Yuko Murayama,^{a,b} Shun-ichi
 Sekine^{a,b*} and Shigeyuki
 Yokoyama^{a,b*}

^aDepartment of Biophysics and Biochemistry,
 Graduate School of Science, University of
 Tokyo, 7-3-1 Hongo, Bunkyo-ku,
 Tokyo 113-0033, Japan, and ^bRIKEN Systems
 and Structural Biology Center, 1-7-22 Suehiro,
 Tsurumi, Yokohama 230-0045, Japan

Correspondence e-mail:
 shunichi.sekine@riken.jp, yokoyama@riken.jp

Received 15 November 2012

Accepted 2 January 2013

Crystallization and preliminary X-ray crystallographic analyses of *Thermus thermophilus* backtracked RNA polymerase

DNA-dependent RNA polymerase (RNAP) synthesizes RNA complementary to the template DNA. During transcript elongation, RNAP often undergoes backward translocation ('backtracking') by dissociating the 3' end of the nascent RNA transcript from the template DNA. While the backtracked state of RNAP is inactive in RNA elongation, it actively hydrolyses the RNA 3' end to regenerate the active elongation complex. To study the structural basis of the backtracked state and its cleavage activity, two backtracked RNAP complexes were reconstituted by assembling *Thermus thermophilus* RNAP with designed nucleic acid scaffolds. The reconstituted backtracked complexes were active in the transcript-cleavage reaction. These complexes were crystallized and X-ray diffraction data sets were obtained at resolutions of 3.4 and 3.7 Å.

1. Introduction

DNA-dependent RNA polymerase (RNAP) plays a fundamental role in gene expression. It transcribes genes to produce protein-coding messenger RNAs and noncoding RNAs with various functions, such as transfer RNAs and ribosomal RNAs. RNAP transcribes genes into RNAs by translocating along DNA. Although transcription is unidirectional, RNAP often undergoes reverse translocation, or 'backtracking', during transcription (Komissarova & Kashlev, 1997a; Nudler *et al.*, 1997). Backtracking occurs when RNAP encounters obstacles to its forward translocation, such as the incorporation of an incorrect nucleotide, damage in the template DNA or specific regulatory sequences (Donahue *et al.*, 1994; Jeon & Agarwal, 1996; Marr & Roberts, 2000). In the backtracked state, one or more nucleotide residues at the 3' end of the nascent RNA are dissociated from the template strand and therefore the backtracked complex is inactive in RNA elongation (Komissarova & Kashlev, 1997b). The backtracked complex can be restored to the active elongation complex (EC) either by its forward translocation or by cleavage of the protruding nucleotide residues from the backtracked RNA using another intrinsic activity of RNAP. In the latter case, RNAP hydrolyses the RNA at its catalytic site to remove the protruding 3'-end residues, allowing it to resume RNA elongation from the newly formed 3' end (Izban & Luse, 1993; Reines, 1992). RNAP backtracking and the subsequent transcript cleavage serve as proofreading mechanisms in the case of nucleotide misincorporation (Erie *et al.*, 1993; Thomas *et al.*, 1998; Zenkin *et al.*, 2006). Thus, the backtracked state is associated with an alternative activity of RNAP and is essential for faithful, well regulated transcription.

Crystallographic studies on *Saccharomyces cerevisiae* RNAP II and *Thermus thermophilus* RNAP have revealed the structures of the eukaryotic and bacterial ECs, respectively (Gnatt *et al.*, 2001; Vassilyev *et al.*, 2007). Although crystal structures of *S. cerevisiae* RNAP II in backtracked states have been reported (Cheung & Cramer, 2011; Wang *et al.*, 2009), the structure of a bacterial backtracked complex is not known. To elucidate the structural basis of backtracking and transcript cleavage by a bacterial RNAP, we crystallized backtracked complexes using the RNAP from *T. thermophilus*. In this report, the reconstitution and crystallization of the backtracked complexes and their preliminary X-ray crystallographic analyses are described.

