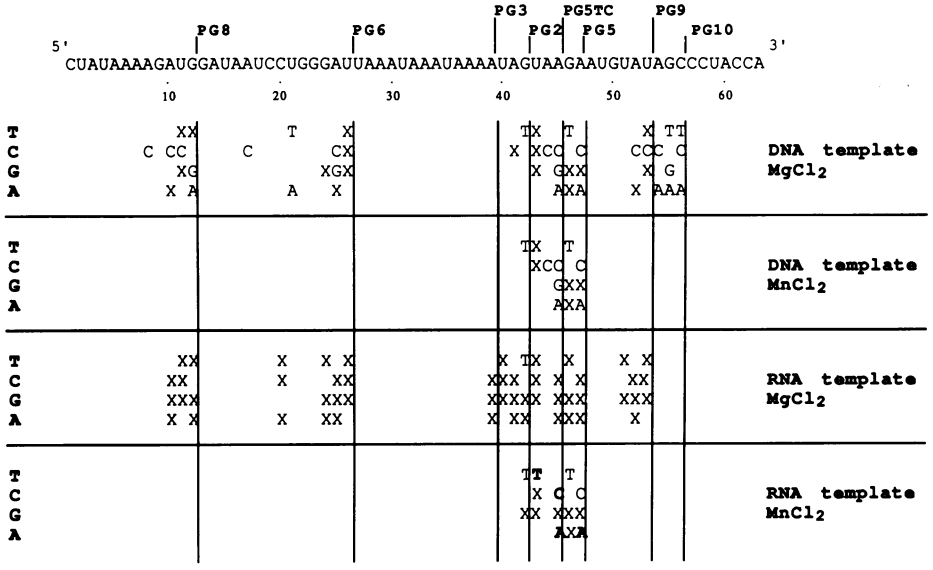

Fig. 6. Misincorporation patterns on DNA and RNA templates in the presence of either $MgCl_2$ or $MnCl_2$. In the upper part, the template sequence is shown (in its RNA form). Arrows indicate the 3′ end of the respective oligonucleotide primers. A capital letter represents a base which is incorrectly incorporated at the indicated position. A cross stands for the absence of incorporation at that position: on the left is indicated the nature of the substrate tested (T, C, G, A stand for dTTP, dCTP, dGTP and dATP, respectively). For the assay performed in $MnCl_2$ (RNA template) mismatches which cannot be further extended in the assay (lack of complementary nucleotide) are indicated in bold characters.

that substitution of Mg^{2+} by Mn^{2+} did not drastically affect the kinetic process for incorporation of correct deoxynucleotides (cf. Table I), we conclude that Mn^{2+} increases markedly the error rates on RNA templates as it does on DNA templates.

Excision of a mismatch

Effective misincorporation results from the balance between three factors: the incorporation of a wrong nucleotide, its non-excision by the 3′→5′ exonuclease activity and the ability of the enzyme to extend the nascent chain downstream of the mismatch. In this respect it was important to assess first the ability of Pol I to excise a mismatch created on an RNA template and second the rate of extension of a terminal mismatch.

Pol I excision ability was compared on two kinds of hybrids, a perfectly matched primer (PG5TC/R2 or PG5TC/D2), or a primer having a terminal mismatch

Table II. Rates of excision of 3′ primer termini by Pol I

Hybrid	Terminal base pair	k_{exo} (s^{-1})		d	
		$MgCl_2$	$MnCl_2$	$MgCl_2$	$MnCl_2$
PG5TC/D2	G_t:C	6×10^{-3}	1.2×10^{-2}	23	17
PG5TT/D2	G_t:T	1.4×10^{-1}	2.0×10^{-1}		
PG5TC/R2	G_t:C	3.3×10^{-3}	1.8×10^{-2}	1	2
PG5TT/R2	G_t:T	3.3×10^{-3}	3.6×10^{-2}		

Standard assay conditions (without addition of a deoxynucleotide). Values of k_{exo} are determined from the slope of a semi-logarithmic plot of the concentration of the annealed hybrid versus time. d is the ratio of the excision rates of mismatched versus matched base pair termini.

