



저작자표시-비영리-변경금지 2.0 대한민국

이용자는 아래의 조건을 따르는 경우에 한하여 자유롭게

- 이 저작물을 복제, 배포, 전송, 전시, 공연 및 방송할 수 있습니다.

다음과 같은 조건을 따라야 합니다:



저작자표시. 귀하는 원저작자를 표시하여야 합니다.



비영리. 귀하는 이 저작물을 영리 목적으로 이용할 수 없습니다.



변경금지. 귀하는 이 저작물을 개작, 변형 또는 가공할 수 없습니다.

- 귀하는, 이 저작물의 재이용이나 배포의 경우, 이 저작물에 적용된 이용허락조건을 명확하게 나타내어야 합니다.
- 저작권자로부터 별도의 허가를 받으면 이러한 조건들은 적용되지 않습니다.

저작권법에 따른 이용자의 권리는 위의 내용에 의하여 영향을 받지 않습니다.

이것은 [이용허락규약\(Legal Code\)](#)을 이해하기 쉽게 요약한 것입니다.

[Disclaimer](#)

이학박사 학위논문

**Single-molecule Fluorescence Study
Discovering 1-Dimensional Diffusion of
RNA Polymerase after Intrinsic
Termination**

내재 종결 이후 RNA 중합효소의 일차원
확산을 발견한 단일 분자 형광 연구

2021 년 2 월

서울대학교 대학원

물리천문학부

강 우 영


Single-molecule Fluorescence Study Discovering 1-Dimensional Diffusion of RNA Polymerase after Intrinsic Termination


지도교수 홍 성 철

이 논문을 이학박사 학위논문으로 제출함
2021 년 01 월

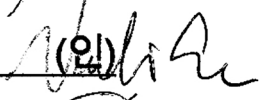
서울대학교 대학원
물리천문학부
강 우 영


강우영의 이학박사 학위논문을 인준함
2020 년 12 월

위 원 장 박 혜 윤 (인) 

부위원장 홍 성 철 (인) 

위 원 홍 승 훈 (인) 

위 원 이 남 기 (인) 

위 원 강 창 원 (인) 

Ph.D. Dissertation

**Single-molecule Fluorescence Study
Discovering 1-Dimensional Diffusion of
RNA Polymerase after Intrinsic
Termination**

Wooyoung Kang

Research Advisor: Professor Sungchul Hohng

February 2021

Department of Physics and Astronomy

Graduate School

Seoul National University

Abstract

Single-molecule Fluorescence Study Discovering 1-Dimensional Diffusion of RNA Polymerase after Intrinsic Termination

Wooyoung Kang

Major in Physics

Department of Physics and Astronomy

The Graduate School

Seoul National University

Transcription is the process of synthesizing RNA from DNA by RNA polymerase (RNAP) and is essential in regulating gene expression. Previously, it has been understood that the transcription consists of three stages: initiation, elongation, and termination. Extensive studies have been conducted on transcription initiation, and much mechanism has been known. On the other hand, there are a few studies on the termination, and the understanding is still insufficient. In particular, it is not known whether RNA or RNAP will dissociate first from DNA or at the same time. We observed the intrinsic terminator using *E. coli* RNAP through the single-molecule fluorescence experiments.

The results showed that the RNA transcript was first released from DNA, and

most RNAP remained on DNA. The remaining RNAP diffuses one-dimensionally on DNA and dissociates from DNA after several tens of seconds after RNA release. Also, since the 1D diffusion coefficient of the RNAP after intrinsic termination increases as the salt concentration increases, it can be found that the remaining RNAP diffuses on DNA through a hopping mechanism.

As the transcription termination refers to the release of an RNA transcript from DNA, the retention of RNAP constitutes a new fourth stage of transcription, named the recycling stage. Also, the remaining RNAP was designated as the recycling RNAP. During the recycling stage, it was observed through the fluorescent signal that the recycling RNAP could initiate transcription again when it encounters promoters, which we defined as the transcription reinitiation. The reinitiation occurs both upstream and downstream of the termination site, and antisense reinitiation also occurs. This new transcription initiation mechanism is expected to increase the efficiency of transcription of adjacent genes in prokaryotes and to increase also the transcription efficiency of the clustered gene in eukaryotes.

Keywords: Single-molecule fluorescence spectroscopy, Protein induced fluorescence enhancement (PIFE), E. coli RNA polymerase, Transcription termination

Student Number: 2014-21368

Contents

Abstract.....	i
----------------------	----------

Contents.....	iii
----------------------	------------

Chapter 1

Introduction.....	1
--------------------------	----------

1.1. Intrinsic termination.....	1
---------------------------------	---

1.2. Single-Molecule Experiments.....	4
---------------------------------------	---

References.....	7
-----------------	---

Chapter 2

E. coli RNAP diffuses one-dimensionally on DNA after intrinsic termination.....	11
--	-----------

2.1. Introduction.....	11
------------------------	----

2.2. Materials and Methods.....	13
---------------------------------	----

2.2.1. DNA templates preparation.....	13
---------------------------------------	----

2.2.2. Single-molecule transcription termination experiment.....	17
--	----

2.3. Results	19
--------------------	----

2.3.1. Fluorescent detection of intrinsic termination.....	19
--	----

2.3.2. One-dimensional diffusion of RNA free RNAP on DNA.....	26
---	----

2.4. Conclusion.....	28
----------------------	----

References.....	30
-----------------	----

Chapter 3

The recycling RNAP one-dimensional diffuses on DNA using a hopping mechanism.....33

3.1. Introduction.....	33
3.2. Materials and Methods.....	35
3.2.1. Protein preparation.....	35
3.2.2. DNA templates preparation.....	35
3.2.3. Single-molecule transcription termination experiments.....	39
3.3. Results.....	41
3.3.1. Measurement of the recycling RNAP retention time.....	41
3.3.2. 1D diffusion coefficient of the recycling RNAP.....	46
3.4. Conclusion.....	52
References.....	53

Chapter 4

The recycling RNAP can initiate transcription again without dissociation from DNA.....55

4.1. Introduction.....	55
4.2. Materials and Methods.....	56
4.2.1. DNA preparation.....	56
4.2.2. Single-molecule experiments for detecting reinitiation.....	61
4.3. Results.....	63
4.3.1. Fluorescent detection of upstream and downstream reinitiation.....	63
4.3.2. Fluorescent detection of antisense reinitiation.....	72

3.4. Conclusion.....	80
References.....	83
Appendix.....	87
A. The measurements of EcoRI E111Q binding efficiency.....	87
B. The modeling of the recycling RNAP diffusing on DNA.....	89
C. The measurement of σ^{70} retention efficiency.....	93
References.....	95
Abstract in Korean (국문초록).....	96

Chapter 1

Introduction

1.1. Intrinsic termination

A transcription terminator is a DNA region that indicates the end of genes or operon during transcription, which is a transcription termination signal. This process is essential for gene regulation and removal of RNA polymerase (RNAP), allowing new transcription to begin. In bacteria, there are two types of transcription terminator; Rho-dependent terminator and intrinsic terminator. Rho-dependent terminator requires the adenosine triphosphate (ATP)-dependent RNA-DNA helicase, known as Rho, which binds to nascent RNA and dissociates the RNAP-RNA-DNA complex. In contrast, an intrinsic terminator, or factor-independent terminator, only needs the RNA structure without any proteins.

The intrinsic terminator consists of a GC-rich hairpin structure and a 7-8 nucleotide U-rich tract (Figure 1.1). Transcription termination caused by the

intrinsic terminator, named as intrinsic termination, is accomplished through the following steps; (i) pausing at the U-rich tract, (ii) hairpin structure formation, (iii) hybrid-shearing or hyper-translocation, and (iv) EC dissociation (Ray-soni et al., 2016).

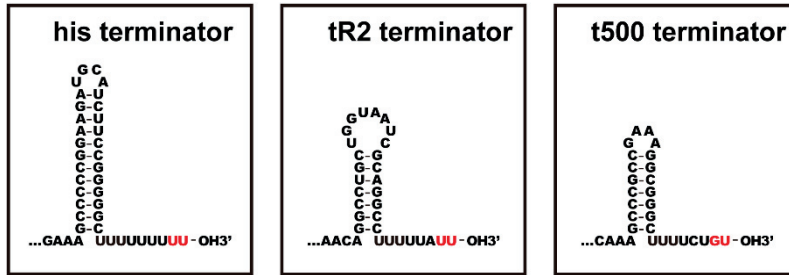


Figure 1.1. Three representative intrinsic terminator sequences and expected secondary structures. General intrinsic terminators have a hairpin loop and a U-tract. The transcript termination positions are marked in red.