2. Methods

2.1. Nucleic acid scaffolds

The nucleic acid scaffolds for the reconstitution of the backtracked complexes were designed to include a 9–10 bp DNA/RNA hybrid with 5' and 3' overhangs of single-stranded RNA and an 18 bp downstream DNA duplex (Fig. 1). In Bt-1 and Bt-2, one and two nucleotide residues, respectively, of the RNA 3' end were designed to be mismatched with the template nucleotides. The sequences of the nucleic acids are as follows. The RNA oligomers were: RNA1, CCAGCCGGCGCUCGCA; RNAS1, CCAGCCGGCGCUCsGsCA (where 's' denotes the phosphorothioate linkage); RNA2, GCCAGCCGGCGCUCGA; RNAS2, GCCAGCCGGCGCUCsGsGA. The DNA oligomers were: Temp1, GTCACCTACCAAGCTACGCGAGCGCCG; NT1, GTAGCTTGTTGGTAGTGA-CGAG; Temp2, TCACTACCACAAGCTACGAGAGCGCCGG; NT2, CGTAGCTTGTTGGTAGTGAAGA. Synthetic nucleic acids were purchased from Thermo Scientific, Sigma–Aldrich or Japan Bio Services.

2.2. Assembly of backtracked complexes

The RNAP core enzyme was purified from *T. thermophilus* cells as described previously (Vassilyeva *et al.*, 2002). For the *in vitro* reconstitution of the backtracked complexes, 25 μ M RNA, 25 μ M template DNA and 25 μ M nontemplate DNA were mixed with 20 μ M RNAP in 10 mM Tris–HCl buffer pH 7.7 containing 150 mM sodium chloride and 1% glycerol. The mixture was incubated at 343 K for 5 min and gradually cooled to 293 K. The formation of the complexes was confirmed by native PAGE (Fig. 2).

2.3. Transcript-cleavage assays

For the measurement of the intrinsic RNA-cleavage activity, backtracked complexes (BC-1 and BC-2) were assembled by mixing 1 μ M of the nucleic acid scaffold with 1.5 μ M RNAP in 50 mM Tris–HCl buffer pH 7.7 containing 100 mM sodium chloride and preheated at 328 K. To start the cleavage reaction, magnesium chloride was added to the solution to a final concentration of 10 mM. Aliquots (4 μ l) were withdrawn at the 2, 5, 10 and 20 min time points and the reactions were quenched by mixing the aliquot with an equal volume of Tris–borate electrophoresis buffer containing 50 mM EDTA and 10 M urea. The nucleic acids were separated by denaturing PAGE on

a 20% gel containing 7 M urea and were stained with SYBR Gold (Molecular Probes).

2.4. Crystallization

To avoid RNA cleavage during crystallization, the backtracked complexes were crystallized either in the absence of magnesium ions or by using RNAs with phosphorothioate bonds instead of phosphodiester bonds near the cleavage sites (RNAS1 and RNAS2; Fig. 3). For screening of the initial crystallization conditions, the Crystal Screen, Crystal Screen 2, Natrix, MembFac (Hampton Research) and Wizard I and II (Emerald BioSystems) kits were used. The screening was performed by the sitting-drop vapour-diffusion technique at 293 K. For optimization of the crystallization conditions, the hanging-drop vapour-diffusion technique was used. A 1 μ l portion of sample solution consisting of 20 μ M RNAP and 25 μ M nucleic acid scaffold was mixed with an equal volume of reservoir solution and was equilibrated against 200 μ l reservoir solution at 293 K. The final reservoir conditions for the backtracked complexes were 50 mM Tris–HCl buffer pH 8.5, 150 mM calcium acetate, 2.7–3.0% polyethylene glycol (PEG) 8000, 30% 1,6-hexanediol. In the final conditions, normal RNA (RNA1 or RNA2), rather than the RNA containing phosphorothioate bonds (RNAS1 or RNAS2), was used in the absence of magnesium. For the diffraction experiments, crystals grown for 4 d were washed in reservoir solution supplemented with 10% ethanol and flash-cooled in liquid nitrogen.