(PG5TT/R2 or PG5TT/D2). (As shown in Figure 6, this misincorporation occurs by elongating the PG5/D2 hybrid in the presence of Mg^{2+} or Mn^{2+}; it occurs weakly, and in $MnCl_2$ only, during PG5 elongation on the R2 template). In the presence of 6 mM $MgCl_2$, the degradation profiles follow an exponential course and go to completion. The corresponding rates are given in Table II. For DNA template, discrimination occurs in favour of the mismatched hybrid (k_{exo} respectively equal to 1.4×10^{-1} s^{-1} and 6×10^{-3} s^{-1} resulting in a discrimination factor d of 23). On the RNA templates, excision occurs more slowly and discrimination is lost ($k_{exo} = 3.3\times10^{-3}$ s^{-1} in both cases; cf. Table II).

Experiments were repeated in 1 mM $MnCl_2$. In this case, the reaction does not go to completion; estimates of the rate constants are taken from the initial slopes of the degradation profile and are subjected to larger errors. But the results are qualitatively the same, discrimination is practically lost on the RNA template, and the limited amount of excision which is observed occurs rather slowly (cf. Table II again).

In summary, the fidelity of Pol I on an RNA template in $MgCl_2$ is not due to a specific and efficient excision of a newly created RNA–DNA mismatch, and the reappearance of an error pattern in the presence of Mn^{2+} is not due to a drastic loss of this proofreading capacity.

Propagation of a mismatch

The rate of extension of a terminal mismatch on a DNA template has been evaluated for Pol I (Kuchta *et al.*, 1988). The incorporation rate is strongly affected when compared with the one occurring after a correct pairing (K_m is 15 times higher and k_{cat} is reduced by 3500-fold). We compare here the process of extension of homologous terminal mismatches created on an RNA versus a DNA template.

In a first set of experiments the mismatched hybrid PG5TT/R2 was incubated with Pol I in the presence of high concentrations of the next complementary deoxynucleotide (dTTP 1 mM, which is ~20-fold higher than the corresponding apparent K_m value) at 25°C for times ranging from 15 s to 10 min. Under these conditions, and in the presence of $MgCl_2$, no extension of the hybrid was visible. In the presence of $MnCl_2$, only 4% of the molecules were elongated in 15 s. This amount did not increase after 10 min. The control experiment using the correctly paired PG5TC/R2 showed the expected elongation rate. The other control experiment, performed on PG5TT/D2 showed also extension of the hybrid in the presence of both $MgCl_2$ and $MnCl_2$. We conclude that this RNA–DNA mismatch cannot be extended. It was

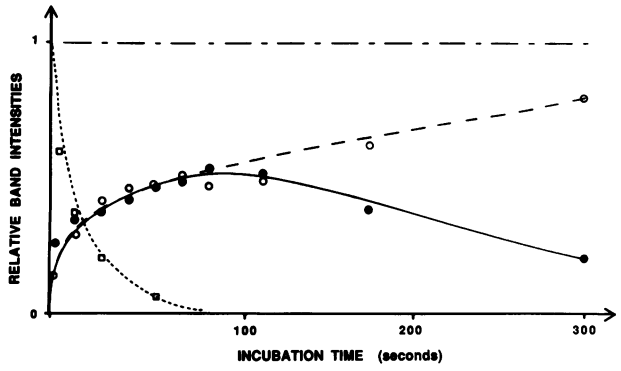

Fig. 7. Synthesis and extension of an RNA–DNA mismatch by Pol I. In our standard buffer complemented with 1 mM Mn^{2+}, a complex was formed between Pol I (70 nM) and the PG5/R2 hybrid (3 nM). Reaction was started by addition of 1 mM dCTP and the concentrations of the original hybrid and of the +1 and +2 products followed as a function of time. They are plotted here after normalization for the total amount of enzyme–hybrid initially formed (original hybrid, □; +1, ●; +2, ○).

introduced at a position which corresponded to a marked pause site when faithful incorporation was followed during run-off experiments (cf. Figures 2a and b). We suspected therefore that Pol I had an even stronger tendency to dissociate when a terminal mismatch was created on an RNA–DNA hybrid. This irreversible dissociation would account for the plateau of incorporation observed in $MgCl_2$ and the absence of elongation in $MnCl_2$.

