Acta Crystallographica Section D
Biological Crystallography
ISSN 0907-4449

Crystallization and preliminary crystallographic analysis of T7 RNA polymerase elongation complex

Dmitry Temiakov,[a]† Tahir H. Tahirov,[b]† Michael Anikin,[a] William T. McAllister,[a] Dmitry G. Vassylyev[c,d]* and Shigeyuki Yokoyama[c,d,e,f]*

[a]Morse Institute of Molecular Genetics, Department of Microbiology and Immunology, SUNY Health Science Center at Brooklyn, 450 Clarkson Avenue, Brooklyn, New York 11203-2098, USA, [b]High Throughput Factory, RIKEN Harima Institute at SPring-8, 1-1-1 Kouto, Mikazuki-cho, Sayo, Hyogo 679-5148, Japan, [c]Cellular Signaling Laboratory, RIKEN Harima Institute at SPring-8, 1-1-1 Kouto, Mikazuki-cho, Sayo, Hyogo 679-5148, Japan, [d]Structurome Project, RIKEN Harima Institute at SPring-8, 1-1-1 Kouto, Mikazuki-cho, Sayo, Hyogo 679-5148, Japan, [e]Department of Biophysics and Biochemistry, Graduate School of Science, Univeristy of Tokyo, 7-3-1 Hongo, Bukyo-ku, Tokyo 113-0033, Japan, and [f]RIKEN Gemomic Sciences Center, 1-7-22 Suehiro-cho, Tsurumi, Yokohama 230-0045, Japan

† These authors contributed equally to this work.

Correspondence e-mail: dmitry@yumiyoshi.harima.riken.go.jp, yokoyama@biochem.s.u-tokyo.ac.jp

Received 29 July 2002
Accepted 29 October 2002

Stable transcription-elongation complexes consisting of T7 RNA polymerase (molecular mass 99 kDa) in association with a nucleic acid scaffold consisting of an 8 bp RNA–DNA hybrid and 10 bp of downstream DNA were assembled and crystallized by the sitting-drop vapour-diffusion technique under near-physiological conditions. The crystals diffract beyond 2.6 Å resolution and belong to space group $P1$, with unit-cell parameters a = 79.91, b = 84.97, c = 202 Å, α = 90.36, β = 92.97, γ = 109.94°. An unambiguous molecular-replacement solution was found using the C-terminal portion of the T7 RNA polymerase structure from the early initiation complex as a search model. Model building and structure refinement are now in progress.

1. Introduction

Although the DNA-dependent RNA polymerase (RNAP) encoded by bacteriophage T7 consists of a single subunit (883 amino acids, 99 kDa), this enzyme is able to carry out all the steps of the transcription cycle in the same manner as the multisubunit RNAPs found in bacteria and eukaryotic cells (McAllister, 1997). The C-terminal portion of phage RNAP (residues 267–883, which contains the active site) is structurally related to the pol I-like DNA polymerases and it has been suggested that the N-terminal domain of the phage RNAP confers upon the catalytic core the ability to carry out the steps that are specific to the transcription process (Cheetham *et al.*, 1999; Cheetham & Steitz, 2000; Jeruzalmi & Steitz, 1998; Sousa *et al.*, 1993).

To initiate RNA synthesis, the enzyme must bind to a specific promoter DNA sequence that lies upstream of the start site for transcription, melt apart the two strands of the DNA in the vicinity of the start site (forming a transcription bubble) and begin RNA synthesis using the coding strand of the downstream DNA as a template and a single ribonucleotide as a primer. A number of biochemical observations indicate that the organization of the transcription complex during elongation (EC) is distinct from the organization of the early initiation complex (IC), suggesting that substantial changes in structure may occur during this transition. These include changes in the footprint of the enzyme on the DNA, changes in the pattern of protease cleavage and changes in nucleic acid–protein contacts as probed by nuclease protection and protein–nucleic acid crosslinking experiments (Brieba & Sousa, 2001; Ikeda & Richardson, 1986; Ma, Temiakov *et al.*, 2002; Place *et al.*, 2000; Sousa *et al.*, 1992; Temiakov *et al.*, 2000).

Four structures of T7 RNAP have been reported: free enzyme, RNAP bound to the transcription inhibitor T7 lysozyme, a binary RNAP–promoter complex and an early IC in which the first three bases of the template strand have been transcribed (Cheetham *et al.*, 1999; Cheetham & Steitz, 1999; Jeruzalmi & Steitz, 1998; Sousa *et al.*, 1993). All these structures exhibit a similar conformation of the enzyme, indicating that the initial stages of transcription may be achieved without serious changes in enzyme structure. However, these structures were not in agreement with biochemical observations with regard to the organization of the EC and did not provide plausible mechanisms for further transcription progress (Huang & Sousa, 2000; Liu & Martin, 2001; Ma *et al.*, 2002; Severinov, 2000; Temiakov *et al.*, 2000).

