FEBS Letters 439 (1998) 302–306 FEBS 21147

Synthesis of mixed ribo/deoxyribopolynucleotides by mutant T7 RNA polymerase

S.O. Gudima, D.A. Kostyuk, O.I. Grishchenko, V.L. Tunitskaya, L.V. Memelova, S.N. Kochetkov*

Engelhardt Institute of Molecular Biology, Russian Academy of Sciences, Vavilov st., 32, 117984 Moscow, Russia

Received 12 October 1998

Abstract Synthesis of deoxynucleotide-containing RNA-like single-stranded polynucleotides (dcRNAs) using the Y639F, S641A mutant of T7 RNA polymerase (T7 RNAP) was studied. A number of different T7 promoter-containing plasmids were tested as templates for dcRNA synthesis. The dcRNA synthesis efficiency strongly depended on the sequence of the first 8–10 nucleotides immediately downstream of the promoter and increased with the distance of the first incorporated dNMP from the transcription start. The incorporation of dGMP which is obligatory for most T7 promoters in positions +1–+2(3) was practically negligible. Using the constructed plasmid pTZR7G containing seven dG links in the non-coding chain immediately downstream of the promoter, the synthesis of all possible dcRNAs (except dG-containing) was achieved with high yields.

Key words: T7 RNA polymerase; dNTP utilization; Mixed ribo/deoxypolynucleotides

1. Introduction

Bacteriophage T7 RNA polymerase (T7 RNAP) is widely used as a tool for synthesis of specific transcripts. Due to its molecular properties, the enzyme is also a suitable model for the study of mechanisms of nucleic acid biosynthesis [1–3]. Some T7 RNAP mutants are able to utilize both rNTPs and dNTPs as substrates or, in other words, to incorporate dNMP into the nascent RNA chain instead of corresponding rNMP [4,5]. In particular we have recently described such a mutant bearing two amino acid substitutions: Y639F and S641A [5]. This observation led to the enzymatic synthesis of the novel intermediate class of nucleic acids – deoxynucleotide-containing RNA-like single-stranded polynucleotides (dcRNAs), i.e. mixed ribo/deoxyribopolynucleotides containing rNMPs along with dNMPs in the synthesized polynucleotide chain. A number of such polynucleotides were obtained and used as substrates in the reactions catalyzed by HIV-1 reverse transcriptase [6] and aminoacyl-tRNA synthetases [7]. However, the yields of dcRNAs strongly depended on the template used. Thus the dcRNAs containing deoxypyrimidine nucleotides were obtained in high yields while the synthesis of dG-containing transcripts was hampered [8]. In this paper we used a number of the T7-promoter-containing plasmids as templates in the transcription-like promoter-dependent reaction catalyzed by mutant T7 RNAP. Based on the data of comparative analysis we have formulated general principles of synthesis of dcRNA of definite nucleotide composition.

2. Materials and methods

The wild-type and Y639F, S641A mutant T7 RNAP were prepared as described previously [9]. Restriction endonucleases, mung bean nuclease and T4 DNA ligase were purchased from Promega (USA) and used according to the instructions of the manufacturer.

Plasmids pPV18 and pPV19 (derivatives of the plasmids pTZ18R and pTZ19R (USB), containing the *Sal*GI-*Sph*I fragment of pBR322) were constructed earlier [10]. Plasmids pBSKS II, pGEM-2, and pSELECT-1 were purchased from Stratagene and Promega, respectively. Plasmids pLys,3 and pT5Fo were kind gifts of Dr. S. Litvak (Institut de Biochimie et Génétique Cellulaire du CNRS, Bordeaux, France) and Dr. V. Ksendzenko (Institute of Biochemistry and Physiology of Microorganisms, Pushchino, Russia). Plasmids of the pK series were provided courtesy of Prof. W.T. McAllister (State University of New York, Brooklyn, NY, USA).

Plasmid pTZR7G was constructed from the pTZ18R as follows: pTZ18R was digested by endonucleases *Eco*RI and *Sma*I. The large fragment was treated with the mung bean nuclease in the presence of an excess of dCTP and dGTP (10 mM final concentration of each) to blunt the protruding 5′ ends and ligated by T4 DNA ligase [11]. The sequence pTZR7G was confirmed by dideoxy sequencing [11].