A second set of experiments, done with the PG5/R2 hybrid in the presence of $MnCl_2$ strengthened this hypothesis. This hybrid, M_0, is perfectly matched. Addition of dCTP alone allows creation of a strong A_t:C mismatch, m_1, which can eventually be followed by the correct pairing G_t:C (M_2). In this case, as shown in Figure 7, depletion of the original complex EM_0 to form m_1 takes place at a rate $k_1 = 0.055$ s^{-1}. Appearance of the consecutive products is biphasic and does not correspond to a simple scheme of two consecutive reactions A → B → C. In the first phase of the reaction (incubation times <70 s) the m_1 and M_2 products appear at similar rates. Their respective concentrations reach equivalent plateau values. In the second phase of the reaction, m_1 is slowly converted into M_2 (rate <0.01 s^{-1}). We suspected that the arrest of the reaction in the first phase was due to a fast dissociation of the Em_1 complex, competing with the production of M_2. Completion of the reaction will then be limited by the reassociation of the enzyme at m_1. Indeed, when calf thymus DNA was

introduced in the reaction mixture at the same time as the deoxynucleotide, the enzyme dissociating from m_1 is irreversibly trapped and appearance of the two products m_1 and M_2 reaches a common plateau value. We propose therefore the reaction scheme:

$$\begin{array}{ccccc} & & k_1 & & k_2 & \\ & EM_0 & \rightarrow & Em_1 & \rightarrow & EM_2 \\ & & k_3 \downarrow & \uparrow & k_4 & \\ & & E + m_1 & & & \end{array}$$

with $k_1 = 0.055\ s^{-1}$, $k_2 \sim k_3$ (fast with respect to k_1) and $k_4 < 0.01\ s^{-1}$.

The overall biphasic behaviour of creation and elongation of the A_t:C mismatch m_1, is therefore due to the instability of the complex between Pol I and the mismatched DNA−RNA hybrid at m_1 which hampers its elongation.

Discussion

During the initial binding step *E.coli* DNA polymerase I does not really discriminate between DNA and RNA templates. We have found that the equilibrium association constant is only decreased by a factor of four and the residence time is basically unaffected when PG5/D2, our synthetic DNA template, was replaced by a homologous RNA sequence. Association pace constants, k_{on}, computed for a simple equilibrium, are in the range of $10^6\ M^{-1}\ s^{-1}$. Our measurements on DNA templates are in the range of those given in the literature. An equilibrium constant of 200 nM was found for the association of Pol I on the homopolymer poly(dA)·oligo(dT) (Bryant *et al.*, 1983), and the Klenow fragment was reported to have a K_e of 5 nM on a synthetic 13/20mer DNA (Kuchta *et al.*, 1987). This fluctuation is probably due to the nature of the nucleotide sequence. Note that the equilibrium constant found for our RNA template is included between these two values. Basically the enzyme is devoid of selectivity during the initial binding step.

When template is in excess, incorporation of a single incoming deoxynucleotide proceeds via biphasic kinetics, as for the homologous DNA, indicating that the rate limiting step in the steady state occurs after formation of the phosphodiester bond (cf. McClure and Jovin, 1975). The time characteristic of the second slow phase is compatible with the residence time of the enzyme on its template.

Those steady-state experiments, as well as run-off assays reveal three salient features. First, the maximum velocity of the reaction is within experimental error the same on both matrices. It is limited by the recycling process after synthesis at +1 (in the first case) or by the rates which limit processivity (in the second case). Second, the largest difference observed for deoxynucleotide incorporation on the various hybrids resides in the apparent Michaelis constant for the second substrate, 0.1−1 μM on a DNA oligomer, and 25−30 μM on an RNA template. By contrast, on the same hybrids, HIV-1 reverse transcriptase has generally a better affinity for dNTP on an RNA than on a DNA template (M.Ricchetti, unpublished observations). It is possible that those differences reflect true changes in the affinity for the incoming substrate when Pol I is bound to an RNA instead of a DNA template. However, in order to be firmly established, this point requires analysis of the kinetics of the burst, using a rapid quenching apparatus. Third, the overall maximum rate of elongation is limited by the poor processivity of Pol I on the two templates. Release of the enzyme occurs at rather precise positions, corresponding to intermediates which accumulate during run-off assays performed in the absence of competing DNA. The poor overall maximal velocity observed, in all conditions tested, for the elongation of synthetic primers is certainly due to this lack of processivity and to the necessity for the enzyme to recycle at a pace which has been estimated to be of the order of 20 s to 1 min at well-defined positions of the template. Clearly what is missing to make *E.coli* Pol I a good reverse transcriptase is not its poor ability to bind and retrotranscribe RNA but rather its lack of processivity on both matrices.