When the RNAP encounters the intrinsic terminator U-rich tract, it pauses for a short time (Gusarov et al., 1999; Yarnell et al., 1999). This pausing time provides the time window for RNA to form hairpin structures, consequently allowing the elongation complex (EC) to go to the termination pathway (Gusarov et al., 1999). When the mutation occurred in the U-rich tract makes the pause at the termination site abrogate, transcription termination also does not happen (Gusarov et al., 1999).

During the pausing of EC at the U-rich tract, hairpin structure rapidly forms except for the bottom 2-3 base pairs (Lubkowska et al., 2011). It is known that

these partial formations of hairpin are remarkably similar to a hairpin pause, which is expected to stabilize the pre-termination pause before complete hairpin formation (Chan et al., 1997; Lubkowska et al., 2011). The complete hairpin formation causes the upstream of the DNA-RNA hybrid to melt 3-4 base pairs (Gusarov et al., 1999; Komissarova et al., 2002). Since uracil-lacking in this part makes the energetic barrier for hairpin formation higher, U's in the first three positions of the U-rich tract of intrinsic terminator are highly conserved (d'Aubenton Carafa et al., 1990).

There are two models of how hairpin formation, which induces the partial melting of DNA-RNA hybrid, affects the EC. The first model is the hybrid-shearing model, in which RNA slips in a DNA-RNA hybrid as the hairpin formation pulls it (Komissarova et al., 2002). The second one is hyper-translocation, where RNAP moves from the termination site to downstream by 2-4nt without RNA elongation (Yarnell et al., 1999; Santangelo et al., 2004). The single-molecule magnetic tweezer experiment revealed that both models might be possible according to the termination sequence (Larson et al., 2008). Thus, the hairpin formation causes rearrangement of RNA, DNA, and RNAP, which probably irreversibly inactivates EC, and, in turn, causes termination to occur.

The final step of intrinsic termination is EC dissociation, meaning the dissociation of RNAP and RNA from DNA. Previous studies have shown that conformational changes that loosen contact with DNA-RNA hybrid are essential in this process (Gnatt et al., 2001; Tagami et al., 2010; Chakraborty

et al., 2012; Weixlbaumer et al., 2013). However, how RNAP rearrangement destabilizes the EC remains unclear, and besides, the order of RNA, DNA, and RNAP release also remains unknown.

1.2. Single-Molecule Experiments

In order to understand the life process, it is necessary to understand the dynamics of biological molecules, such as DNA, RNA, and protein. While traditional biological technics has given the averaged information of biological molecules, advances in technology have allowed us to observe the dynamics of biological molecules at the single-molecule level. In particular, the single-molecule technics based on fluorescence microscopy has been used in various fields of biology since its potential shown in the mid-1990s (Ha et al., 1996). This technic uses two phenomena to measure the distance between molecules or within molecules.

At first, fluorescence resonance energy transfer (FRET) refers to a phenomenon in which energy is transferred between two fluorescent dyes, termed as donor and acceptor, through non-radiative dipole-dipole coupling (Helms 2008). The efficiency of FRET can be measured as the ratio of the intensity of the acceptor to total emission intensity. It is very suitable for measuring the conformation change of the biological molecules because it varies very sensitively in the range of 5 to 10 nm. In general, two fluorescent dyes are used to measure the distance between them. However, when three or four fluorescent dyes are used simultaneously, it is possible to understand

more complex dynamics by simultaneously measuring three or six different distances (Hohng et al., 2004; Lee et al., 2010).

Protein-induced fluorescence enhancement (PIFE) is a phenomenon in which some organic dyes, especially cyanine dyes, increase their intensity when protein approaches dyes. There are cis- or trans- isomers in cyanine dye, and their quantum yields are different. Also, these two isomers' ratio depends on the local environmental viscosity and temperature (Aramendia et al., 1994). The protein's approach near the cyanine dye acts as an additional viscosity factor, increasing the intensity of the dye. A previous study showed that this phenomenon could be used to detect the distance between a fluorescent dye and a protein at the single-molecule level (Luo et al., 2007). After that, it has been widely used to measure the dye-protein distance at 0-3nm, which is less than that of FRET.

Total internal reflection fluorescence (TIRF) microscopy is widely used to measure single-molecule fluorescence. TIRF microscopy can dramatically reduce background fluorescence by creating an evanescent wave so that excitation occurs only at 100-200 nm from the surface to which the sample is attached. The fluorescent signal is separated into donors and acceptors by a dichroic mirror collected by an electron-multiplying charge-coupled device (EM-CCD) camera with high quantum efficiency and low readout noise (Figure 1.2).

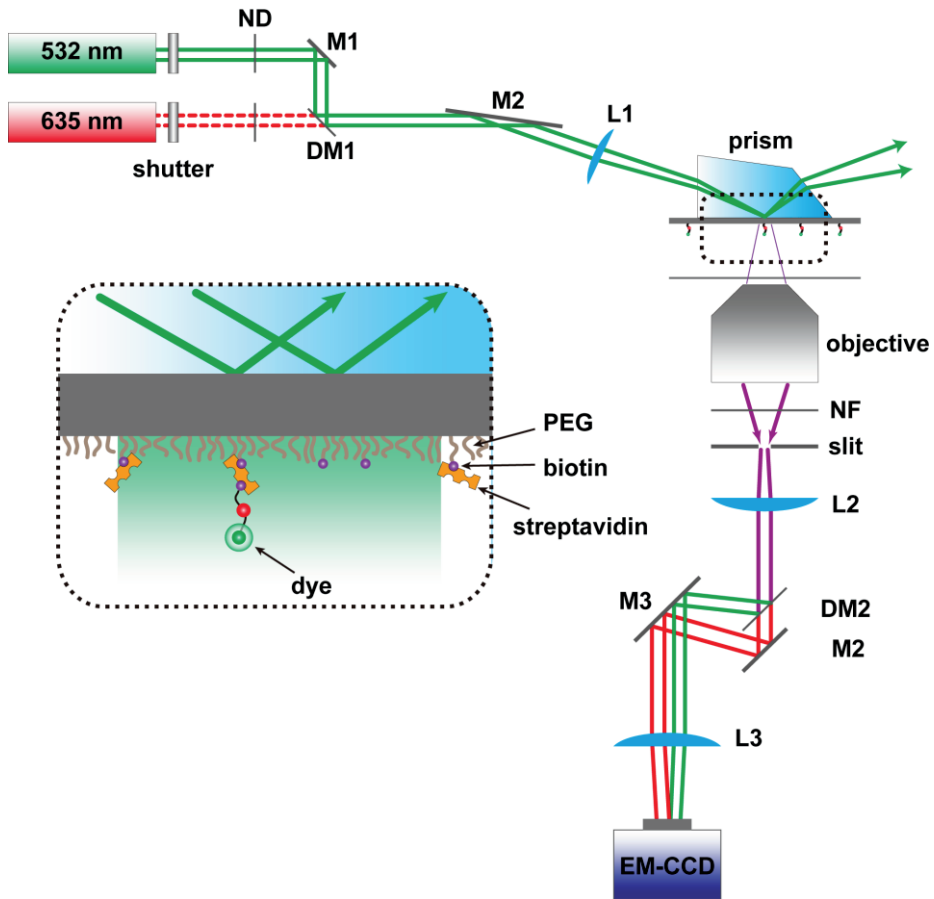


Figure 1.2. Schematics of Prism-type TIRF Microscope. Dye-excitation laser beam paths were united by using a dichroic mirror (DM1). Neutral density filters (ND) were used for adjusting laser power. Shutters were controlled via DAQ board for ALEX mode. Notch filters (NF) were placed after the objective lens (UPlanSApo 60X, Olympus) to block reflected laser beams. Dichroic mirror (DM2) was also used to separate each dye signal spatially.