3. Results

3.1. Reconstitution of the backtracked complexes

We designed two nucleic acid scaffolds, Bt-1 (RNA1/Temp1/NT1) and Bt-2 (RNA2/Temp2/NT2): the RNA 3'-end base is designed to mismatch the template DNA base in Bt-1, while the two bases at the RNA 3' end are mismatched with the template in Bt-2 (Fig. 1). The backtracked transcription complexes were reconstituted by assembling *T. thermophilus* RNAP and the nucleic acid scaffolds Bt-1 and Bt-2. The assembly of the backtracked complexes was confirmed by native PAGE (Fig. 2). The complexes assembled with scaffolds Bt-1 and Bt-2 are hereafter referred to as BC-1 and BC-2, respectively.

Figure 1

Schematic representation of the nucleic acid scaffolds for reconstitution of the backtracked complexes.

Figure 2

Reconstitution of BC-1 and BC-2. The complexes were analysed by 5% native PAGE. The nucleic acids were stained with SYBR Gold and the proteins were stained with SimplyBlue SafeStain (Invitrogen).

Figure 3
RNA cleavage by the backtracked complexes. (a) BC-1 and (b) BC-2 were incubated in the presence of 10 mM magnesium chloride for the indicated times at 328 K. In the lanes marked <S>, BC-1 and BC-2 assemblies using phosphorothioate-containing RNAs, RNAS1 and RNAS2, respectively, were analysed. (c) Schematic representation of the translocational states of BC-1 and BC-2.

3.2. Characterization of the *in vitro* assembled complexes

To determine whether the assembled complexes properly represent a backtracked state, their transcript-cleavage activities were analysed. When BC-1 was incubated in the presence of magnesium ions, the initial 16-nucleotide RNA (RNA1) was converted to a 14-nucleotide RNA (Fig. 3a). This indicated that BC-1 is in the proper one-nucleotide backtracked state (Fig. 3c, left) and the hydrolysis of the 16-nucleotide RNA at the RNAP catalytic site yielded the 14-nucleotide RNA by removing a dinucleotide RNA. On the other

hand, the transcript cleavage was less efficient in BC-2 than in BC-1. The cleavage reaction in BC-2 yielded two RNA products of 13 and 14 nucleotides in length, indicating that RNAP hydrolytically removes three and two nucleotides, respectively, from the original 16-nucleotide RNA (RNA2; Fig. 3b). Thus, BC-2 appears to oscillate between the one- and two-nucleotide backtracked states (Fig. 3c, right). Judging from the electrophoretic band intensities, the two states exist in nearly the same proportions. In the one-nucleotide backtracked state of BC-2, the penultimate nucleotide of the RNA (residue 15) is presumed to reside in the 'pre-translocation position' (position +1 relative to the catalytic site) and is mismatched with the template nucleotide.

3.3. Crystallization

We began our study by screening initial crystallization conditions for the backtracked complexes, as the 3' overhang of RNA potentially causes some changes in the RNAP conformation, which may differ from that in the normal EC. The initial screening for the crystallization conditions of BC-1 and BC-2 was performed by the standard vapour-diffusion method with commercially available screening kits. Bipyramidal crystals of up to 1 mm in size were obtained after 4–8 weeks of incubation with various precipitants containing 1,6-hexanediol or 2-methyl-2,4-pentanediol. The precipitant and the shape of the crystals were similar to those of previously reported EC crystals (Kashkina *et al.*, 2006; Vassilyev *et al.*, 2007). However, the crystals of the backtracked complexes grew quite slowly and they diffracted X-rays poorly, in spite of their large size. The optimal crystallization conditions for these backtracked complexes, which included the 3' overhang of RNA, may differ from those of the active ECs. Therefore, we further optimized the crystallization conditions and the nucleic acid scaffolds as follows.

We searched for additives to accelerate the crystal growth. As a result of the additive screening, we found that high-molecular-weight polyethylene glycols such as PEG 4000 and PEG 8000 efficiently promoted crystal growth. In the presence of 2–4% PEG 8000 and 30% 1,6-hexanediol, crystals appeared in 1–2 d and grew in 4–6 d with high reproducibility. We decided to use the precipitant in the final conditions: 50 mM Tris-HCl buffer pH 8.5, 150 mM calcium acetate, 2.7–3.0% PEG 8000, 30% 1,6-hexanediol.