Run-off transcription (the synthesis of 'parental' RNA and dcRNA was carried out as follows. Incubation mixture contained 200 mM HEPES-KOH, pH 7.5; 30 mM $MgCl_2$, 40 mM DTT, 2 mM spermidine, 100 mg/ml acetylated BSA, 200 U/ml RNasin, 150 mg/ml of DNA template (plasmid treated with appropriate restriction enzymes, see Table 1), 500–2000 U/ml of wild-type or mutant T7 RNAP, and various concentrations of rNTPs and dNTPs. For 'parental' RNA synthesis 4 mM rGTP and 2 mM of each rATP, rCTP and rUTP were used; for dcRNA synthesis a 4-fold excess of an appropriate dNTP instead of rNTP was added. After 4–5 h incubation (38°C) the transcripts were purified by conventional methods [11] and resuspended in an appropriate volume of RNase-free sterile water. If necessary, the transcripts were purified by preparative electrophoresis and redissolved in TE buffer [11].

The base sequences of the RNA and dcRNAs were analyzed by reverse transcriptase sequencing as in [12].

The relative efficiency of RNA and dcRNAs production was determined by measuring the ^{32}P-labelled NMP incorporation in the nascent polynucleotide product chain. In the kinetic experiments with pK plasmids the variable concentrations of rCTP (25–500 μM) and dCTP (25–1500 μM) were used. K_m values for rCTP and dCTP were calculated using the computer program Microcal Origin 4.1 (Microcal Software).

3. Results

Our earlier studies [6–8] have demonstrated that the template sequence immediately downstream of the promoter greatly contributes to the efficiency of dcRNA synthesis. To evaluate this supposition in this work a number of plasmids was tested as templates. The plasmids used (Table 1) contained the T7 RNAP consensus promoter and G in +1 posi-

*Corresponding author. Fax: (7) (095) 135 14 05,
E-mail: kochet@genome.eimb.rssi.ru