References

- Aramendia, P.F. Negri, R.M. & Sanroman, E. (1994) Temperature-Dependence of Fluorescence and Photoisomerization in Symmetrical Carbocyanines - Influence of Medium Viscosity and Molecular-Structure. *The Journal of Physical Chemistry C*. 98, 3165–3173.
- Chakraborty, A., Wang, D., Ebright, Y. W., Korlann, Y., Kortkhonjia, E., Kim, T., Chowdhury, S., Wigneshweraraj, S., Irschik, H., Jansen, R., Nixon, B. T., Knight, J., Weiss, S., & Ebright, R. H. (2012). Opening and closing of the bacterial RNA polymerase clamp. *Science*, 337(6094), 591–595.
- Chan, C. L., Wang, D., & Landick, R. (1997). Multiple interactions stabilize a single paused transcription intermediate in which hairpin to 3' end spacing distinguishes pause and termination pathways. *Journal of molecular biology*, 268(1), 54–68.
- d'Aubenton Carafa, Y., Brody, E., & Thermes, C. (1990). Prediction of rho-independent Escherichia coli transcription terminators. A statistical analysis of their RNA stem-loop structures. *Journal of molecular biology*, 216(4), 835–858.
- Gnatt, A. L., Cramer, P., Fu, J., Bushnell, D. A., & Kornberg, R. D. (2001). Structural basis of transcription: an RNA polymerase II elongation complex at 3.3 Å resolution. *Science*, 292(5523), 1876–1882.
- Gusarov, I., & Nudler, E. (1999). The mechanism of intrinsic transcription termination. *Molecular cell*, 3(4), 495–504.
- Ha, T., Enderle, T., Chemla, S., Selvin, R., & Weiss, S. (1996). Single Molecule Dynamics Studied by Polarization Modulation. *Physical review letters*, 77(19), 3979–3982.
- Helms, V. (2008). Fluorescence Resonance Energy Transfer. Principles of Computational Cell Biology (pp 202). Wiley-Blackwell.
- Hohng, S., Joo, C., & Ha, T. (2004). Single-molecule three-color FRET. *Biophysical journal*, 87(2), 1328–1337.

Chapter 1

Komissarova, N., Becker, J., Solter, S., Kireeva, M., & Kashlev, M. (2002). Shortening of RNA:DNA hybrid in the elongation complex of RNA polymerase is a prerequisite for transcription termination. *Molecular cell*, 10(5), 1151–1162.

Larson, M. H., Greenleaf, W. J., Landick, R., & Block, S. M. (2008). Applied force reveals mechanistic and energetic details of transcription termination. *Cell*, 132(6), 971–982.

Lee, J., Lee, S., Ragunathan, K., Joo, C., Ha, T., & Hohng, S. (2010). Single-molecule four-color FRET. *Angewandte Chemie (International ed. in English)*, 49(51), 9922–9925.

Lubkowska, L., Maharjan, A. S., & Komissarova, N. (2011). RNA folding in transcription elongation complex: implication for transcription termination. *The Journal of biological chemistry*, 286(36), 31576–31585.

Luo, G., Wang, M., Konigsberg, W. H., & Xie, X. S. (2007). Single-molecule and ensemble fluorescence assays for a functionally important conformational change in T7 DNA polymerase. *Proceedings of the National Academy of Sciences of the United States of America*, 104(31), 12610–12615.

Ray-Soni, A., Bellecourt, M. J., & Landick, R. (2016). Mechanisms of Bacterial Transcription Termination: All Good Things Must End. *Annual review of biochemistry*, 85, 319–347.

Santangelo, T. J., & Roberts, J. W. (2004). Forward translocation is the natural pathway of RNA release at an intrinsic terminator. *Molecular cell*, 14(1), 117–126.

Tagami, S., Sekine, S., Kumarevel, T., Hino, N., Murayama, Y., Kamegamori, S., Yamamoto, M., Sakamoto, K., & Yokoyama, S. (2010). Crystal structure of bacterial RNA polymerase bound with a transcription inhibitor protein. *Nature*, 468(7326), 978–982.

Weixlbaumer, A., Leon, K., Landick, R., & Darst, S. A. (2013). Structural basis of transcriptional pausing in bacteria. *Cell*, 152(3), 431–441.

Yarnell, W. S., & Roberts, J. W. (1999). Mechanism of intrinsic transcription termination and antitermination. *Science*, 284(5414), 611–615.

Chapter 2

E. coli RNAP diffuses one- dimensionally on DNA after intrinsic termination

2.1. Introduction

Transcription is the first gene expression process in which RNAP copies DNA into RNA, especially messenger RNA. Transcription consists of three stages. The first stage is the transcription initiation. In this stage, RNAP binds to specific DNA sequences, known as the promoter, and generates a transcription bubble. After that, RNAP starts to synthesize RNA, but it repeats the synthesis and release of short RNA, called abortive initiation (Revyakin et al., 2006). When RNAP escapes the promoter region, it forms a stable elongation complex consisting of RNA, DNA, and RNAP. In the second stage, named transcription elongation, RNAP adds nucleotide triphosphate (NTP), complementary to the one DNA strand, to RNA. The final stage is transcription termination. When the elongation complex encounters the specific region known as the DNA terminator, it causes the temporal pause

and disassembly of the elongation complex (Figure 2.1).

In prokaryotic, RNAP consists of a core enzyme made of five subunits and σ factor, essential in the transcription initiation. Although transcription factors, such as NusA or NusG, affect the bacterial transcription, the entire transcription process can occur with the only RNAP. Initiation of transcription requires the promoter and σ factor. Also, intrinsic termination occurs only by the RNA structure consisting of GC-rich hairpin and U-tract (Farnham et al., 1981; Wilson et al., 1995; Czyz et al., 2014; Mondal et al., 2016), and transcription factors can only affect the efficiency of intrinsic termination.

All termination models suggest a structural change (Gusarov et al., 1999; Ray-Soni et al., 2017) during intrinsic termination. Although the structural changes have not been characterized yet, it is considered that the structural changes destabilize the DNA-RNA hybrid resulting in partial separation of RNA and DNA and the formation of RNA hairpin. However, despite these models, it remains unknown how the elongation complexes are disassembled. In particular, it is unknown in what order RNAP, DNA, and RNA will dissociate from the elongation complex or at the same time. Using single-molecule fluorescence measurements, we observed the release of RNA and RNAP from DNA and examined their post-terminational fate during *E. coli* intrinsic termination.

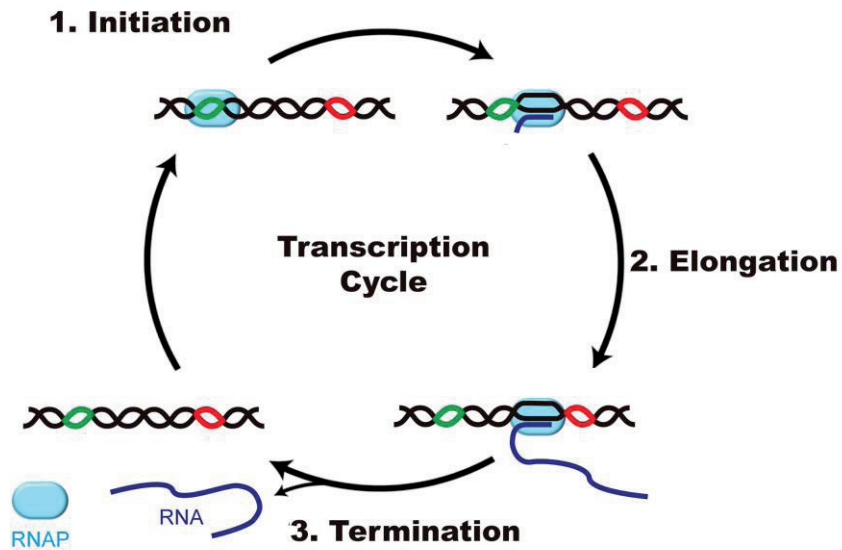


Figure 2.1. Three stages of the transcription cycle. The first stage is initiation, where RNAP (cyan oval) binds DNA (black line) at a promoter (green line). After initiation, RNAP incorporates several NTP into RNA (blue line) and escapes promoter. The second stage is elongation, where RNAP advances downward and extends RNA. The third stage is termination, where RNAP pauses at a terminator (red line), and RNA and RNAP dissociates from DNA. RNAP removed from DNA can rebind to other promoters and start transcription again.