The 3′→5′ exonuclease activity of Pol I is not abolished when DNA is replaced by RNA as a template. On a DNA template, we found that the ratio of the excision rates between a mismatched primer terminus and the homologous correct base pair, *d*, was of the order of 24 (larger than those found by Kuchta *et al.*, 1988). A 1.5-fold drop in discrimination was observed when Mg^{2+} was replaced by the mutagenic ion Mn^{2+}, while 4- to 6-fold changes were found by El-Deiry *et al.* (1984, 1988) using a different assay. Substituting RNA for DNA has a much more drastic effect than changing the nature of the divalent ion since now discrimination in favour of the mismatched primer terminus is lost. The rates of excision are faster in the presence of $MnCl_2$ than in $MgCl_2$. In each case they are of the same order of magnitude as those found for a perfectly matched DNA−DNA hybrid as if both terminal RNA−DNA base pairs were recognized as canonical by the exonuclease catalytic centre.

Kuchta *et al.* (1988) have proposed a kinetic mechanism governing incorrect DNA synthesis: it implies three different stages of selection. First, during dNTP binding and phosphodiester bond formation, a selectivity in the range of 10^4-10^6 is realized, as measured by comparison of the relevant kinetic constants during pre-steady state. Second, a slower conformational change following the chemical reaction allows additional time for the proofreading system to act, resulting in an increase in selectivity of four to 60. Last, extension of the primer by addition of the next base is also slower (gain by a factor estimated between six and 340) (as reviewed in Carroll and Benkovic, 1991). El Deiry *et al.* using steady-state measurements have implied the same three stages for the decreased fidelity observed in the presence of Mn^{2+} (cf. also Eger *et al.*, 1991).

Replacing the DNA template strand by RNA leads to a paradoxical situation: though the discriminating ability of Pol I editing is abolished, replication remains extremely accurate in the presence of Mg^{2+}. Such a high fidelity in the forward process implies that the first stage described by Kuchta *et al.* is extremely selective (cf. below). From the study of the extension and excision of two mismatches in $MnCl_2$, it appears that the two remaining steps are coupled: in one case, extension of the mismatched terminus was very weak; in the other, it took place as a biphasic process. In the first phase extension of the mismatch was incomplete, and limited by the rate of misincorporation which was comparable with the excision rate and to the rate of dissociation of the enzyme from its template. The second phase (which disappears in the presence of a competitor) was attributed to a slow reassociation of the enzyme on the mismatched intermediate followed by extension, a

mechanism consistent with a lower affinity of DNA Pol I for a hybrid terminated by a mismatch. Hence, low processivity limits the extension of a newly incorporated mismatch, which is already disfavoured by an efficient competition existing between synthesis and excision.

Our assay for mismatch incorporation takes into account both extended and non-extended products. In this test, error rates therefore reflect selectivity occurring during the first and second stages defined above. Since excision is non-specific, we infer that the very low level of misincorporation occurring in Mg^{2+} (error rates lower than 5×10^{-7} on the sequences examined) is mainly due to the great selectivity of the first stage of the kinetic mechanism when the template is an RNA.

This finding has two consequences. A practical one: Pol I appears to be an excellent enzyme to retrotranscribe faithfully RNA into DNA provided long incubation times are used to overcome the poor processivity of this enzyme. A theoretical one: if our guess is correct, comparison between the kinetics of correct and incorrect incorporations in the pre-steady-state phase of a reaction performed on an RNA–DNA hybrid would practically account for the extreme selectivity found in the overall process. An exonuclease mutant would not greatly affect the overall fidelity of this reaction. The study of exonuclease deficient mutants has already documented the importance of an induced fit mechanism, followed by a slow rate limiting chemistry for limiting the extent of misincorporation during this first stage of the reaction. For T7 DNA polymerase, these steps account for a discrimination of the order of 10^5-10^6 in the polymerization process on DNA templates (cf. Wong *et al.*, 1991).