Here, we report crystallization of the T7 RNAP EC, the structure of which may help to resolve the principal differences between previous biochemical and structural data (Cheetham & Steitz, 1999; Severinov, 2000; Temiakov *et al.*, 2000) and provide further insight into mechanisms of transcription at various stages of transcription cycle.

2. Experimental procedures and results

2.1. T7 RNAP purification and assembly of the elongation complex

Histidine-tagged T7 RNAP was purified as previously described (He *et al.*, 1997; Temiakov *et al.*, 2000) with an additional purification step on phosphocellulose introduced to ensure high purity of the enzyme. Briefly, after elution from

Ni^{2+}-agarose beads the protein was dialyzed against phosphate buffer (20 m*M* sodium phosphate pH 7.7, 50 m*M* NaCl, 5 m*M* *β*-mercaptoethanol) and mixed with 1 ml of phosphocellulose slurry (Whatman P11; prepared as recommended by the supplier). After 1 h incubation on a rotary shaker at 277 K, the slurry was filtered through an empty PD10 column (Pharmacia). The flowthrough was discarded and the polymerase was eluted with four 1 ml aliquots of elution buffer (200 m*M* sodium phosphate pH 7.7, 400 m*M* NaCl, 5 m*M* *β*-mercaptoethanol). Peak fractions were identified by gel electrophoresis, pooled and dialyzed against buffer containing 10 m*M* Tris pH 8.1, 200 m*M* NaCl, 20 m*M* $MgCl_2$ and 5 m*M* *β*-mercaptoethanol. The protein was concentrated to 50 mg ml^{-1} on a Vivaspin concentrator (Sartorius), divided into aliquots and stored at 203 K. We estimate the purity of the polymerase to be greater than 99% by gel electrophoresis.

Synthetic DNA oligonucleotides (18 nt template strand and 10 nt non-template strand) and 12-mer synthetic RNA (Dharmagon Research) were dissolved in water (all at 3 m*M* concentration), incubated for 7 min at 343 K and then allowed to cool to 293 K over a period of 2–3 h to give a solution of scaffold template having an eight-base-pair RNA–DNA hybrid with an unpaired 5′ RNA 'tail' of 4 nt and 10 bp of downstream DNA (Fig. 1*a*). Transcription

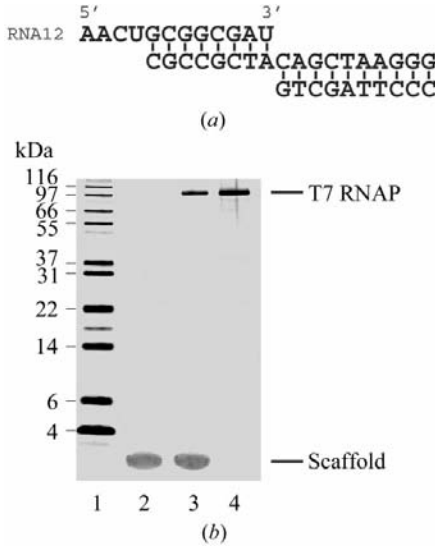

Figure 1
(*a*) Nucleic acid scaffold used to assemble T7 RNA polymerase elongation complex. (*b*) PAGE analysis of the crystals. Samples were resolved in 4–12% gradient gels (Nu-PAGE, Invitrogen) using an MES buffer system and visualized by silver staining. Lane 1, Mark-12 molecular-weight standards (Invitrogen); lane 2, RNA–DNA scaffold; lane 3, crystal; lane 4, T7 RNA polymerase.

complexes were assembled by mixing T7 RNAP with the scaffold in a 1:1 molar ratio in buffer containing 10 m*M* Tris pH 8.1, 200 m*M* NaCl, 20 m*M* $MgCl_2$ and 5 m*M* *β*-mercaptoethanol. Such complexes have been shown to be stable and fully active *in vitro* and to have all of the properties of ECs formed by initiation at a T7 promoter in a double-stranded DNA template (Temiakov *et al.*, 2003).

2.2. Crystallization

We elected to use low-salt solutions in search of crystallization conditions for the T7 RNAP EC in order to ensure stability of the assembled complex (Temiakov *et al.*, 2003). In previous crystallization attempts (Chen *et al.*, 1999), it was demonstrated that PEG 8000 was the only precipitant that produced crystals of wild-type and His-tagged T7 polymerase. The same precipitant was successfully used in crystallization of promoter and initiation complexes of T7 RNAP, even though a higher salt concentration was used (Cheetham & Steitz, 1999; Jeruzalmi & Steitz, 1998). We have tried a variety of polyethylene glycol precipitants and found that solutions of PEG 5000 MME, PEG 6000 and PEG 8000 produced crystals, the latter crystals being suitable for crystallographic studies. Crystallization was carried out by the sitting-drop vapour-diffusion technique at 293 K. The drop, containing 2 µl of 10 mg ml^{-1} complex solution, was mixed with 2 µl of well solution containing 10% PEG 8000, 10% glycerol, 5 m*M* *β*-mercaptoethanol, 100 m*M* Tris buffer pH 8.1. Thin plate-like crystals (0.35 × 0.35 × 0.05 mm) grew within 3–5 d.