2.2. Materials and Methods

2.2.1. DNA templates preparation

In order to observe the transcription termination by single-molecule

fluorescence, biotin and Cy5 double-labeled DNA template was prepared. This DNA template consists of a 50-bp upstream part including the strong promoter A1 of bacteriophage T7, a 38-bp transcription unit with intrinsic terminator tR2 of phage λ , and a 15-bp downstream part. It was pre-labeled with biotin at 5'-end of the nontemplate strand and Cy5 at 5'-end of the template strand. This DNA was named L+15 (Figure 2.2). The DNA whose downstream part was extended to 112-bp, including EcoRI binding site, was named L+112, and the DNA where the position of biotin and Cy5 were changed in L+112 was named L+112R (Figure 2.3). The DNA template with a 50-bp upstream part, including the strong promoter A1 of bacteriophage T7, and a 31-bp transcription unit without any terminator was named as L+15M. The DNA template with a 50-bp upstream part, including the strong promoter A1 of bacteriophage T7, a 257-bp transcription unit with intrinsic terminator tR2 of phage λ , and a 15-bp downstream part was named as T257/L+15 (Figure 2.3). Other DNA templates in which the tR2 terminator was replaced with *E. coli* his attenuator or phage ϕ 82 t500 terminator were named his+62 and t500+62, respectively.

All DNA templates were prepared by polymerase chain reactions (PCR) using AccuPower ProFi Taq PCR premix from Bioneer, Korea. All PCR products were purified using the Cleanup kit Expin PCR SV mini from GeneAll, Korea. All amplification templates and primers labeled with Cy5 or biotin were purchased from Integrated DNA Technologies, USA (Table 2.1 and 2.2).

For L+15, DNA_template_0, B_primerF and 5_primerR_L+15 were used. For L+112, DNA_template_1, B_primerF and 5_primerR_L+112 were used. For L+112R, DNA_template_1, 5_primerF and B_primerR_L+112 were used. For L+15M, DNA_template_M, B_primerF and 5_primerR_L+15M were used. For T257/L+15, DNA_Template_257_1, DNA_Template_257_2, B_primerF and 5_primerR_T257/L+15 were used. For his+62, DNA_Template_his, B_primerF and 5_primerR_OT were used. For t500+62, DNA_Template_t500, B_primerF and 5_primerR_OT were used.



Figure 2.2. The sequence of DNA template L+15. L+15 with T7A1 promoter (yellow box) and tR2 terminator was prelabeled with biotin (black dot) at the 5'-end of the nontemplate strand and Cy5 (red dot) at the 5'-end of the template strand.

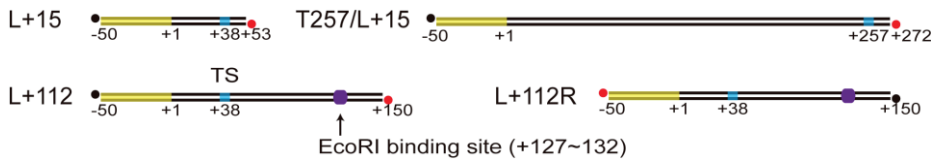


Figure 2.3. Design of various DNA templates. All DNA templates include T7A1 promoter (yellow box), termination site (TS, cyan oval), biotin (black dot), and Cy5 (red dot). L+112 and L+112R also include EcoRI recognition site (purple box).

Chapter 2

Table 2.1. Sequences of amplification templates used for PCR

Oligo name	Sequence (5' → 3')
DNA_template_0	TATCA AAAAG AGTAT TGACT TAAAG TCTAA CCTAT AGGAT ACTTA CAGCC ATCGA ACAGG CCTGC TGGTA ATCGC AGGCC TTTT ATTTG GGGGA GAGGG AAGTC ATGAA AAAAC TAACC TTTGA AATTC GATCT CCAGG ATCCA CCACC
DNA_template_1	TATCA AAAAG AGTAT TGACT TAAAG TCTAA CCTAT AGGAT ACTTA CAGCC ATCCC AAAGC CCGCC GAAAG GCGGG CTTT CTGTT TCTGG GCGGT GAAGT CATGA AAAAA CTAAC CTTTG AAATT CGATC TCCAG GATCC ACCAC C
DNA_template_M	TATCA AAAAG AGTAT TGACT TAAAG TCTAA CCTAT AGGAT ACTTA CAGCC ATGGC CTGCT GGTGA CTGAC TGACT GACTG AC
DNA_template_257_1	TATCA AAAAG AGTAT TGACT TAAAG TCTAA CCTAT AGGAT ACTTA CAGCC ATCGA ACAGG CCTCA AACAA AAGAA TGGAA TCAAA GTAA CTTCA AAATT AGACA CAACA TTGAA GATGG AAGCG TTCAA CTAGC AGACC ATTAT CAACA AAATA CTCCA ATTGG CGATG GCCCT GTCCT TTTAC CAGAC AACCA TTACC
DNA_template_257_2	TTTAC CAGAC AACCA TTACC TGTCC ACACA ATCTG CCCTT TCGAA AGATC CCAAC GAAAA GAGAG ACCAC ATGGT CCTTC TTGAG TTTGT AACAA CAGGC CTGCT GGTAA TCGCA GGCCT TTTTA TTTGG GGGAG AGGGA AG
DNA_template_his	TATCA AAAAG AGTAT TGACT TAAAG TCTAA CCTAT AGGAT ACTTA CAGCC ATCCG AAAGC CCCC GAAGA UGCAU CUUCC GGGGG CUUUU UUUUU TGGGC GGTGA AGTCA TGAAA AACT AACCT TTGAA ATTCG ATCTC CAGGA TCCAC CACC
DNA_template_t500	TATCA AAAAG AGTAT TGACT TAAAG TCTAA CCTAT AGGAT ACTTA CAGCC ATCCC AAAGC CCGCC GAAAG GCGGG CTTT CTGTT TCTGG GCGGT GAAGT CATGA AAAAA CTAAC CTTTG AAATT CGATC TCCAG GATCC ACCAC C

Table 2.2. Sequences of primers

Oligo name	Sequence (5' → 3')
B_primerF	Biotin-TATCA AAAAG AGTAT TGACT TAAAG TC
5_primerR_L+15	Cy5-CTTCC CTCTC CCCCC AATAA AAAG
5_primerR_L+112	Cy5-GCGAG ATTAC CATT A AGTGA A
5_primerF	Cy5-TATCA AAAAG AGTAT TGACT TAAAG TC
B_primer_L+112	Biotin-GCGAG ATTAC CATT A AGTGA A
5_primerR_L+15M	Cy5-GTCAG TCAGT CAGTC AGTCA CCAGC AG
5_primerR_T257/L+15	Cy5-GGTCT AAGCT TGTGG TGTCT TGTGT GT
5_primerR_OT	Cy5-GGTGG TGGAT CCTGG AGATC G

2.2.2. Single-molecule transcription termination experiments

In order to construct the fluorescent elongation complex, 50 nM of DNA template was mixed with *E. coli* RNA holoenzyme (20 nM, NEB) ATP, CTP, GTP (each 20 μ M, GE Healthcare), and Cy3-labeled ApU (250 μ M, TriLink). This mixture was incubated for 30 min at 37 °C in an initiation buffer (10 mM Tris-HCl, pH 8.0, 20 mM NaCl, 20 mM MgCl₂ and 1 mM dithiothreitol). Transcription was mostly initiated with Cy3-ApU, incorporated into the +1 and +2 positions of RNA, but transcription elongation stalled with 12-nucleotide RNA because of the absence of UTP.

Before immobilizing the stalled elongation complex on the quartz slides, quartz slides were coated with polyethylene glycol (PEG) to suppress non-specific binding (Sofia et al., 1998). At first, quartz slides were cleaned using piranha solution to remove organic residues, incubated with (3-aminopropyl)

trimethoxysilane (United Chemical Technologies), and coated with polymers by incubation in a 1:40 mixture of biotin-PEG-5000 and m-PEG-5000 (Laysan Bio.) (Roy et al., 2008). After that, the slides were treated with 0.2 mg/ml streptavidin (Invitrogen) for 5 min. The stalled elongation complexes were immobilized on polymer-coated quartz slides through biotin-streptavidin conjugation. Unbound RNAPs and un-immobilized elongation complexes were removed by washing with an initiation buffer and an imaging buffer extensively.