In 1990, Delarue *et al.* presented an attempt to unify the structure of polymerases on the basis of sequence comparison. As noted by Sousa (1991), this proposal can be split into two hypotheses: (i) many polymerases may fold like the known tertiary structure of Pol I; (ii) they might present their most conserved motifs in a similar spatial arrangement. Comparison of the three dimensional structure of the Klenow fragment of Pol I (Ollis *et al.*, 1985) and of the DNA dependent RNA polymerase from T7 bacteriophage (cf. Chung *et al.*, 1990; Sousa, 1991; Wang *et al.*, 1992) has validated both proposals in this particular case. Six invariant common amino acid residues, belonging to the so-called A, B and C motifs, are positioned within ±1 amino acid in the same way in these two structures. The active site of T7 RNA polymerase is composed of those three different partners which can be arranged in a similar manner to the Klenow fragment of Pol I. Site directed (cf. for most recent results, Osumi-Davis *et al.*, 1992; Maksimova *et al.*, 1991; Bonner *et al.*, 1992) or linker insertion mutagenesis (A.Gross, W.-J.Chen and W.T.McAllister, submitted) have confirmed that these residues are mainly implied in the binding and the incorporation of the deoxynucleotide substrate (as the homologous residues in Pol I; cf. Polesky *et al.*, 1990; for a detailed comparison see Bonner *et al.*, 1992). A non-adjacent residue, K172, causes altered template specificity of T7 RNA polymerase: when mutated into leucine it makes the enzyme able to utilize poly(rC) or poly(U) as a template (Lyakhov *et al.*, 1992).

Converting a DNA dependent RNA polymerase into a replicase can therefore be achieved by mutating a single amino acid residue, a good example of the adaptability of an enzyme to a change in the nature of the template. This recent finding corroborates an old observation: the specificity of RNA polymerases for their templates can be significantly affected by minor changes in medium conditions like the replacement of Mg^{2+} by 1 mM Mn^{2+} (reviewed in Lazcano *et al.*, 1988).

Comparison of structures of the Klenow fragment of Pol I and of HIV-1 reverse transcriptase has led Kohlstaedt *et al.* to similar conclusions. The overall tertiary folding of the 66 kDa subunit of reverse transcriptase and of Klenow fragment are essentially similar in their so-called 'palm' subdomains, containing the catalytic site. Crucial homologous residues are located in a similar manner. An RNA–DNA hybrid duplex, in its A form, can be built into a cleft which extends between the RNase H and the polymerase active sites. Its orientation is reported to be compatible with the one found in a recently determined structure of Klenow fragment complexed with duplex DNA (Kohlstaedt *et al.*, 1992). It is therefore not too surprising that an RNA–DNA hybrid could also be positioned in the same Pol I cleft, though Pol I lacks three other conserved determinants, thought to be specific for RNA dependent polymerases (cf. Poch *et al.*, 1989).

Surprisingly, however, Pol I binds an RNA–DNA hybrid (probably in its A conformation) with only a 4-fold drop in affinity with respect to the cognate B DNA–DNA structure. A subsequent step is affected, the binding and incorporation of the second substrate. Otherwise Pol I maintains its major characteristics on this 'improper' template, a limited processivity and an excellent selectivity. Thus, DNA Pol I appears to be even more adaptable than T7 RNA polymerase (cf. again Lazcano *et al.*, 1988).

Retrotranscription by DNA Pol I could still be operative in *E.coli* in specific cases, like the formation of palindromic units (as suggested by Gilson *et al.*, 1990). More likely, its poor stringency could be a remnant of ancestral enzymes involved both in RNA dependent and DNA dependent replication (cf. for example Darnell and Doolittle, 1986; Poch *et al.*, 1989; Ricchetti, 1991). In any case we predict that few mutations directed at the catalytic site would convert Pol I into an efficient and faithful reverse transcriptase.

Materials and methods

Escherichia coli DNA polymerase I was purchased from Pharmacia and the Klenow fragment from Boehringer. The deoxynucleotide primers and templates were synthesized using β-cyanoethylphosphoramidite supports with a LKB Gene Assembler Synthetizer (Pharmacia). The dNTP substrates, FPLC pure, were purchased from Pharmacia. Their purity were checked by FPLC analysis. Radioactive [γ-^{32}P]ATP (10 Ci/mol; 10 mCi/ml) was purchased from Amersham, T4 polynucleotide kinase from BRL, T7 RNA polymerase from USB and RNasin from Promega. Poly (rA) was purchased from Boehringer.