We also modified the nucleic acid scaffolds to obtain better diffraction. To maximize the crystal contacts, we used a downstream DNA duplex that was 5 bp longer than that in the previously reported EC (Vassilyev *et al.*, 2007) and the nontemplate strand was further

Figure 4
Crystals of (a) BC-1 and (b) BC-2.

Table 1

Statistics of the diffraction data for the BC-1 and BC-2 crystals.

Values in parentheses are for the highest-resolution shell.

Complex	BC-1	BC-2
Beamline	SPRING-8 BL41XU	Photon Factory NE3A
Space group	$P4_12_12$	$P4_12_12$
Unit-cell parameters (Å)	$a = b = 156.0$, $c = 495.0$	$a = b = 156.9$, $c = 496.2$
Resolution (Å)	50–3.7 (3.83–3.70)	50–3.4 (3.52–3.40)
Unique reflections	65538 (6390)	85469 (8353)
Multiplicity	6.3 (4.8)	5.8 (4.1)
Mean $I/\sigma(I)$	7.2 (2.7)	7.7 (2.6)
$R_{\text{merge}}^{\dagger}$ (%)	17.4 (56.3)	17.6 (53.9)
Completeness (%)	98.9 (98.6)	99.0 (98.8)

$\dagger R_{\text{merge}} = \sum_{hkl} \sum_i |I_i(hkl) - \langle I(hkl) \rangle| / \sum_{hkl} \sum_i I_i(hkl)$, where $\langle I(hkl) \rangle$ is the mean value of multiple observations of the intensity $I_i(hkl)$ of the symmetry-related reflections.

extended by three nucleotide residues. The modified scaffolds improved the crystal quality, allowing us to obtain the 3.4 Å diffraction data set for BC-2. We also examined crystallization with a scaffold containing RNA with a nine-nucleotide 3' overhang. However, the crystals of the 'long backtracked' complex diffracted X-rays poorly (data not shown).

3.4. Data collection and preliminary crystallographic analyses

Diffraction data sets were collected from cryocooled crystals of BC-1 and BC-2 (Fig. 4), using synchrotron radiation on beamlines BL41XU of SPRING-8, Harima, Japan and NE3A of the Photon Factory, Tsukuba, Japan. The resolutions were 3.7 and 3.4 Å, respectively. The data were indexed, integrated and scaled with the *HKL-2000* program (Otwinowski & Minor, 1997). The data-collection statistics are provided in Table 1. The estimated mosaicities per frame were within 0.6–1.0° for BC-1 and 0.4–0.5° for BC-2. Although the crystals appeared to belong to space group $P4_12_12$, this is owing to merohedral twinning of $P4_1$ crystals. Twinning analyses with *phenix.xtriage* from the *PHENIX* suite (Adams *et al.*, 2010) revealed high twin-fraction values. For the BC-1 crystal, the values were estimated to be 0.30, 0.28 and 0.44 according to the Britton plot, the *H*-test and the maximum-likelihood method, respectively. For the BC-2 crystal, they were 0.35, 0.31 and 0.44, respectively. Moreover, the values of the second moment of intensity ($\langle I^2 \rangle / \langle I \rangle^2$) for acentric reflections were 1.3 for the BC-1 data and 1.4 for the BC-2 data. These values are indicative of perfectly twinned data, as the value should be 2 for untwinned data and 1.5 for perfectly twinned data. Therefore, we decided to process our data by assuming perfect twinning. We scaled the data in the $P4_12_12$ space group and expanded them to the $P4_1$ space group, as in the case of the previously reported EC structures (Vassylyev *et al.*, 2007). Refinement was performed with these expanded data, assuming perfect twinning. These data provided better electron density than the data scaled in the $P4_1$ space

group (not shown). For both the BC-1 and the BC-2 crystals, the Matthews coefficient (Matthews, 1968) and the solvent content were 3.8 Å³ Da^{−1} and 68%, respectively. The preliminary crystallographic refinement revealed clear electron densities for RNAP and its bound nucleic acids, including the backtracked residue(s) of RNA (not shown). Structure determination and refinement are under way.