PII: S0014-5793(98)01393-3

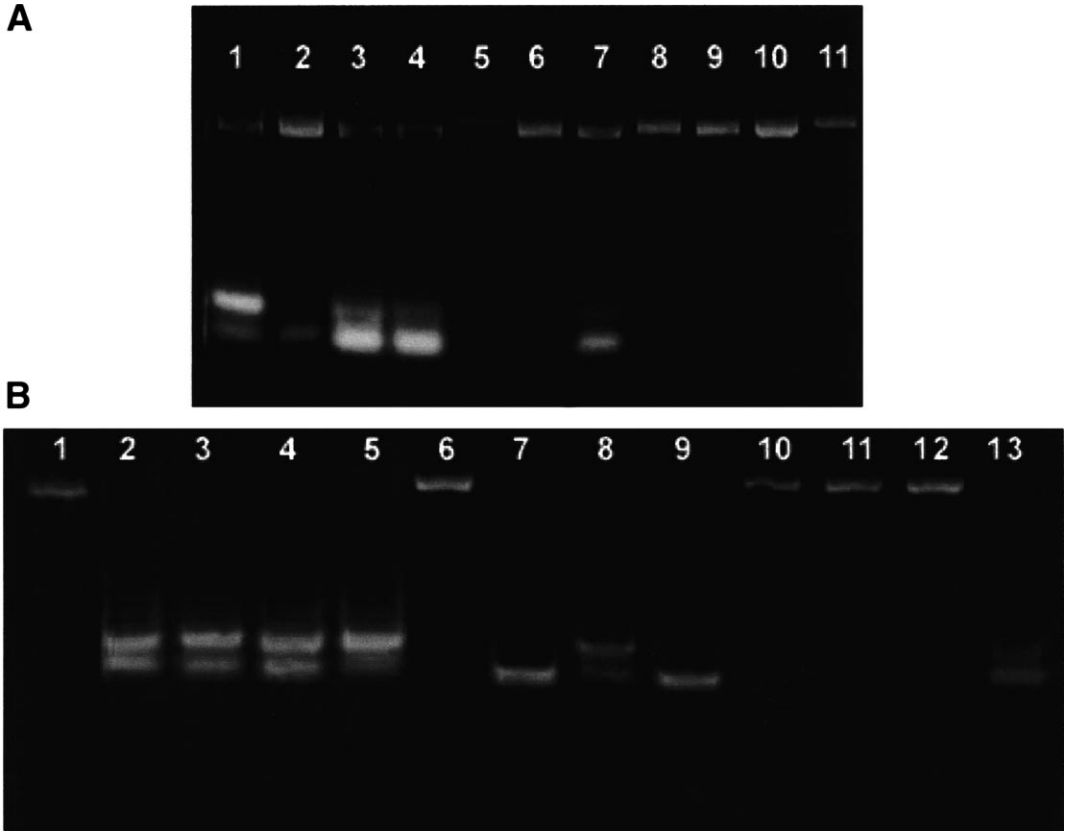

Fig. 1. Synthesis of RNA and dcRNAs by the mutant of T7 RNAP. The details of in vitro transcription and electrophoresis procedure are described in Section 2. 1% Agarose gels were stained with ethidium bromide. A: pPV19/*Xba*I used as the template. The nucleotide composition of reaction mixture: lanes: 1, rGTP, rATP, rUTP, rCTP; 2, rGTP, **dATP**, rUTP, rCTP; 3, rGTP, rATP, **dTTP**, rCTP; 4, rGTP, rATP, rUTP, **dCTP**; 5, **dGTP**, rATP, rUTP, rCTP; 6, rGTP, rATP, rUTP (control without rCTP); 7, rGTP, rATP, **dTTP, dCTP**; 8, rGTP, **dATP, dTTP**, rCTP; 9, rGTP, **dATP**, rUTP, **dCTP**; 10, rGTP, **dATP, dTTP**, dCTP; 11, **dGTP**, rATP, dTTP, **dCTP**. B: pTZR7G/*Pvu*II used as template, 1%. The nucleotide composition of reaction mixture: lanes: 1, control (no NTPs); 2, rGTP, rATP, rUTP, rCTP; 3, rGTP, **dATP**, rUTP, rCTP; 4, rGTP, rATP, **dTTP**, rCTP; 5, rGTP, rATP, rUTP, **dCTP**; 6, **dGTP**, rATP, rUTP, rCTP; 7, rGTP, **dATP**, rUTP, **dCTP**; 8, rGTP, rATP, **dTTP**, **dCTP**; 9, rGTP, **dATP**, **dTTP**, rCTP; 10, **dGTP**, **dATP**, rUTP, rCTP; 11, **dGTP**, rATP, rUTP, **dCTP**; 12, **dGTP**, rATP, **dTTP**, rCTP; 13, rGTP, **dATP, dTTP, dCTP**.

tion (G in this position is obligatory for the transcription) [1]. As is seen the plasmids contained differed in the first 15 nucleotides downstream of the promoter.

To carry out dcRNA synthesis, the previously described optimized protocol was used [6]. The use of mutant T7 RNAP in most cases resulted in the de novo production of single-stranded RNA or dcRNA polynucleotides. Sometimes the patterns of non-denaturing agarose gel electrophoresis demonstrated doubling of the bands corresponding to the full-size transcript (see Fig. 1). It was shown previously that these double bands corresponded to the conformers of the same product [6].

Fig. 1A demonstrates typical agarose gel electrophoresis of the products obtained by the run-off transcription with one of the templates (*Xba*I-restricted pPV19). As is seen not all combinations of r/dNTPs resulted in the efficient synthesis of the full-size product. Thus, no detectable amounts of dG-containing transcripts were observed. The yields of other products followed the order RNA > dT-RNA > dC-RNA > dTdC-RNA > dA-RNA thus correlating with the position of the first dNMP incorporation (dT +8, dC +7, dA +4).