Preparation of RNA

Template RNA was synthesized using T7 RNA polymerase and synthetic DNA templates, as described by Milligan *et al.* (1987) with the following modifications: 0.2 mM template DNA–primer hybrid was mixed with 420 units of T7 RNA polymerase in the presence of 4 mM ribonucleotides in 40 mM Tris–HCl pH 8.1, 20 mM $MgCl_2$, 1 mM spermidine, 5 mM DTT (dithiothreitol), 50 mg/ml BSA, 0.01% Triton X-100 (v/v), 8% PEG (mol. wt 6000), 20 units of RNasin, and incubated at 37°C for 4 h.

Purification of DNA and RNA

DNA and RNA synthesized as described above were purified by preparative gel electrophoresis; a small amount of radioactive fragment was co-

electrophoresed in order to determine the position of the correct band. The specific band was then cut out and the nucleic acid fragment eluted in H_2O, phenolized and precipitated three times in ethanol. Alternatively short DNA fragments, as the primers, were purified through HPLC using a PVD I column (10 cm x 5 mm). After adsorption they were eluted by a linear gradient of NaCl from 0.1 to 0.5 M in CH_3COONa 20 mM, 10% acetonitrile at pH 7.5. DNA and RNA were quantitated by spectrophotometric measurement. The molar absorption coefficient was calculated according to the length and base composition of the nucleic acid fragment.

Primer labelling and primer–template hybrid formation

The 5′ terminus of the DNA primer was labelled with ^{32}P in a 20 μl reaction mixture containing 40 mM Tris–HCl pH 7.6, 60 mM $MgCl_2$, 30 mM DTT, 10 mM ATP, 10 pmol [γ-^{32}P]ATP, 10 pmol of primer DNA and 0.5 units of T4 polynucleotide kinase. The solution was incubated at 37°C for 45 min. The reaction was terminated by heating at 90°C for 10 min. Hybridization of the labelled primer with the template was obtained adding to the 20 μl of the labelling mixture 40 pmol of the template (DNA or RNA) and NaCl at a final concentration of 0.1 M, then incubating at 90°C for 105 s; the temperature was smoothly reduced to 20°C. Poly(rA)·$(dT)_{15}$ was prepared using the same procedure (primer concentration: 33 ng/ml, template concentration: 33 μg/ml).

Nucleotide purity

FPLC pure deoxynucleotides were used. Their purity was checked by FPLC analysis as described in Ricchetti and Buc (1990). The FPLC eluted peak was dried and resuspended in H_2O at a final concentration of 10 mM. To avoid degradation of the nucleotides, in particular the deamination of cytosine, samples were aliquoted and stocked at −20°C. Each aliquot was used only once. The absence of incorporation of contaminant deoxyribonucleotides was also crosschecked via the lack of unexpected incorporation on RNA templates in the presence of $MgCl_2$ (incubation time of 5 min with the corresponding deoxynucleotide at 1 mM concentration). This test would allow detection of a contaminant present at 1 p.p.m.

Enzymatic assay

In the standard reaction, the enzyme was incubated with the hybrid, usually for 30 s at 25°C, in the following buffer: 50 mM Tris–HCl, 6 mM $MgCl_2$, 50 mM KCl, 1 mM DTT, pH 7.8. (Alternatively a manganese containing buffer was used where the 6 mM $MgCl_2$ were replaced by 1 mM $MnCl_2$). dNTPs were added and incubated for 30 s at 25°C. The reaction 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 a 20% polyacrylamide gel containing 8 M urea.

Autoradiography and densitometric analysis by direct radioactivity measurements

Autoradiography of samples labelled with ^{32}P was performed by exposing gels to photostimulable storage phosphor imaging plates in a Phosphorimager instrument (Molecular Dynamics) (Johnson *et al.*, 1990). Densitometric analysis was performed with ImageQuant Software v3.0 provided by Molecular Dynamics.

