FEBS Letters 440 (1998) 264–267 FEBS 21235

Minireview

Recent studies of T7 RNA polymerase mechanism

S.N. Kochetkov*, E.E. Rusakova, V.L. Tunitskaya

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

Received 15 October 1998

Abstract Bacteriophage T7 RNA polymerase (T7 RNAP) is known to be one of the simplest enzymes catalyzing RNA synthesis. In contrast to most RNA polymerases known, this enzyme consists of one subunit and is able to carry out transcription in the absence of additional protein factors. Owing to its molecular properties, the enzyme is widely used for synthesis of specific transcripts, as well as being a suitable model for studying the mechanisms of transcription. In this minireview the recent data on the structure and mechanism of T7 RNAP, including enzyme-promoter interactions, principal stages of transcription, and the results of functional studies are discussed.

Key words: T7 RNA polymerase; Transcription; Enzyme-promoter interaction; Initiation; Elongation; Mutagenesis; Mechanism

1. Basic properties of T7 RNA polymerase

Bacteriophage T7 RNA polymerase (T7 RNAP) is known to be one of the simplest enzymes catalyzing RNA synthesis. In contrast to most RNAPs known, this enzyme (as well as those encoded by bacteriophages T3, SP6 and K11 [1,2]) is composed of one subunit. T7 RNAP transcribes late genes of bacteriophage T7 in the absence of additional protein factors. Owing to its molecular properties, the enzyme is widely used as a tool for synthesis of specific transcripts, as well as being a suitable model for studying the mechanisms of transcription [3]. Recent years have seen a substantial progress in our understanding of the structure and mechanism of T7 RNAP [3,4]. This review describes current structural and functional information on this enzyme.

T7 RNAP was first isolated from bacteriophage T7-infected *Escherichia coli* cells in 1968 [5]. The polypeptide chain of the enzyme consists of 883 amino acid (aa) residues (MW 98 092 Da) [6]. T7 RNAP is structurally related to the members of a superfamily of nucleotide polymerases that includes single-subunit DNAPs and RNAPs such as *E. coli* DNAP I and reverse transcriptases. X-ray studies have demonstrated a marked resemblance between the three-dimensional structures of T7 RNAP and many DNAPs (Fig. 1) [7–9]. Thus, despite the almost complete lack of sequence homology, T7 RNAP and Klenow fragment of *E. coli* DNAP I demonstrate a very high structural similarity: when polymerization domains of these enzymes are superimposed, all α-helices and β-strands (except one) in the two structures correspond to each other. The shapes of these domains resemble the right arm of a man and consist of the subdomains 'palm', 'thumb' and 'fingers' [7]. A deep cleft formed by the subdomains is the binding site for the DNA template. In T7 RNAP the dimensions of this cleft allow the placing in it of almost two full turns of dsDNA [9]. Inside this cleft, structural motifs A, B, and C [10] conservative for most single-subunit nucleotide polymerases and containing functionally essential aa residues are located. These residues form the putative active site of T7 RNAP.

2. T7 RNAP-promoter interactions

For the first stages of transcription catalyzed by any RNAP, Scheme 1 is generally accepted [11], where R and P represent RNAP and promoter, and RPc is the first specific 'closed' complex. Melting of promoter dsDNA results in formation of the initiation-competent 'open' complex (RPo). The consecutive binding of complementary NTPs and synthesis of several initial links of the RNA chain (RPi) lead to the formation of a transcriptionally competent elongation complex (RPt). While for *E. coli* RNAP a number of intermediate complexes can be isolated [12], the high rate of RNA chain elongation (about 230 nucl/s at 37°C [13]) prevents such experiments with T7 RNAP.

T7 RNAP catalyzes the transcription from late promoters of bacteriophage T7 recognizing the 23 bp consensus sequence (Fig. 2). In spite of the fact that its binding affinity is 2–3 orders lower than for *E. coli* RNAP (K_d=10^{-7} M), T7 RNAP is absolutely specific to T7 promoters, and exhibits no affinity even to T3 promoters [5]. Their consensus sequences differ only in the triplet (−10 to −12), so the latter is believed to play the major role in the recognition by the respective RNAPs [14]. Mutagenesis studies indicate that T7 promoters apparently consist of two functional sites: binding (−17 to −6) and initiating (−6 to +6). Mutations in the former result in a decrease of T7 RNAP affinity without visible effect on the initiation rate, whereas mutations in the latter have weak influence on the enzyme, but greatly affect RNA synthesis [15,16].