In order to image the single-molecule fluorescence from Cy3 and Cy5-labeled elongation complex, a home-made TIRF microscopy was used. A 532-nm green laser (EXLSR-532-50-CDRH, Spectra-physic) and a 640-nm red laser (EXLSR-640C-60-CDRH, Spectra-physic) were used for Cy3 excitation and Cy5 excitation, respectively. A single-molecule image was obtained by EM-CCD (Ixon DV897, Andor Technology) controlled by a customized C# program. All experiments were performed at 37 °C with 0.05 s exposure time in an alternating laser excitation (ALEX) mode. Because the experiments were performed in an ALEX mode, the actual time resolution was 0.1 s.

All experiments were performed in an oxygen scavenging buffer, or a transcription buffer (10 mM Tris-HCl, pH 8.0, 20 mM NaCl, 20 mM MgCl₂, 1 mM dithiothreitol, 5 mM 3,4-protocatechuic acid and, 100 nM protocatechuate-3,4-dioxygenase) with 200uM NTP each at 37 °C. When needed, 500nM NusA, NusG, or both were added to the buffer. For EcoRI roadblock experiments, the immobilized elongation complexes were

incubated with 1nM EcoRI E111Q, a mutant of restriction enzyme EcoRI, for 10 min. After that, the complexes were incubated with wild-type EcoRI (10 units/ μ l, Beams Biotechnology) to remove the DNA template without EcoRI E111Q binding. EcoRI roadblock experiments were performed in the transcription buffer additionally containing 1nM EcoRI E111Q. The results were analyzed using IDL (7.0, ITT), Matlab (R2014b, The MathWorks), and Origin (8.5, OriginLab).

2.3. Results

2.3.1. Fluorescent detection of intrinsic termination

The stalled elongation complexes, which were immobilized on quartz slide, resumed elongation by injecting all four NTPs, while the fluorescence signals of Cy3-RNA and Cy5-DNA in individual complexes were monitored. Using L+15 as a DNA template, both readthrough or termination at the termination site (TS) were observed after injecting NTPs at 0 s. In readthrough events, RNAP continues transcription without dissociation at TS, eventually reaching the downstream end of DNA where Cy5 is located. As a result, the Cy3-RNA signal does not vanish, and the PIFE of Cy5 occurs by RNAP with red laser excitation. In addition, FRET from Cy3 to Cy5 is also observed with green laser excitation, indicating the 5'-end of RNA approaches to the downstream end of DNA (Figure 2.4).

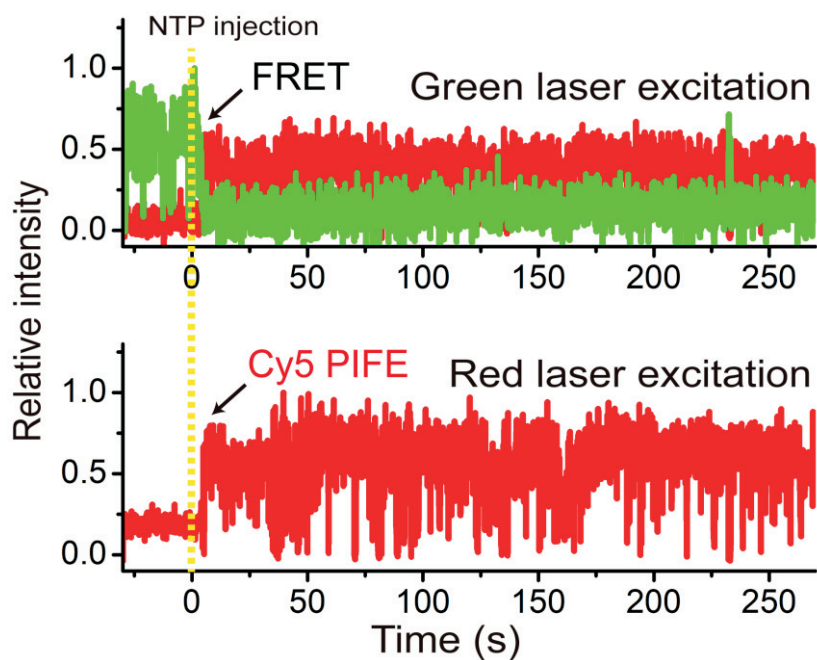


Figure 2.4. Representative fluorescence traces of the readthrough event. Cy3 signal and Cy5 signal were represented as a green line and a red line, respectively. RNAP arrives at the downstream end of DNA without releasing RNA. Both protein-induced fluorescence enhancement (PIFE) of Cy5 and fluorescence resonance energy transfer (FRET) from Cy3 to Cy5 were monitored. The yellow vertical line indicates NTP injection timing.

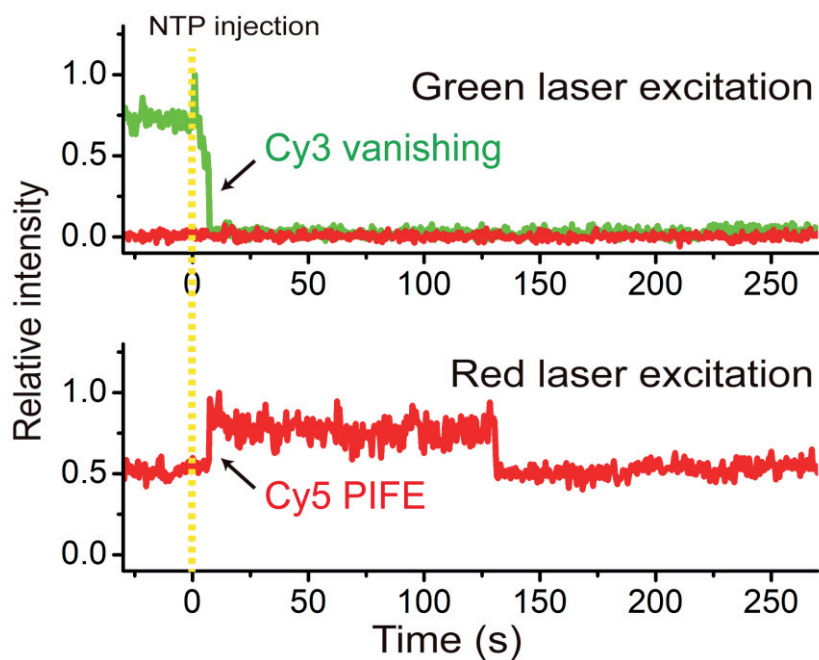


Figure 2.5. Representative fluorescence traces of the termination event with RNAP retention. Cy3 signal and Cy5 signal were represented as a green line and a red line, respectively. Cy3-RNA signal vanished right after NTP injection (yellow line), indicating RNA dissociation from DNA and RNAP. After that, 91% of RNA-free RNAP occurred protein-induced fluorescence enhancement (PIEF) of Cy5.