To analyze the composition of the complex prior to collection of diffraction data, one large single crystal was removed from the sitting drop, washed several times in a stabilizing solution (12.5% PEG 8000, 12.5% glycerol, 5 m*M* *β*-mercaptoethanol in 100 m*M* Tris buffer pH 8.1), dissolved in SDS and subjected to electrophoresis in a 4–12% Nu-PAGE gradient gel using an MES buffer system (Invitrogen). After electrophoresis, the gel was stained with silver stain (Wako) (Fig. 1*b*). The crystal sample (lane 3) contained a large protein band (100 kDa) corresponding to T7 RNAP (lane 4) as well as a fast-migrating band (<4 kDa) that corresponded to the nucleic acid scaffold (lane 2).

The size and the quality of the crystals depended upon the ionic strength of the solution, the pH and the presence of certain additives. The final conditions (10% PEG 8000, 8% glycerol, 5 m*M* *β*-mercaptoethanol, 100 m*M* Tris buffer pH 8.1) yielded high-quality crystals (Fig. 2).

Table 1
Data-collection statistics for the T7 RNAP elongation complex.

Values in parentheses are for the highest resolution shell.

Unit-cell parameters	
a (Å)	79.91
b (Å)	84.97
c (Å)	202.00
α (°)	90.36
β (°)	92.97
γ (°)	109.94
Temperature (K)	100
Space group	*P*1
Molecules in asymmetric unit	4
Solvent content (%)	57.3
Resolution (Å)	40–2.6 (2.69–2.6)
Observations	488245
Unique reflections	143689
R_{merge}†	0.081 (0.451)
Completeness (%)	94.2 (73.6)
$I/\sigma(I)$	14.7 (2.9)

† $R_{merge} = \sum|I_j - \langle I_j\rangle|/\sum\langle I_j\rangle$, where I_j is the intensity of reflection j and $\langle I_j\rangle$ is the average intensity of reflection j.

2.3. Data collection and processing

The quality of the crystals was first checked using a laboratory X-ray generator (Rigaku FR-D, rotating anode, Cu radiation, focal size 0.15 × 1.5 mm, confocal mirrors) operated at 50 kV and 100 mA. Diffraction was measured using a Rigaku R-AXIS IV++ imaging-plate detector. For data collection at room temperature (293 K), the crystals were mounted in glass capillaries. For data collection at cryotemperature (100 K), the crystals were soaked in a stock solution in which the concentration of cryoprotectant was gradually increased over a period of ∼2–3 min. The crystals then were mounted in nylon loops (Hampton Research) and flash-cooled in a stream of cold nitrogen gas. The best cryoprotectant solution contained 10%(*w*/*v*) PEG 8000, 10%(*v*/*v*) glycerol, 20%(*v*/*v*) PEG 400, 100 m*M* sodium chloride, 10 m*M* magnesium chloride and 5 m*M* *β*-mercaptoethanol in 50 m*M* Tris buffer pH 8.1. The reduction of salt

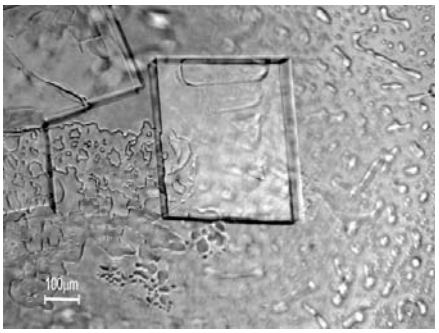

Figure 2
Crystals of T7 RNA polymerase elongation complex.

concentration in the cryoprotectant helped to prevent crystal cracking and improved diffraction quality.

A complete diffraction data set at 2.6 Å resolution was collected using synchrotron radiation at SPring-8 beamline BL45XU equipped with a Rigaku R-AXIS V imaging-plate detector. The diffraction data were processed using the *HKL*2000 program package (Otwinowski & Minor, 1997) (Table 1).

3. Discussion

Although the shape and thickness of the different crystals were variable, the unit-cell parameters were nearly identical. Crystals belong to space group *P*1 and the solvent content (Matthews, 1968) (Table 1) showed that the asymmetric unit could accommodate four molecules of the complex.