T7 RNAP interacts with promoter asymmetrically, contacting the sequences −17 to −13, −7 to −1 and −14 to −9, and −3 to +2 of the coding and non-coding chains, respectively [17]. These contacts seem to be located mainly in the DNA

Scheme 1.

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

Abbreviations: RNAP, RNA polymerase; DNAP, DNA polymerase

PII: S0014-5793(98)01484-7

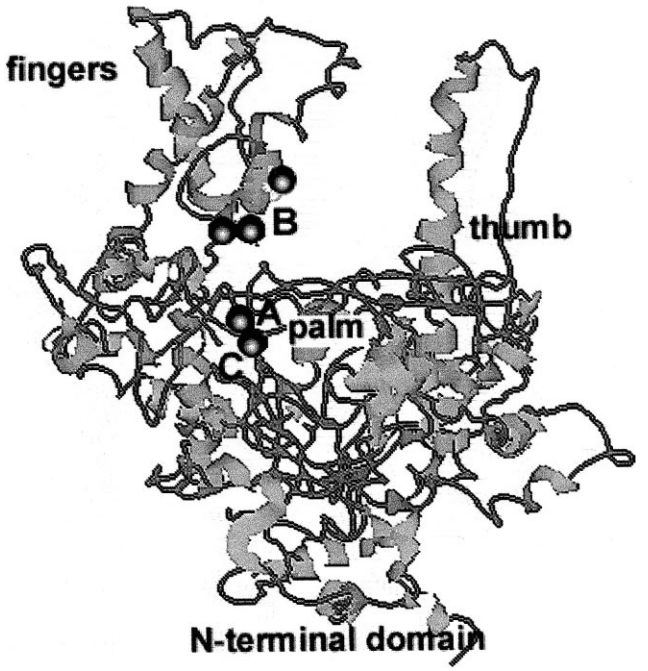

Fig. 1. Three-dimensional structure of T7 RNAP [9]. The subdomains 'palm', 'fingers', and 'thumb', as well as structural motifs A, B, and C, are indicated. The black spheres correspond to the Cα atoms of invariant aa residues D537 (motif A), K631, Y639, G640 (motif B) and D812 (motive C).

major groove, first with the non-coding chain in positions −10 and −11, and then with the coding one in positions −6 to −8 [15,16,18,19].

These interactions are supposed to facilitate DNA melting preceding the formation of the 'open' complex (RPo) [20]. S1-nuclease footprinting has demonstrated the digestion of the non-coding chain fragment (−6 to +2) with major breaks at positions −6 and −4, whereas the coding chain was protected by T7 RNAP [17]. When the promoter in the region −5 to +3 was partially single-stranded (owing to the deletion of nucleotides from any DNA chain), its binding to the enzyme was strengthened. Thus, DNA unwinding in this region evidently promotes further DNA melting thus stabilizing the complex. [19].

The 'open' complex formation is thermodynamically unfavorable, and initiation of transcription (RPi) competes with RPo renaturation to RPc [21] (see Scheme 1). The addition of GTP, the initiating NTP for all T7 RNAP promoters, followed by other rNTPs, enlarges the T7 RNAP-protected area, stabilizes the complex, and facilitates its transition to RPi [12,22,23]. These results (obtained using MPE Fe(II) footprints) suggest consecutive conformational changes of T7 RNAP during transition from RPc to RPi complexes [22,23]. However, these conclusions were criticized by Muller et al. [18]. Formation of the ladder of oligo(rG)nucleotides (3–14 nt in length) was interpreted in favor of RNA product 'slippage' for one position backward after the synthesis of (pppGpGpG) and the repeated use of the sequence CCC to add more rG links. Thus, the initiating complex seems to be transformed not to a single, but to a number of different ternary complexes, and the footprinting patterns reflect the 'averaged' situation.