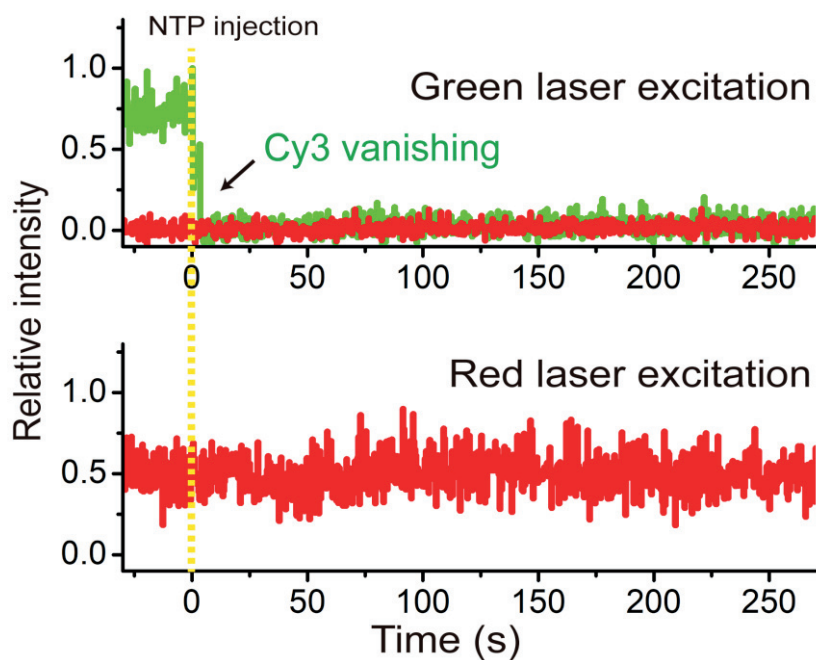


Figure 2.6. Representative fluorescence traces of the termination event without RNAP retention. Cy3 signal and Cy5 signal were represented as a green line and a red line, respectively. RNA dissociated from DNA and RNAP, showing Cy3-RNA signal is vanishing right after NTP injection (yellow line). 9% of RNA-free RNAP dissociated from DNA without protein-induced fluorescence enhancement (PIFE) of Cy5.

In termination events, RNAP recognizes the terminator and ends transcription at TS. Therefore, the Cy3-RNA signal is vanishing after NTP injection (Figure 2.5 and 2.6). From termination events, the time difference between NTP injection and Cy3-RNA vanishing was 6.1 ± 0.9 s (Figure 2.7), defined as the termination time. The termination efficiency, which is the frequency of the termination events, was 33 ± 4 % (Table 2.3). To confirm that an intrinsic terminator caused this result, we used inosine triphosphate (ITP) instead of guanosine triphosphate (GTP). Unlike guanosine, inosine forms only two hydrogen bonds with cytosine; therefore, a weaker hairpin terminator is formed, preventing termination. When replacing GTP with ITP, the termination efficiency was reduced to 12 ± 2 %, although the termination time did not affect (Table 2.3). Furthermore, using the terminator-lacking DNA template L+15M, the Cy3-RNA signal vanishing was not observed (Table 2.3). Based on these results, the Cy3 vanishing indicated RNA release from the elongation complex caused by the intrinsic terminator.

Most of the termination complexes showed Cy5 PIFE with red laser excitation after Cy3 vanishing (Figure 2.5), while the other termination complexes showed only Cy3 vanishing without Cy5 PIFE (Figure 2.6). Like Cy5 PIFE observed in readthrough events, Cy5 PIFE observed after Cy3 vanishing was also caused by RNAP, suggesting that RNAP keep contacting DNA after intrinsic terminator, as its possibility has been predicted before (Bellecourt et al., 2019). The PIFE occurrence, defined as the rate of Cy5 PIFE after RNA release, was 91 ± 5 % using L+15 and was not affected by

Chapter 2

the weaker hairpin structure caused by ITP (Table 2.3). Using T257/L+15, which has a 257-bp transcription unit, PIFE occurrence is similar (Table 2.3), indicating that RNAP retention does not occur only in promoter-proximal termination. After termination, the RNAP retention is also observed with other intrinsic terminators, such as *E. coli* his attenuator and phage ϕ 82 t500 terminator (Table 2.3), suggesting that post-terminational retention of RNAP is general.

As previously suggested (Santangelo et al., 2011), the termination efficiency was raised with NusA (Table 2.4). However, the PIFE occurrence and the termination time were little affected by NusA (Table 2.4). Also, NusG did not affect the termination time and efficiency and the PIFE occurrence. When both NusA and NusG are present, all properties are similar to those of NusA alone (Table 2.4).

Table 2.3. Termination properties of various DNA templates.

DNA template	Termination time ^a (s)	Termination efficiency (%)	PIFE occurrence (%)
L+15	6.1 ± 0.9	33 ± 4	91 ± 5
L+15 + ITP	8.6 ± 2.5	12 ± 2	83 ± 3
L+15M	Not available	Not available	Not available
T257/L+15	25.0 ± 6.6	36 ± 4	86 ± 2
his+62	3.0 ± 0.5	84 ^b	87 ^b
t500+62	3.5 ± 0.6	68 ^b	70 ^b

^aThese results was obtained from a single exponential fitting. ^bThese experiments were performed only one; therefore, the standard deviation was not obtained.

Table 2.4. Termination properties with various transcription factors.

Transcription factor	Termination time ^a (s)	Termination efficiency (%)	PIFE occurrence (%)
NusA	7.9 ± 0.2	58 ± 10	86 ± 6
NusG	4.1 ± 0.2	35 ± 3	90 ± 11
NusA + NusG	7.0 ± 0.8	58 ± 4	87 ± 3

^aThese results was obtained from a single exponential fitting.

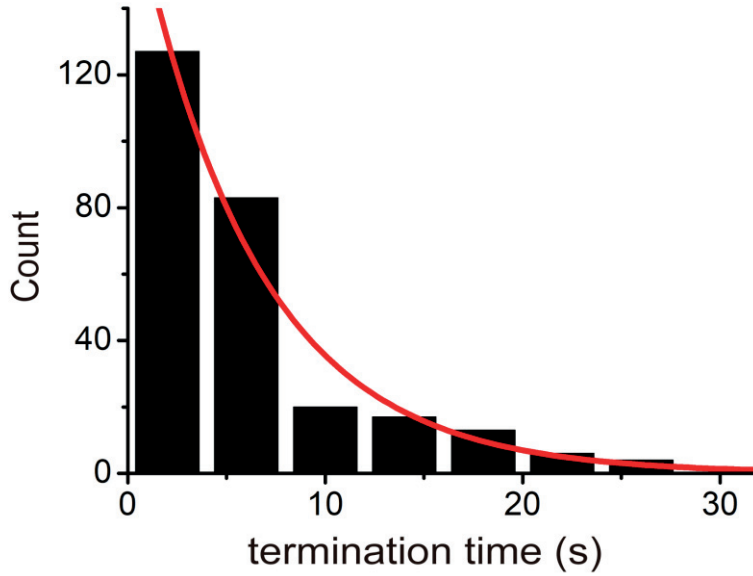


Figure 2.7. Termination time of L+15. We measured the time difference between Cy3-RNA signal vanishing time and NTP injection using L+15. The distribution was fitted to a single-exponential decay function (red line), and the fitted decay time is 6.9 ± 0.9 s. The number of data for measurement of termination time is 270.

2.3.2. One-dimensional diffusion of RNA free RNAP on DNA

Because TS and Cy5 are too close, it was unknown whether RNAP remaining on DNA after termination moves on DNA or not from the result using L+15. In order to examine that RNAP can reach farther Cy5-end of DNA, we used the DNA template L+112, of which the distance between TS and Cy5-end is 112-bp. In the experiment using L+112, the termination efficiency and the termination time were not changed from the experiment with L+15. The PIFE occurrence was also similar (Table 2.5), suggesting that RNAP escape TS after RNA release and moves on DNA.

However, it was still unknown whether RNAP reach the Cy5-end of DNA through three-dimensional (3D) diffusion after dissociation from DNA or through one-dimensional (1D) diffusion without dissociation from DNA. When the protein binds to L+112 between TS and Cy5, RNAP cannot reach Cy5 through 1D diffusion. The mutant of EcoRI, which replaced 111th glutamic acid (E) with glutamine (Q), retains the ability to bind to a specific DNA sequence, but cannot digest DNA. Therefore, EcoRI E111Q binding to DNA blocks the procession of *E. coli* RNAP (Epshtein et al., 2003). Similarly, this binding would obstruct 1D diffusion downward from TS but does not affect 3D diffusion. In experiments using L+112 with EcoRI E111Q binding, this roadblock reduced PIFE occurrence from $89 \pm 2 \%$ to $25 \pm 7 \%$ (Figure 2.8), whereas the termination time was not affected. Also, because RNAP cannot pass the EcoRI E111Q, readthrough events were not observed. Therefore, termination efficiency was not measured (Table 2.5). PIFE

occurrence decreased by about 72 %, and this rate was similar to the rate of EcoRI E111Q binding to EcoRI recognition site (71 ± 5 %, See Appendix A). When EcoRI E111Q bind to DNA, almost all RNAPs fail to reach the Cy5-end of DNA, which supports that RNAPs reach the Cy5-end of DNA mostly through 1D diffusion, much more often than 3D diffusion.