Practically the same regularities in dcRNA synthesis were observed for most of the plasmids tested. Fig. 2A demonstrates relative yields of dcRNAs with different combinations

Table 1
Plasmids used as templates in dcRNA synthesis

Plasmid/restriction enzyme	5′-end terminal sequence of the RNA transcript	Transcript length (nt)
	1 5 10 15	
pPV18/*Hin*dIII	GGGAAUUCGAGCUCG	134
pPV19/*Xba*I	GGGAAGCUUGCAUGC	109
pBSKSII-/*Eco*RV	GGGCGAAUUGCAGCU	73
pSELECT-1/*Eco*RV	GGGCGAAUUCGAGCU	642
pGEM-2/*Bam*HI	GGGAGACCGGAAGCU	42
pT5Fo/*Eco*RI	GCGCCUUUAGCUGAG	77
pLys,3/*Eco*RI	GCCCGGAUAGCUCAG	86
pTZR7G/*Pvu*II	GGGGGGGAUCCACUA	131

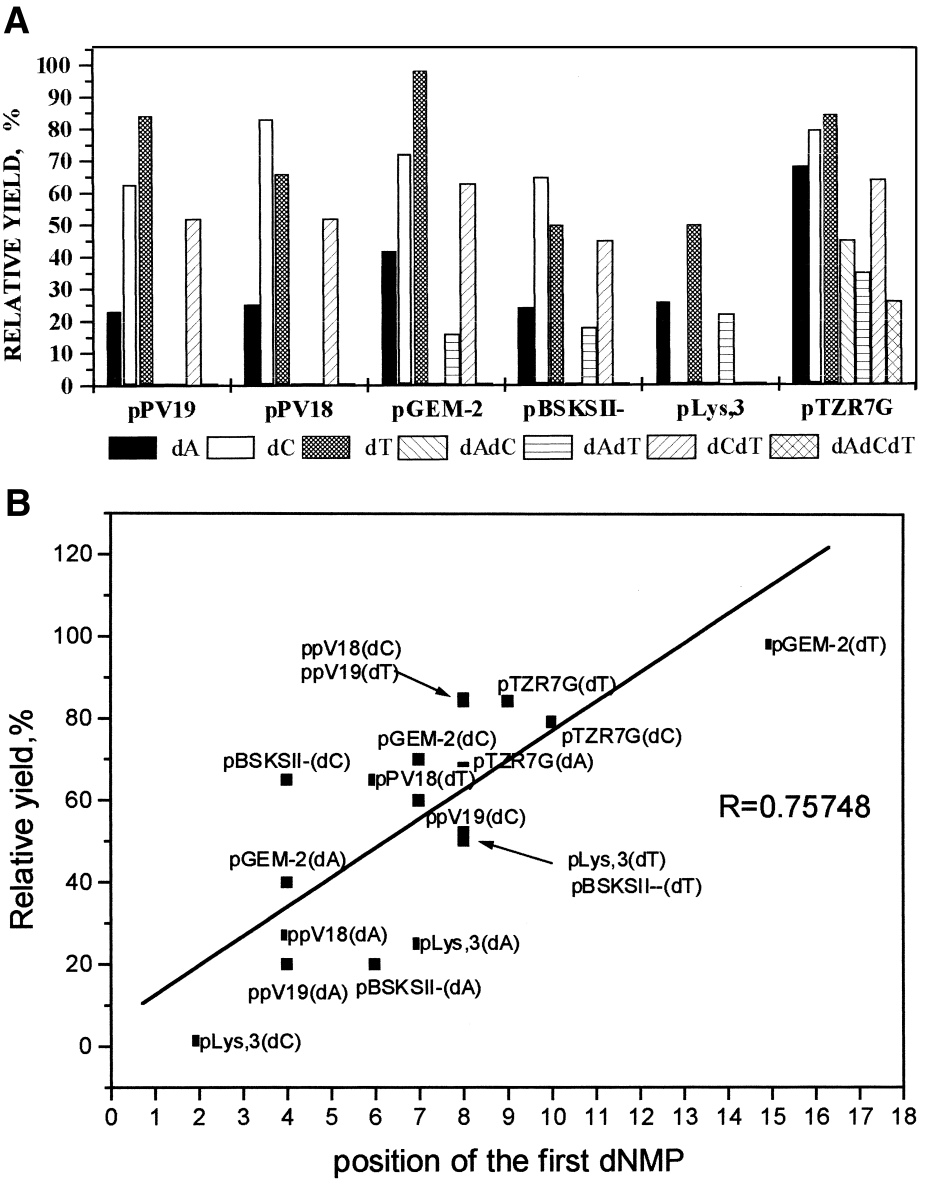

Fig. 2. Relative efficiencies of the promoter-dependent dcRNAs synthesis. A: Yields of dcRNA products for different templates used. The yield of 'parental' RNA was estimated as 100% in all cases. B: Dependence of the yield of monodeoxy-substituted dcRNAs on the position of the first deoxynucleotide to be introduced.