3. The stages of the T7 RNAP transcription cycle

The transcription initiation stage can be determined as synthesis of the first and several next phosphodiester bonds until the reaction reaches the stable elongation mode (RPi to RPt, Scheme 1). For T7 RNAP this stage is characterized by a short lag period (10–15 s), followed by continuous (at least 30 min) RNA synthesis at a constant rate [24].

During the initiation, an abortive cessation of RNA synthesis, caused by enzyme dissociation, is highly probable [18,25]. The distinctive feature of this process is formation of short oligonucleotides (2–14 nt) representing 'broken' RNA chains. The accumulation of substantial amounts of oligonucleotides is indicative of multiple reinitiations, so the abortive transcription is clearly distinct from elongation pauses [23,26]. The probability of abortive transcription depends on the starting RNA sequence, and is minimal after GTP and maximal after UTP incorporation. Abortive transcription correlates with template topology: it is more prominent with linear or relaxed templates, while with supercoiled and single-stranded ones the full-length transcripts predominate [27].

As the length of the abortive transcripts rarely exceeds 14 nt, it may be supposed that the ternary complex dissociation competes with the incorporation of subsequent nucleotides [25]. When the length of the RNA growing chain reaches 9–12 nt, it apparently binds to the specific site in the N-terminal domain of T7 RNAP [18]. RNA binding is evidently followed by conformational changes in the enzyme, resulting in transition of the reaction from the abortive to the processive mode [18,22,25]. So, T7 RNAP may exist in two functionally different conformations. For the 'abortive' conformer, the low affinity to RNA and the hampered translocation along DNA is distinctive, while the 'processive' conformer firmly binds the RNA product and easily moves along the template [22].

According to this model, the termination is considered to be the inversion of the initiation stage. Formation of the intrinsic terminator loop disturbs the contact between RNA and polymerase due to the decreased affinity of the latter to double-stranded structures, thus shifting the equilibrium to the abortive conformation. In other words, RNA acts as a specific effector of allosteric transformations in the T7 RNAP molecule.

As mentioned above, the high rate of RNA synthesis creates obstacles to the isolation and characterization of the elongation complex. In this connection, some model systems mimicking the latter were proposed. Thus, some DNA-binding ligands (i.e. antibiotics) [28] as well as synthetic templates containing the interstrand psoralen cross-links [29] were used to fix T7 RNAP at particular sites of the template. Such

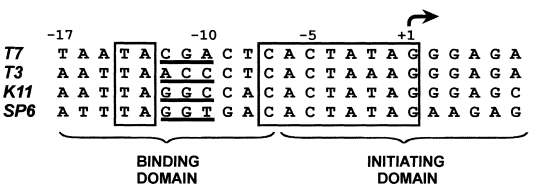

Fig. 2. Consensus sequences of class III promoters of bacteriophages T7, T3, K11, and SP6. The identical sites are boxed. The 'specificity triplets' (see text) are underlined.

'stopped' complexes appear to be stable enough and long-lived. In a number of recent works the elongation complex was modeled by using synthetic RNA-DNA heteroduplexes [30,31]. One of such model systems consisted of dsDNA (60 bp) containing non-complementary 12 bp which were, in turn, complementary to the RNA primer. After addition of NTPs, Mg^{2+}, and RNAP T7 (or *E. coli* RNAP) the elongation complex was formed and correct RNA primer extension was observed [30,31].

For quantitative analysis of the elongation rate the kinetic method based on the usage of a set of linearized templates of different length was proposed [32]. This technique enabled the determination of the time of synthesis of a single copy of full-length transcript as well as the time of a single nucleotide incorporation (i.e. an elementary act of elongation). The rate values thus calculated are in good agreement with the data of other authors [13].

4. T7RNAP interaction with NTPs. The molecular mechanism of the enzyme

A large body of information on the functionally important aa residues of T7 RNAP was obtained mainly from affinity labeling and mutagenesis (random and site-specific) experiments. This information first of all concerns the residues localized in conservative structural motifs A, B, and C, [10] (see Section 1). The T7 RNAP residues invariant in DNAPs and RNAPs are D537 (motif A), K631, Y639, G640 (motif B), and D812 (motif C). Site-directed mutagenesis data have shown that the substitutions D537N and D812N resulted in the complete loss of T7 RNAP activity [33,34]. These data, together with X-ray studies, firmly support the supposed function of these residues in the formation of Mg^{2+}-binding centers involved in the catalysis of the phosphodiester bond synthesis [9].