Initial attempts to solve the structure by the molecular-replacement method using the coordinates of the whole T7 RNA polymerase molecule from the initiation complex (PDB code 1qln) were unsuccessful, indicating that large conformational changes occur upon the transition of T7 RNA polymerase from the early IC to the EC. Indeed, the 8 bp DNA–RNA hybrid (as observed in the eukaryotic EC; Gnatt *et al.*, 2001) superimposed on the 3 bp hybrid of the T7 RNAP early IC clashes with the RNAP N-terminal domain, suggesting that a major rearrangement of this domain occurs upon the transition from the early IC to the EC. When the N-terminal domain (residues 2–267) and the protruding parts of C-terminal domain (residues 344–393, 590–612, 742–767) were removed from the search model, clear rotation and translation solutions for the four molecules in the asymmetric unit (related by non-crystallographic symmetry) were obtained using the *AMoRe* program (Navaza, 2001) (Table 2). The rigid-body refinement, as implemented in *AMoRe*, yielded an *R* factor of 43.7% at 3.0 Å resolution. Model building of the nucleic acids and the rearranged portions of the protein and refinement of the structure are in progress.

Table 2
Molecular-replacement solution in *P*1 space group for T7 RNAP elongation complex.

(*a*) Rotation search. Resolution range 20.0–4.0 Å, search radius 30.0 Å. CC_P and CC_I are Patterson and intensity-based correlation coefficients as defined in the *AMoRe* program (Navaza, 2001).

	CC_P	CC_I	*R* factor (%)
1st peak	28.1	24.8	52.4
2nd peak	27.4	23.6	52.4
3rd peak	14.9	21.5	53.7
4th peak	13.5	20.3	54.2
5th peak	8.7	18.2	54.5

(*b*) Translation search. Resolution range 8.0–4.0 Å.

	Correlation coefficient	Translation function	*R* factor (%)
2 molecules in ASU			
1st peak	30.0	19.8	47.0
2nd peak	20.3	5.3	50.3
3 molecules in ASU			
1st peak	35.0	16.5	46.0
2nd peak	26.5	4.9	48.7
4 molecules in ASU			
1st peak	38.2	15.6	45.2
2nd peak	30.9	5.0	47.5

(*c*) Rigid-body refinement.

Resolution range (Å)	8.0–3.0
R factor (%)	43.7

References

Brieba, L. G. & Sousa, R. (2001). *EMBO J.* **20**, 6826–6835.
Cheetham, G., Jeruzalmi, D. & Steitz, T. A. (1999). *Nature (London)*, **399**, 80–83.
Cheetham, G. & Steitz, T. A. (1999). *Science*, **286**, 2305–2309.
Cheetham, G. & Steitz, T. A. (2000). *Curr. Opin. Struct. Biol.* **10**, 117–123.
Chen, C. J., Liu, Z. J., Rose, J. P. & Wang, B. C. (1999). *Acta Cryst.* D**55**, 1188–1192.
Gnatt, A. L., Cramer, P., Fu, J., Bushnell, D. A. & Kornberg, R. D. (2001). *Science*, **292**, 1876–1882.
He, B., Rong, M., Lyakhov, D. L., Gartenstein, H., Diaz, G. A., Castagna, R. C., McAllister, W. T. & Durbin, R. K. (1997). *Protein Expr. Purif.* **9**, 142–151.
Huang, J. & Sousa, R. (2000). *J. Mol. Biol.* **303**, 347–358.
Ikeda, R. A. & Richardson, C. C. (1986). *Proc. Natl Acad. Sci. USA*, **83**, 3614–3618.
Jeruzalmi, D. & Steitz, T. A. (1998). *EMBO J.* **17**, 4101–4113.
Liu, C. & Martin, C. T. (2001). *J. Mol. Biol.* **308**, 465–475.
Ma, K., Temiakov, D., Jiang, M., Anikin, M., & McAllister, W. T. (2002). *J. Biol. Chem.* **277**, 43206–43215.
McAllister, W. T. (1997). *Nucleic Acids Mol. Biol.* **11**, 15–25.
Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
Navaza, J. (2001). *Acta Cryst.* D**57**, 1367–1372.
Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
Place, C., Oddos, J., Buc, H., McAllister, W. T. & Buckle, M. (2000). *Biochemistry*, **38**, 4948–4957.
Severinov, K. (2000). *Proc. Natl Acad. Sci. USA*, **98**, 5–7.
Sousa, R., Chung, Y. J., Rose, J. P. & Wang, B. C. (1993). *Nature (London)*, **364**, 593–599.
Sousa, R., Patra, D. & Lafer, E. M. (1992). *J. Mol. Biol.* **224**, 319–334.
Temiakov, D., Anikin, M. & McAllister, W. T. (2003). In the press.
Temiakov, D., Mentesana, P. E., Ma, K., Mustaev, A., Borukhov, S. & McAllister, W. T. (2000). *Proc. Natl Acad. Sci. USA*, **97**, 14109–14114.