of r/dNTPs. The dGMP incorporation was negligible in all cases, and relative yields of synthesized products correlated again with the position of the first dNMP to be incorporated. Fig. 2B clearly demonstrates the latter effect. Thus, for plasmids pT5Fo and pLys,3, no detectable amounts of dC-containing products were obtained (the position of the first C is +2) while for pGEM-,2 where the first dTMP insertion occurs at position +15, the dT-RNA synthesis efficiency was practically the same as that of corresponding 'parental' RNA (98%). It seems that the enzyme becomes insensitive to the difference between r- and dNTP when the distance from the transcription start passes a certain value.

Though the transcripts from plasmids pPV18, pPV19 and pGEM-2 contain the first AMP residue in position +4 the yield of dA-RNA from the latter is approximately double. This is probably connected with the position of the next A residue. In pPVs the next A is in position +5, so the dAdA cluster should be formed, while in pGEM-2 it is located in position +6 (G is inserted between two As). Thus, the synthesis of dAdA in positions +4–+5 is rather ineffective while the shift of dA residue to the next position increases the yield substantially. One can conclude that position +5 is a 'crucial point' downstream of which the synthesis of dcRNA becomes more efficient.

For templates pBSKS II and pSELECT-I (note that its initial nine-nucleotide sequences are practically identical, Table 1), the row of product yields evidently does not correlate with the position of the first dNMP to be incorporated (see Fig. 2A). One could expect a yield decrease in the order RNA > dT-RNA > dA-RNA > dAdT-RNA > dC-RNA. The first C appears in position +4, so dC-, dCdT- and dAdC-RNA synthesis should be very low-efficient if detectable at all. However, the real order observed is: RNA > dC-RNA > dT-RNA > dCdT-RNA > dA-RNA > dAdT-RNA and the syn-

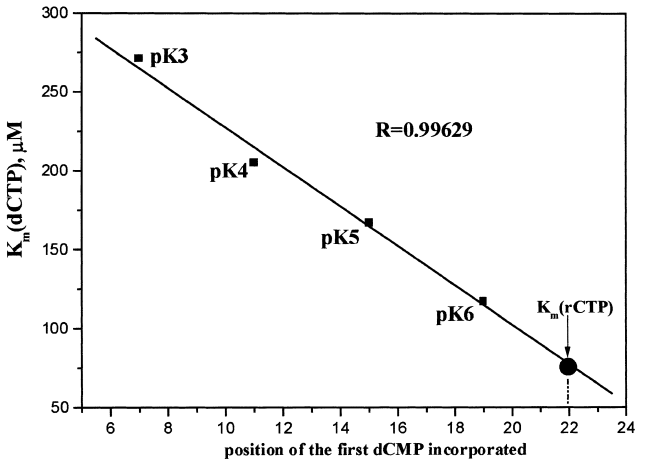

Fig. 3. K_m(dCTP) dependence on the position of the first dCMP to be introduced using the pK plasmids as templates.

thesis of dC-RNA is rather efficient (corresponding yield is approximately 65%). Two possible explanations for the observed discrepancy can be proposed.

First, it was supposed [13,14] that after the third rG incorporation the nascent RNA chain should slip in register −1 relatively to the coding DNA chain. The slippage results in the formation of short oligoG nucleotides (2–14 nt) independent of a correct complementary template sequence. These oligo(rG)s can re-initiate transcription through recurring hybridization with the same template. If such an oligonucleotide serves as a 'primer' in pBSKS II- (pSELECT-I)-driven transcription, the next dCMP incorporation will occur only at position +11 and dC-RNA synthesis should proceed with a high efficiency. Second, in contrast to other plasmids tested the first CMP residue (+4) in pBSKS II-(pSELECT-I) is located within an oligoG cluster GGGCG. As the Y639F, S641A T7 RNAP mutant was shown to have an increased level of misincorporations in the region +2–+5 [8], an elevated level of dC-containing RNA-like products due to GGGGG cluster formation may be suggested. Anyhow, the result appears to be the same – the elongation of the dcRNA chain resulting in high yields of dC-containing transcripts.