The affinity labeling of K631 also resulted in enzyme inactivation [35], whereas the activity of mutants by this residue depended on the type of substitution. Thus, K631G, K631L, K631M retained only 1–2% activity, whereas K631R retained about 20% [34,35]. Binding of these mutants to promoter was close to wild-type T7 RNAP.

Affinity of Y639 mutants to promoter was also close to that of the wild-type enzyme but most of them (except the fully active Y639F and 10% active Y639K) could not synthesize RNA [36,37]. Y639F acquired the striking ability to use both rNTP and dNTP as substrates [3]. The introduction of the additional mutation S641A greatly increased this property, so the 'double' mutant Y639F, S641A was able to synthesize long mixed polynucleotides where one, two, or three types of rNTP were substituted for respective the dNTPs [38,39]. The utilization of dGTP by these two mutants was insignificant, apparently due to the special role of GTP in transcription initiation [39].

The results of mutagenesis of the residues located in motif B (K631, Y639 as well as T636P, G645A, F646C) [35,36,40] allow us to propose that they are involved in NTP binding, and the invariant Y639 residue, in particular, plays a role in rNTP/dNTP discrimination.

Some aa residues located outside motifs A, B, and C were also shown to be essential for T7 RNAP function. Thus, H811Q mutation (H811 is conservative in most phage RNAPs and bacterial DNAPs) resulted in a 75% decrease in the synthesis of both the full-length and abortive transcripts [41]. It is suggested that H811 may increase the nucleophilicity of D812 in the process of catalysis [10].

The last two C-terminal T7 RNAP residues, F882 and A883, forming the so-called foot, are believed to affect both the processivity and the catalytic efficiency of the enzyme. X-ray data [9] suggest that C-terminal residues form an α-helix, located in close proximity to the active site. The 'foot' mutant (with deleted F882 and A883) binds the promoter approximately 30 times more weakly than the wild-type T7 RNAP, but exhibits an increased affinity to non-specific DNA. Its transcription activity is dramatically decreased and so altered that the mutant synthesizes considerable amounts of poly(G) even in the presence of all four NTPs. The authors explain the weak activity of the 'foot' mutant both by the decrease in the processivity of the enzyme due to the destabilization of T7 RNAP*DNA*RNA triple complex, and by the decrease in the rate of phosphodiester bond synthesis [42].

The extended T7 RNAP region (730–770) greatly defines its specificity to promoter. [43]. D748 and Q758 were suggested to play a role in the discrimination between phage RNAPs promoters [43,44]. Mutations P563T, P563A and Y571S located in the sequence 563–575, conserved for phage RNAPs, demonstrated a very low activity with promoter-containing templates (Y571S was completely inactive). At the same time, all the mutants efficiently used the 'promoterless' templates poly(dC) and poly(dI-dC) thus indicating its possible participation in the specific enzyme-promoter interactions. Besides, the mutations of the D569A(N) residue belonging to this cluster tremendously affect the catalytic activity while the affinity to promoter is largely retained [40,45]. It should be noted that both regions 730–770 and 563–575, which are likely to be involved in the T7 RNAP-promoter interactions, have no counterparts in DNAPs.

The X-ray structure of T7 RNAP solved at a relatively low resolution (3.3 Å) contains no information on the ligands bound [9]. In this connection, the analysis of the recently published structure of closely related T7 DNAP complexed with template primer and ddGTP [8] is of great value. The results of such analysis indirectly support the proposed roles of some essential T7 RNAP residues. Thus, ddGTP is bound in a cleft between the 'thumb' and 'fingers' and is paired with the template cytidine residue whose 5′-phosphate group forms a hydrogen bond with the amide nitrogen of G533 (respective residue in T7 RNAP G642). The heterocyclic base of cytosine is located over the Cα atom of G527 (in T7 RNAP T636). Phosphate residues in ddGTP are coordinated with two ions of divalent metal contacting the invariant residues D475 and D654 (in T7 RNAP D537 and D812).