Assessment of the various constants

Run-off assays. For a given incubation time, t, the intensity of each band, $I_i(t)$, corresponding to the substrate (0) and to the n consecutive products was recorded. The material remaining at the origin was corrected for the amount of non-displacable primer (obtained by similar measures performed in enzyme excess and for very long incubation time). The average number of steps $<i>_t$ performed at time t by the hybrid enzyme complex is computed as:

$$<i>_t = \frac{\sum_{i=1}^{n} i\, I_i(t)}{\sum_{i=0}^{n} (I_i(t))}$$

For times short enough, $V_i = <i>_t/t$. V_i is expressed as a function of dNTP concentration [B] by a Michaelis Menten formula. The average apparent maximal velocity, $\bar{V}_{max}$, expressed in steps (or added nucleosides) per second, and the average apparent Michaelis constant, $\bar{K}_m$, are obtained from a fit of V_i versus [B] by a hyperbola as described in Ricchetti and Buc (1990). On the other hand $I_0/\Sigma\, I_i$ represents the depletion of the original primer–hybrid template. It follows an exponential course $e^{-v_0 t}$. Plotting v_0 against the concentration of B yields also apparent maximal velocities and Michaelis constants for the first step (after correction for the 3′→5′ exonuclease reaction).

$$v_0(B) = k_{exo} + \frac{V^0_{max}\,[B]}{K^0_m + [B]}$$

Processivity assays. In the presence of 120 μg/ml calf thymus DNA the same assay is used; at saturating deoxynucleotide concentration, when t is increased, $<i>_t$ reaches a plateau, which is taken as the average number of steps accomplished before dissociation.

Incorporation of a single incoming nucleotide. Biphasic curves analogous to the ones shown in Figure 3a are fitted by an exponential for times >5 s (the plateau corresponds to the total amount of displaceable primer, Y_T). The exponential part of the curve extrapolates at $t = 0$ to a positive value taken as the amount of preformed hybrid–enzyme complex [EA] = Y_0. Saturation curves are obtained from the change of |EA| at increasing enzyme concentration $[E]_t$ (total active concentration of enzyme was obtained by a reverse titration with PG5/D2 and was found >95% of expected value). The characteristic time of the exponential, τ, allows determination of τ_c, time required to perform one cycle by the formula:

$$\tau = \tau_c \times \frac{(Y_T - Y_0)}{Y_0}.$$

Acknowledgements

We are very grateful to P.Roux for technical assistance and O.Delpech for secretarial and artwork assistance. We thank also G.Mirambeau for many stimulating discussions. This work was supported by a grant from the Agence Nationale de Recherches sur le SIDA 1991.

References

Arnold,E., Jacobo-Molina,A., Nanni,R.G., Williams,R.L., Lu,X., Ding,J., Clark,A.D.,Jr, Zhang,A., Ferris,A.L., Clark,P., Hizi,A. and Hughes,S.H. (1992) *Nature*, **357**, 85–89.

Bambara,R.A., Uyemura,D. and Lehman,I.R. (1976) *J. Biol. Chem.*, **251**, 4090–4094.

Beese,L.S. and Steitz,T.A. (1989) *Nucleic Acids Mol. Biol.*, **3**, 28–43.

Bonner,G., Patra,D., Lafer,E.M. and Sousa,R. (1992) *EMBO J.*, **11**, 3767–3775.

Bryant,F.R., Johnson,K.A. and Benkovic,S.J. (1983) *Biochemistry*, **22**, 3537–3546.

Carroll,S.J. and Benkovic,J.J. (1991) *Nucleic Acids Mol. Biol.*, **5**, 99–113.

Chamberlin,M. (1965) *Fed. Proc.*, **24**, 1446–1447.

Chang,L.M.S. (1974) *J. Biol. Chem.*, **249**, 7441–7446.

Chung,Y.J., Sousa,R., Rose,J.P., Lafer,E. and Wang,B.C. (1990) In Wu,F.Y.-H. and Wu,C.-W. (eds), *Structure and Function of Nucleic Acids and Proteins*. Raven Press, New York, pp. 55–59.

Darnell,J.E. and Doolittle,F.W. (1986) *Proc. Natl Acad. Sci. USA*, **83**, 1271–1275.

Delarue,M., Poch,O., Tordo,N., Moras,D. and Argos,P. (1990) *Protein Engng*, **3**, 461–467.

Echols,H. and Goodman,M.F. (1991) *Annu. Rev. Biochem.*, **60**, 477–511.

Eger,B.T., Kuchta R.D., Carroll,S.S., Benkovic,P.A., Dahlbeg,M.E., Joyce,C.M. and Benkovic,S.J. (1991) *Biochemistry*, **30**, 1441–1448.

El-Deiry,W.S., Downey,K.M. and So,A.G. (1984) *Proc. Natl Acad. Sci. USA*, **81**, 7378.

El-Deiry,W.S., So,A.G. and Downey,K.M. (1988) *Biochemistry*, **27**, 546–553.