To evaluate quantitatively the correlation between dcRNA yield and the position of the first dNMP incorporation, kinetic experiments were performed. K_m values for rCTP and dCTP were determined using plasmids of the pK series. Briefly, these plasmids contain the $(GGGA)_nGACU$ sequence immediately downstream of the T7 promoter. The values for n vary from 1 (pK3) to 4 (pK6), so r/dCTP is first incorporated at positions +7, +11, +15, and +19 of the message. The results of these experiments are demonstrated in Fig. 3. There is a linear dependence between K_m(dCTP) values and dCMP position. The K_m(rCTP) values remain practically unchanged regardless of the template pK used, while K_m(dCTP) values decrease with the shift of the first dCMP position from the transcription start, thus indicating a simultaneous increase of the enzyme affinity to dCTP. It is also seen that K_m(rCTP) values which are roughly equal for all pK plasmids are also fitted to linear dependence and agree with $n = 22$. This means that after this position the mutant T7 RNAP cannot discriminate between ribo- and deoxyCTP. Thus, the results obtained suggest that the efficiency of dcRNA synthesis is determined (at least partly) by the relative affinities to r/dNTP.

The data represented above suggest that efficient dcRNA synthesis can be achieved when the position of the first dNMP to be incorporated is situated at least 5–6 residues downstream of the transcription start. As was shown above for all templates starting with GTP, the synthesis of dG-containing dcRNAs is practically negligible. Besides, the synthesis of three-dNMP-substituted products is also hampered. To avoid the latter encumbrance, we constructed a novel vector, pTZR7G, containing seven dGMP residues immediately downstream of the promoter in the non-template chain (see Table 1).

Fig. 1B demonstrates the dcRNA synthesis on the pTZR7G template. Efficient synthesis of all types of dcRNA except dG-containing ones was achieved, and the yield of a three-substituted dcRNA, dAdCdT-RNA, was rather high (26%). Thus, pTZR7G seems to be quite efficient as a template for dcRNA synthesis. This efficiency is apparently due to the location of the first incorporated deoxynucleotide in position +8 downstream from the promoter. This point is preceded by the oligo(rG) 'primer' (+1–+7) synthesized by the enzyme at the initial stages of transcription.

4. Discussion

Summarizing briefly the results obtained the following regularities of dcRNA synthesis can be formulated:

1. The efficiency of synthesis of dcRNA of a definite r/dnucleotide composition is fully determined by the sequence of the 5′ region of the message.
2. The mutant T7 RNAP is not able to initiate transcription with dNTP (or its efficiency is very low). As GTP is the initiating NTP for all known T7 promoters the synthesis of dG-containing dcRNAs is practically forbidden.
3. The efficiency of dNTP utilization by mutant T7 RNAP increases as the position of the first dNMP to be incorporated recedes from transcription start. At the beginning of the reaction (positions +1–+3) the dNTP utilization seems to be hampered (if at all possible), in the middle of the initiation region (+4–+7) it is moderate, while further it gradually increases so the mutant T7 RNAP evidently cannot discriminate between ribo- and deoxyNTPs at position +20.

We suggest that these regularities are connected to the known two-step mechanism of the T7 RNAP-catalyzed transcription [15,16]. According to this mechanism, the reaction complex consisting of the enzyme, DNA template, NTP, and RNA product is rather unstable when the length of the latter is shorter than 8–10 nt and demonstrates a high tendency to abort [13]. When next several nucleotides are added to the growing RNA chain the latter is fixed in an appropriate binding site (tentatively on the N-terminal domain of the enzyme). This binding results in reaction complex stabilization and its subsequent transition into the elongation mode. It can be supposed that the 'real' intrinsic K_m(NTP) values for each step during the 'abortive' stage differ from 'overall' K_m values obtained in usual kinetic experiments. In favor of this proposal is the fact that the values of K_m(GTP) (the latter is obligatorily located in positions +1–+2 and also often in +3) are as a rule much higher than that for other rNTPs