A number of conserved residues of T7 DNAP provide additional steric and electrostatic interactions with bound ddGTP. Two oxygens of γ-phosphate interact with R518 (in T7 RNAP R627). Y526 (M635) and K522 (K631) contact with α- and β-phosphates, respectively. The dideoxyribose moiety of ddGTP is wedged between the phenolic ring of Y526 and the aliphatic side chain of E480, forming a hydrogen bond with a hydroxy group of invariant Y530 (in T7 RNAP Y639). These residues form a hydrophobic 'pocket' close to the C2′ atom of ribose, apparently responsible for the discrimination between r- and dNTPs.

Thus, the X-ray structure of T7 DNAP (as well as of other DNAPs) is rather useful for the analysis of the mechanism of

T7 RNAP. However, that does not exclude the necessity of determining the three-dimensional structure of the latter complexed with promoter and some NTP derivative. The fulfilment of this task will certainly be useful in getting the decisive information for further progress in the studies of T7 RNAP and RNAPs as a whole.

Acknowledgements: The work of the authors was supported by HHMI International program (Project 75195-545003) and the Russian Foundation for Basic Research (Projects 97-04-49376 and 98-04-48194).

References

[1] McGraw, N.J., Bailey, J.N., Cleaves, G.R., Dembinsky, D.R., Gocke, C.R., Collifee, L.K., MacWright, R.C. and McAllister, W.T. (1985) Nucleic Acids Res. 13, 6753–6766.
[2] Dietz, A., Weisser, H.J., Kossel, H. and Hausmann, R. (1990) Mol. Gen. Genet. 221, 283–286.
[3] Sousa, R. (1996) Trends Biochem Sci. 21, 186–190.
[4] McAllister, W.T. (1993) Cell. Mol. Biol. Res. 39, 385–391.
[5] Chamberlin, M., McGrath, J. and Waskell, L. (1970) Nature 228, 227–231.
[6] Moffatt, B.A., Dunn, J.J. and Studier, F.W. (1984) J. Mol. Biol. 173, 265–269.
[7] Steitz, T.A. (1993) Curr. Opin. Struct. Biol. 3, 31–38.
[8] Doublie, S., Tabor, S., Long, A.M., Richardson, C.C. and Ellenberg, T. (1998) Nature 39, 251–258.
[9] Sousa, R., Chung, Y.T., Rose, J.P. and Wang, B.-C. (1993) Nature 364, 593–599.
[10] Delarue, M., Poch, O., Tordo, N., Moras, D. and Argos, P. (1990) Protein Eng. 3, 461–467.
[11] Straney, D.C. and Crothers, D.M. (1985) Cell 43, 449–459.
[12] Chamberlin, M. (1982) Enzymes 15, 61–86.
[13] Golomb, M. and Chamberlin, M. (1974) J. Biol. Chem. 249, 2858–2863.
[14] Klement, J.F., Moorefield, M.B., Jorgenson, E., Brown, J.E., Risman, S. and McAllister, W.T. (1990) J. Mol. Biol. 215, 21–29.
[15] Li, T., Ho, H.H., Maslak, M., Schick, C. and Martin, C.T. (1996) Biochemistry 35, 3722–3727.
[16] Schick, C. and Martin, C.T. (1995) Biochemistry 34, 666–672.
[17] Muller, D.K., Martin, C.T. and Coleman, J.E. (1989) Biochemistry 28, 3306–3313.
[18] Muller, D.K., Martin, C.T. and Coleman, J.E. (1988) Biochemistry 27, 5763–5771.
[19] Jorgensen, E.D., Durbin, R.K., Risman, S.S. and McAllister, W.T. (1991) J. Biol. Chem. 266, 645–651.