On the other hand, in experiments using L+112R, the PIFE occurrence and the termination time were similar to those of experiments using L+112 (Figure 2.8 and Table 2.5). These results mean that 1D diffusion of RNAP after termination is bi-directional. Also, because the EcoRI recognition sequence is located between TS and biotin, the PIFE occurrence did not reduce with the EcoRI roadblock (Figure 2.8).

Table 2.5. Termination properties with or without EcoRI roadblock.

DNA template	Termination time (s)	Termination efficiency (%)	PIFE occurrence (%)
L+112	5.9 ± 0.8	41 ± 5	89 ± 2
L+112 with EcoRI E111Q	4.9 ± 0.7	Not available	25 ± 7
L+112R	7.1 ± 1.3	Not available	81 ± 5
L+112R with EcoRI E111Q	6.2 ± 1.6	Not available	84 ± 4

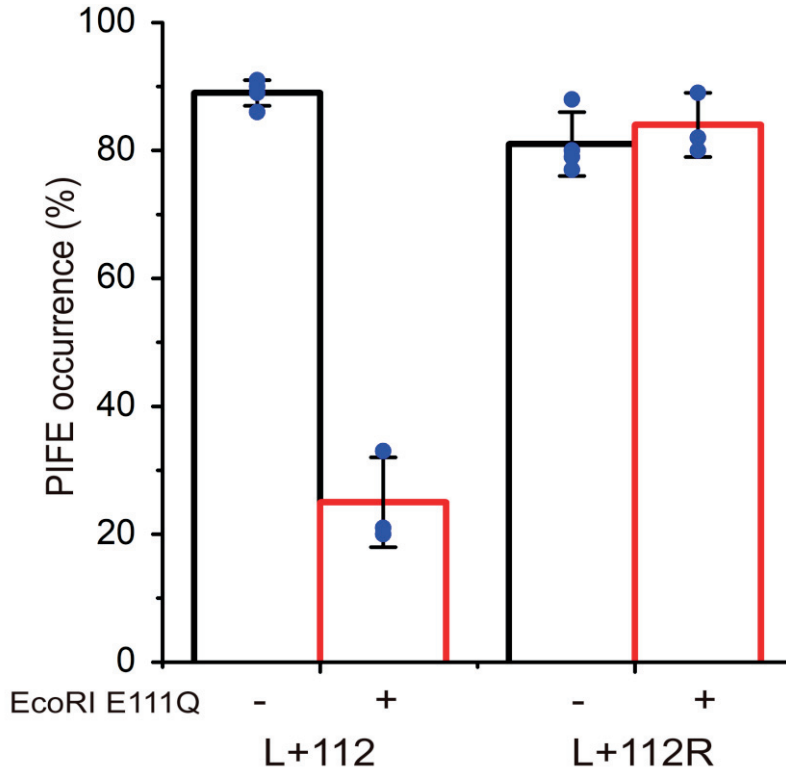


Figure 2.8. Cy5 PIFE occurrence for L+112 and L+112R with or without EcoRI E111Q. EcoRI E111Q roadblock reduced PIFE occurrence on L+112 but not L+112R. Blue dots indicate each result from an independent experiment, and error bars represent the standard deviation of it.

2.4. Conclusion

In this study, we discovered that RNA transcript is released from the elongation complex and RNAP maintains contact with DNA in *E. coli* intrinsic termination. It was also confirmed that RNAP remaining on DNA moves in both upstream and downstream directions through 1D diffusion. The

1D diffusion of the retention RNAP is similar to previously reported translation termination.

When the ribosomes encounter the stop codon in bacterial translation termination, the synthesized peptide chains dissociate first, and the 70S ribosomes remain binding to mRNA. These ribosomes are dissociated into the 30S and 50S ribosome subunits aided by the binding of ribosome recycling factor (RRF) and elongation factor G (EF-G) (Sternberg et al., 2009). Besides, these 70S ribosomes initiate translation at the start codon again through the 70S-scanning initiation pathway without dissociation from the mRNA (Yamamoto et al., 2016).

Like the ribosome, RNAP also has a stage that dissociates from DNA after bacterial intrinsic termination, which is here termed as recycling, and RNAP remaining on DNA after intrinsic termination is defined as recycling RNAP. Furthermore, RNAP, like the ribosome, is expected to be able to start transcription again at promoter without dissociation from the DNA.

References

- Bellecourt, M. J., Ray-Soni, A., Harwig, A., Mooney, R. A., & Landick, R. (2019). RNA Polymerase Clamp Movement Aids Dissociation from DNA but Is Not Required for RNA Release at Intrinsic Terminators. *Journal of molecular biology*, 431(4), 696–713.
- Czyz, A., Mooney, R. A., Iaconi, A., & Landick, R. (2014). Mycobacterial RNA polymerase requires a U-tract at intrinsic terminators and is aided by NusG at suboptimal terminators. *mBio*, 5(2), e00931.
- Epshtein, V., Toulmé, F., Rahmouni, A. R., Borukhov, S., & Nudler, E. (2003). Transcription through the roadblocks: the role of RNA polymerase cooperation. *The EMBO journal*, 22(18), 4719–4727.
- Farnham, P. J., & Platt, T. (1981). Rho-independent termination: dyad symmetry in DNA causes RNA polymerase to pause during transcription in vitro. *Nucleic acids research*, 9(3), 563–577.
- Gusarov, I., & Nudler, E. (1999). The mechanism of intrinsic transcription termination. *Molecular cell*, 3(4), 495–504.
- Mondal, S., Yakhnin, A. V., Sebastian, A., Albert, I., & Babitzke, P. (2016). NusA-dependent transcription termination prevents misregulation of global gene expression. *Nature microbiology*, 1, 15007.
- Ray-Soni, A., Mooney, R. A., & Landick, R. (2017). Trigger loop dynamics can explain stimulation of intrinsic termination by bacterial RNA polymerase without terminator hairpin contact. *Proceedings of the National Academy of Sciences of the United States of America*, 114(44), E9233–E9242.
- Roy, R., Hohng, S., & Ha, T. (2008). A practical guide to single-molecule FRET. *Nature methods*, 5(6), 507–516.
- Santangelo, T. J., & Artsimovitch, I. (2011). Termination and antitermination: RNA polymerase runs a stop sign. *Nature reviews. Microbiology*, 9(5), 319–329.

Sofia, S. J., Premnath, V., V., & Merrill, E. W. (1998). Poly(ethylene oxide) Grafted to Silicon Surfaces: Grafting Density and Protein Adsorption. *Macromolecules*, 31(15), 5059–5070.

Sternberg, S. H., Fei, J., Prywes, N., McGrath, K. A., & Gonzalez, R. L., Jr (2009). Translation factors direct intrinsic ribosome dynamics during translation termination and ribosome recycling. *Nature structural & molecular biology*, 16(8), 861–868.

Wilson, K. S., & von Hippel, P. H. (1995). Transcription termination at intrinsic terminators: the role of the RNA hairpin. *Proceedings of the National Academy of Sciences of the United States of America*, 92(19), 8793–8797.

Yamamoto, H., Wittek, D., Gupta, R., Qin, B., Ueda, T., Krause, R., Yamamoto, K., Albrecht, R., Pech, M., & Nierhaus, K. H. (2016). 70S-scanning initiation is a novel and frequent initiation mode of ribosomal translation in bacteria. *Proceedings of the National Academy of Sciences of the United States of America*, 113(9), E1180–E1189.

Chapter 3

The recycling RNAP one-dimensional diffuses on DNA using a hopping mechanism

3.1. Introduction

In the cell, various proteins bind to the specific sequence of DNA, and these need to find their target sequence in a short time with high affinity. Four diffusion-based mechanisms that contribute to finding specific sequences were proposed (Berg et al., 1981; Von Hippel et al., 1989). The first mechanism is three-dimensional (3D) diffusion, where protein finds its targets through three-dimensional collisions. The second one is one-dimensional (1D) hopping, where protein repeatedly dissociates from DNA, three-dimensional diffuses in a minimal area, and binds back to DNA at nearby segments. The third one is one-dimensional (1D) sliding, where protein moves along the backbone of DNA without dissociation from DNA. The last mechanism is an intersegmental transfer, where protein moves from one site to another via a looped intermediate. The target binding of *E. coli* lac repressor is faster than

the expected binding rate through 3D diffusion (Riggs et al., 1970). It suggests that 1D facilitated diffusion, including hopping, sliding, and intersegmental transfer, accelerates target association rates (Berg et al., 1976; Berg et al., 1981; Von Hippel et al., 1989; Halford et al., 2004).