Gallo,R.C., Yang,S.Y. and Ting,R.C. (1971) *Nature*, **228**, 927–929.

Gilson,E., Perrin,D. and Hofnung,M. (1990) *Nucleic Acids Res.*, **18**, 3941–3952.

Goodman,N.C. and Spiegelman,S. (1971) *Proc. Natl Acad. Sci. USA*, **68**, 2203–2206.

Johnson,R.F., Pickett,S.C. and Barker,D.L. (1990) *Electrophoresis*, **11**, 355–360.

Karkas,J.D., Stravrianopoulos,J.G. and Chargaff,E. (1972) *Proc. Natl Acad. Sci. USA*, **69**, 398–402.

Kohlstaedt,L.A., Wang,J., Friedman,J.M., Rice,P.A. and Steitz,T.A. (1992) *Science*, **256**, 1783–1790.

Kuchta,R.D., Mirzahi,V., Benkovic,P.A., Johnson,K.A. and Benkovic,S.J. (1987) *Biochemistry*, **26**, 8410–8417.
Kuchta,R.D., Benkovic,P. and Benkovic,S.J. (1988) *Biochemistry*, **27**, 6716–6725.
Lazcano,A., Fastag,J., Gariglio,P., Ramirez,C. and Oro,J. (1988) *J. Mol. Evol.*, **27**, 365–376.
Lee-Huang,S. and Cavalieri,L.F. (1964) *Proc. Natl Acad. Sci. USA*, **51**, 1022–1028.
Loeb,L.A., Tartof,K.D. and Travaglini,E.C. (1973) *Nature New Biol.*, **242**, 66–69.
Lyakhov,D.L., Rechinsky,V.O., Tunitskaya,V.L., Chernov,B.K. and Kochetkov,S.N. (1992) *J. Cell. Biochem.*, Suppl. 16E, 166.
Maksimova,T.G. *et al.* (1991) *Eur. J. Biochem.*, **195**, 841–847.
Masaki,S. and Yoshida,S. (1978) *Biochim. Biophys. Acta*, **521**, 74–88.
McClure,W.R. and Jovin,T.M. (1975) *J. Biol. Chem.*, **250**, 4073–4080.
Milligan,J.F., Groebe,D.R., Witherell,G.W. and Uhlenbeck,O.C. (1987) *Nucleic Acids Res.*, **21**, 8783–8798.
Mizrahi,V., Benkovic,P.A. and Benkovic,S.J. (1986) *Proc. Natl Acad. Sci. USA*, **83**, 231–235.
Ollis,D.L., Brick,P., Hamlin,R., Xuong,N.G. and Steitz,T.A. (1985) *Nature*, **313**, 762–766.
Osumi-Davis,P.A., Aguilera,M.C. de, Woody,R.W. and Woody,A.-Y.M. (1992) *J. Mol. Biol.*, **226**, 37–45.
Poch,O., Sauvaget,I., Delarue,M. and Tordo,N. (1989) *EMBO J.*, **8**, 3867–3874.
Polesky,A.H., Steitz,T.A., Grindley,N.D.F. and Joyce,C.M. (1990) *J. Biol. Chem.*, **265**, 14579–14591.
Ricchetti,M. (1991) *Bull. Inst. Pasteur*, **89**, 147–158.
Ricchetti,M. and Buc,H. (1990) *EMBO J.*, **9**, 1583–1593.
Sousa,R. (1991) Ph.D. Thesis, University of Pittsburgh, Pittsburgh, PA.
Travaglini,E.C., Mildvan,A.S. and Loeb,L.A. (1975) *J. Biol. Chem.*, **250**, 8647–8656.
Wang,B.-C., Sousa,R., Chung,Y.J., Rose,J. and Lafer,E.M. (1992) *J. Cell. Biochem.*, Suppl. 16E, 117.
Wells,R.D., Flügel,R.M., Larson,J.E., Schendel,P.F. and Sweet,R.W. (1972) *Biochemistry*, **11**, 621–629.
Wong,I., Patel,S.S. and Johnson,K.A. (1991) *Biochemistry*, **30**, 526–537.
Yoshida,S., Masaki,S. and Koiwai,O. (1981) *Biochim. Biophys. Acta*, **654**, 194–200.

Received on August 20, 1992; revised on November 6, 1992