[20] Ujvari, A. and Martin, C.T. (1997) J. Mol. Biol. 273, 775–781.
[21] Villeman, J., Guajardo, R. and Sousa, R. (1997) J. Mol. Biol. 273, 958–977.
[22] Ikeda, R.A. and Richardson, C.C. (1987) J. Biol. Chem. 262, 3800–3808.
[23] Ikeda, R.A. and Richardson, C.C. (1986) Proc. Natl. Acad. Sci. USA 83, 3614–3618.
[24] Chamberlin, M. and Ring, J. (1973) J. Biol. Chem. 248, 2235–2244.
[25] Morris, C.E., McGraw, N.J., Joho, K., Brown, J.E., Klement, J.F., Ling, M.L. and McAllister, W.T. (1987) in: RNA Polymerase and Regulation of Transcription (Reznikoff, W.S. et al., Eds.), pp. 47–58, Elsevier, New York.
[26] Ling, M.-L., Risman, S.S., Klement, J.F., McGraw, N. and McAllister, W.T. (1989) Nucleic Acids Res. 17, 1605–1618.
[27] Diaz, G.A., Rong, M., McAllister, W.T. and Durbin, R.K. (1996) Biochemistry 35, 10837–10843.
[28] Kassavetis, G.A., Kaya, K.M. and Chamberlin, M.J. (1978) Biochemistry 17, 5798–5804.
[29] Sastry, S.S. and Hearst, J.E. (1991) J. Mol. Biol. 221, 1111–1125.
[30] Daube, S.S. and von Hippel, P.H. (1994) Biochemistry 33, 340–347.
[31] Daube, S.S. and von Hippel, P.H. (1992) Science 258, 1320–1323.
[32] Lyakhov, D.L., Ilgenfritz, H., Chernov, B.K., Dragan, S.M., Rechinsky, V.O., Pokholok, D.K., Tunitskaya, V.L. and Kochetkov, S.N. (1992) Mol. Biol. (Moscow) 26, 679–687.
[33] Osumi-Davis, P.A., Sreerama, N., Volkin, D.B., Middaugh, R.C., Woody, R.W. and Woody, A.-Y.M. (1994) J. Mol. Biol. 237, 5–19.
[34] Tunitskaya, V.L., Bazhulina, N.P., Dragan, S.M., Kuznetsova, N.V., Lyakhov, D.L., Rechinskii, V.O., Morozov, Yu.V. and Kochetkov, S.N. (1993) Mol. Biol. (Moscow) 27, 51–57.
[35] Maksimova, T.G., Mustaev, A.A., Zaychikov, E.F., Lyakhov, D.L., Akbarov, A.K., Luchin, S.V., Rechinsky, V.O., Chernov, B.K. and Kochetkov, S.N. (1991) Eur. J. Biochem. 195, 841–847.
[36] Tunitskaya, V.L., Dragan, S.M., Kostyuk, D.A., Lyakhov, D.L., Memelova, L.V., Rechinsky, V.O. and Kochetkov, S.N. (1994) Biochemistry (Moscow) 59, 355–361.
[37] Bonner, G., Patra, D., Lafer, E.M. and Sousa, R. (1992) EMBO J. 11, 3767–3775.
[38] 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.
[39] Tunitskaya, V.L., Memelova, L.V., Kostyuk, D.A., Gudima, S.O. and Kochetkov, S.N. (1997) Mol. Biol. (Moscow) 31, 353–358.
[40] Rechinsky, V.O., Kostyuk, D.A., Lyakhov, D.L., Chernov, B.K. and Kochetkov, S.N. (1993) Mol. Gen. Genet. 238, 455–458.
[41] Osumi-Davis, P.A., de Aguillera, M.C., Woody, R.W. and Woody, A.-Y.M. (1992) J. Mol. Biol. 226, 37–45.
[42] Mookhtiar, K.A., Peluso, P.S., Muller, D.K., Dunn, J.J. and Coleman, J.E. (1991) Biochemistry 30, 6305–6313.
[43] Raskin, C.A., Diaz, G.A. and McAllister, W.T. (1993) Proc. Natl. Acad. Sci. USA 90, 3147–3151.
[44] Rong, M., He, B., McAllister, W.T. and Durbin, R.K. (1998) Proc. Natl. Acad. Sci. USA 95, 515–519.
[45] Rechinsky, V.O., Chernov, B.K., Dragan, S.M., Kostyuk, D.A., Tunitskaya, V.L. and Kochetkov, S.N. (1995) Mol. Gen. Genet. 247, 110–113.