This work is based on experiments performed at SPRING-8 with the approval of the Japan Synchrotron Radiation Research Institute (proposal Nos. 2008A1244, 2008B1193, 2009A1118, 2009B1354, 2010B1465, 2011A1264, 2011B1382, 2012A1334 and 2012B1392) and at the Photon Factory with the approval of the Photon Factory Program Advisory Committee (proposal No. 2011G554). We thank the beamline staff of BL41XU at SPRING-8 and NE3A at the Photon Factory for assistance with data collection. This work was supported in part by a Japan Society for the Promotion of Science (JSPS) Grant-in-Aid for Scientific Research (to SS and SY) and the Targeted Proteins Research Program (TPRP), the Ministry of Education, Culture, Sports, Science and Technology (MEXT) of Japan. YM was supported by the RIKEN Junior Research Associate Program and the JSPS Global Centers of Excellence Program (Integrative Life Science Based on the Study of Biosignaling Mechanisms).

References

- Adams, P. D. *et al.* (2010). *Acta Cryst.* **D66**, 213–221.
 Cheung, A. C. M. & Cramer, P. (2011). *Nature (London)*, **471**, 249–253.
 Donahue, B. A., Yin, S., Taylor, J. S., Reines, D. & Hanawalt, P. C. (1994). *Proc. Natl Acad. Sci. USA*, **91**, 8502–8506.
 Erie, D. A., Hajiseyedjavadi, O., Young, M. C. & von Hippel, P. H. (1993). *Science*, **262**, 867–873.
 Gnat, A. L., Cramer, P., Fu, J., Bushnell, D. A. & Kornberg, R. D. (2001). *Science*, **292**, 1876–1882.
 Izban, M. G. & Luse, D. S. (1993). *J. Biol. Chem.* **268**, 12874–12885.
 Jeon, C. & Agarwal, K. (1996). *Proc. Natl Acad. Sci. USA*, **93**, 13677–13682.
 Kashkina, E., Anikin, M., Tahirov, T. H., Kochetkov, S. N., Vassylyev, D. G. & Temiakov, D. (2006). *Nucleic Acids Res.* **34**, 4036–4045.
 Komissarova, N. & Kashlev, M. (1997a). *Proc. Natl Acad. Sci. USA*, **94**, 1755–1760.
 Komissarova, N. & Kashlev, M. (1997b). *J. Biol. Chem.* **272**, 15329–15338.
 Marr, M. T. & Roberts, J. W. (2000). *Mol. Cell*, **6**, 1275–1285.
 Matthews, B. W. (1968). *J. Mol. Biol.* **33**, 491–497.
 Nudler, E., Mustaev, A., Lukhtanov, E. & Goldfarb, A. (1997). *Cell*, **89**, 33–41.
 Otwinowski, Z. & Minor, W. (1997). *Methods Enzymol.* **276**, 307–326.
 Reines, D. (1992). *J. Biol. Chem.* **267**, 3795–3800.
 Thomas, M. J., Platas, A. A. & Hawley, D. K. (1998). *Cell*, **93**, 627–637.
 Vassylyev, D. G., Vassylyeva, M. N., Perederina, A., Tahirov, T. H. & Artsimovitch, I. (2007). *Nature (London)*, **448**, 157–162.
 Vassylyeva, M. N., Lee, J., Sekine, S., Laptchenko, O., Kuramitsu, S., Shibata, T., Inoue, Y., Borukhov, S., Vassylyev, D. G. & Yokoyama, S. (2002). *Acta Cryst.* **D58**, 1497–1500.
 Wang, D., Bushnell, D. A., Huang, X., Westover, K. D., Levitt, M. & Kornberg, R. D. (2009). *Science*, **324**, 1203–1206.
 Zenkin, N., Yuzenkova, Y. & Severinov, K. (2006). *Science*, **313**, 518–520.