[1,8]. Fig. 3 demonstrates the monotonous decrease of K_m(dCTP) with a shift of the first introducing deoxynucleotide position away from the transcription origin. This result as a whole substantiates the above-mentioned data and allows us to propose that reaction complex stability depends on the affinity of the incoming NTP which is determined by the distance of its incorporation site from the transcription start. Thus, we observe a tight connection between NTP affinity and the sequence of the 5′ end (initiation) region. One can speculate that the regularities of dcRNA synthesis are valid for its 'pure transcription' prototype. Indeed, for both 'pure' RNA and dcRNA synthesis the 'abortive' stage seems to be the limitation step of the transcription. The difference between these two reactions is that in the 'pure' RNA synthesis the stabilization is achieved in the vicinity of 8–14 nt [15,17], while for dNTPs, the enzyme affinity to which is much lower than that to rNTPs, the full stabilization comes only after position +20. Thus, we observe the 'broadening' of T7 RNAP specificity towards nucleotides as the enzyme switches from the initiation to the elongation mode.

Acknowledgements: This work was supported by HHMI International Program (Project 75195-545003) and the Russian Foundation for Basic Research (Project 97-04-49376).

References

[1] Chamberlin, M.J. and Ryan, T. (1982) in: The Enzymes (Boyer, P.D., Ed.), Vol. XV, pp. 87–108, Academic Press, New York.
[2] McAllister, W.T. and Raskin, C. (1993) Mol. Microbiol. 10, 1–6.
[3] Sousa, R. (1996) Trends Biochem. Sci. 21, 186–190.
[4] Sousa, R. and Padilla, R. (1995) EMBO J. 14, 4609–4621.
[5] Kostyuk, D.A., Dragan, S.M., Lyakhov, D.L., Rechinsky, V.O., Tunitskaya, V.L., Chernov, B.K. and Kochetkov, S.N. (1995) FEBS Lett. 369, 165–168.
[6] Gudima, S.O., Kazantseva, E.G., Kostyuk, D.A., Shchaveleva, I.L., Grishchenko, O.I., Memelova, L.V. and Kochetkov, S.N. (1997) Nucleic Acids Res. 25, 4614–4618.
[7] Aphasizhev, R., Theobald-Dietrich, A., Kostyuk, D., Kochetkov, S.N., Kisselev, L.L., Giege, R. and Fasiolo, F. (1997) RNA 3, 893–904.
[8] Tunitskaya, V.L., Memelova, L.V., Kostyuk, D.A., Gudima, S.O. and Kochetkov, S.N. (1997) Mol. Biol. (Moscow) 31, 353–358.
[9] Rechinsky, V.O., Dragan, S.M., Kostyuk, D.A., Lyakhov, D.L. and Kochetkov, S.N. (1991) Mol. Biol. (Moscow) 25, 1588–1593.
[10] Pokholok, D.K., Gudima, S.O., Yesipov, D.S., Dobrynin, V.N., Rechinsky, V.O. and Kochetkov, S.N. (1993) FEBS Lett. 325, 237–241.
[11] Sambrook, J., Fritsch, E.F. and Maniatis, T. (1989) Molecular Cloning: A Laboratory Manual, Cold Sping Harbor Laboratory Press, Cold Spring Harbor, NY.
[12] Bordier, B., Helene, C., Barr, P.J., Litvak, S. and Sarih-Cottin, L. (1992) Nucleic Acids Res. 20, 5999–6006.
[13] Bonner, G., Lafer, E.M. and Sousa, R. (1994) J. Biol. Chem. 267, 11322–11328.
[14] Moroney, S.E. and Piccirilli, J.A. (1991) Biochemistry 30, 10343–10349.
[15] Ikeda, R.A., Lin, A.C. and Clarke, J. (1992) J. Biol. Chem. 267, 2640–2649.
[16] Steiz, T.A., Smerdon, S.J., Jager, J. and Joyce, C.M. (1994) Science 266, 2022–2025.
[17] Ling, M.L., Rismann, S.S., Klement, J.F., McGraw, N.J. and McAllister, W.I. (1989) Nucleic Acids Res. 17, 1605–1618.