Like DNA-binding protein, the recycling RNAP is also expected to diffuse on DNA via hopping or sliding. For the hopping mechanism, counter-ion re-condenses on DNA to compensate the microscopic dissociation of the hopping protein, whereas counter-ion density does not change for the sliding mechanism (Manning 1977). Therefore, since high salt concentrations speed up the ionic condensation, the 1D diffusion coefficients of hopping proteins increase with rising salt concentration (Hedglin et al., 2010; Gorman et al., 2010; Brown et al., 2016; Cheon et al., 2019), but those of sliding proteins are not affected by salt concentrations (Winter et al., 1981; Blainey et al., 2006; Gorman et al., 2007; Lin et al., 2014).

In this chapter, we introduced the method to measure the 1D diffusion coefficient under various salt conditions using a single-molecule fluorescence experiment. As a result, the 1D diffusion coefficient of the recycling RNAP increases with rising salt concentrations, confirming that RNAP hops during the observed post-termination diffusion on DNA.

3.2. Materials and Methods

3.2.1. Protein preparation

E. coli RNAP with a C-terminal SNAP-tagged RpoZ, known as omega subunit, was expressed in *E. coli* BL21 (DE3) cell and purified previously described (Wang et al., 2016). In order to construct Cy5-labeled O6-benzylguanine, BG-NH₂ (New England Biolab) was incubated with Cy5-NHS ester (12.5 µg/µl, GE Healthcare) for 90 min at room temperature. This Cy5-labeled O6-benzylguanine is incubated with SNAP-tagged RNAP (3.9 mg/ml) for 2 hours at room temperature in a storage buffer (20 mM Tris-HCl, pH 7.5, 50 mM NaCl, 5mM EDTA). Cy5-NHS ester, BG-NH₂, and Cy5-BG, which are not bound to RNAP, were removed using Amino ultra-centrifugal filter (Merck).

3.2.2. DNA template preparation

To measure the 1D diffusion coefficient of the recycling RNAP after intrinsic termination, we prepared six DNA templates with various lengths of the downstream part (Figure 3.1). These DNAs are denoted by L+# with TS-Cy5 basepair number. All DNAs were prepared by PCR using AccuPower ProFi Taq PCR premix from Bioneer, Korea, and all PCR products were purified using the Cleanup kit Expin PCR SV mini from GeneAll, Korea. All amplification templates and primers labeled with biotin or Cy5 were purchased from Integrated DNA Technologies, USA (Table 3.1 and Table 3.2).

Chapter 3

For DNA template L+15, DNA_template_0, B_primerF and 5_primerR_L+15 were used. For L+62, DNA_template_0, B_primerF and 5_primerR_L+62 were used. For L+112, DNA_Template_1, B_primerF and 5_primerR_L+112 were used.

Since L+212 and L+312 are too long to be synthesized, an amplification template was produced through ligation. First, L+112 and additional_template_1 were annealed with DNA_splint_1 by cooling from 90 °C to 30 °C for 2 hours in an annealing buffer (10 mM Tris-HCl, pH 8.0, 50 mM NaCl). Afterward, it was ligated by incubating with T4 DNA ligase 2 (NEB) at 16 °C for 16 hours and used for application reactions. For L+212, B_primerF and 5_primerR_L+212 were used. For L+312, B_primerF and 5_primerR_L+312 were used. For L+512 construction, L+312 and additional_template_2 were annealed with DNA_splint_2, and the same ligation reaction was repeated. L+512 was prepared using B_primerF and 5_primerR_L+512.

In order to measure the retention time of the recycling RNAP, DNA with enough long downstream part was needed. For this, we firstly prepared a long_tail DNA with HindIII recognition site using λ -DNA (NEB), primerF_lambda, and primerR_lambda. Next, long_tail DNA and L+512 were digested with HindIII (NEB) for 1 hour at 37 °C in the Cutsmart™ buffer (NEB). Afterward, HindIII was deactivated for 20 min at 80 °C. The cleaved L+512 and long_tail DNA were annealed and ligated by the same protocol as above. The ligated product, B_primerF and, primerF_lambda were used to

construct long DNA named as $L+\lambda$.

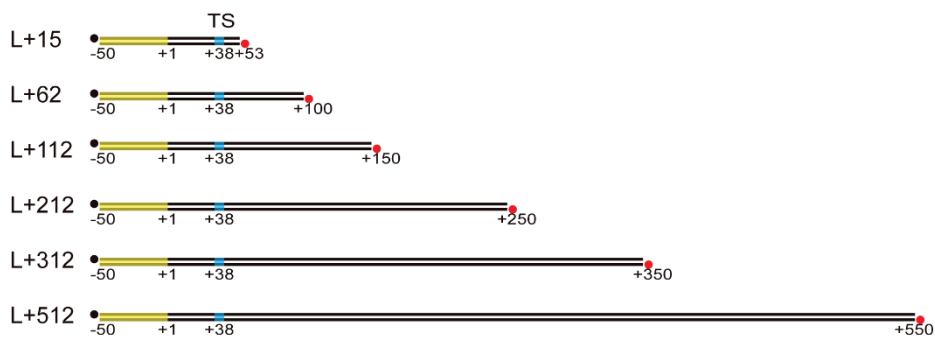


Figure 3.1. Design of various DNA templates. All DNA templates include T7A1 promoter (yellow box), termination site (TS, cyan oval), biotin (black dot), and Cy5 (red dot). Each DNA template has a different distance between TS and Cy5 and is named according to the distance.

Chapter 3

Table 3.1. Sequences of amplification templates used for PCR.

Oligo name	Sequence (5' → 3')
DNA_template_0	TATCA AAAAG AGTAT TGACT TAAAG TCTAA CCTAT AGGAT ACTTA CAGCC ATCGA ACAGG CCTGC TGGTA ATCGC AGGCC TTTT ATTG GGGGA GAGGG AAGTC ATGAA AAAAC TAACC TTTGA AATTC GATCT CCAGG ATCCA CCACC
DNA_template_1	TATCA AAAAG AGTAT TGACT TAAAG TCTAA CCTAT AGGAT ACTTA CAGCC ATCCC AAAGC CCGCC GAAAG GCGGG CTTTT CTGTT TCTGG GCGGT GAAGT CATGA AAAAA CTAAC CTTTG AAATT CGATC TCCAG GATCC ACCAC C
additional_part_1	pAATTC TTACA ATTTA GACCC TAATA TCACA TCAGA CACTA ATTGC CTCTG CCAAA ATTCT GTCCA CAAGC GTTTT AGTTC GCCCC AGTAA AGTTG TCAAT AACGA CCACC AAATC CGCAT GTTAC GGGAC TTCTT ATTAA TTCTT TTTTC GTGGG GAGCA GCGGA TCTTA ATGGA TGGCG CCAGG TGGTA TGGAA GC
additional_part_2	pGGGCT GAAAG TAGCG CCGGG TAAGG TACGC GCCTG GTATG GCAGG ACTAT GAAGC CAATA CAAAG GCTAC ATCCT CACTC GGGTG GACGG AAACG CAGAA TTATG GTTAC TTTT GGATA CGTGA AACAT GTCCC ATGGT AGCCC AAAGA CTTGG GAGTC TATCA CCCCT AGGAC ACACA AGACA CCACA AGCTT AGACC
DNA_splint	TGTGA TATTA GGGTC TAAAT TGTA GAATT GCGAG ATTAC CATTA AGTGA ATTCG AAAAA
DNA_splint_2	GCGTA CCTTA CCCGG CGCTA CTTTC AGCCC GCTTC CATAC CACCT GGC GC CATCC ATTAA

Phosphate is indicated as p.

Table 2.2. Sequences of primers

Oligo name	Sequence (5' → 3')
B_primerF	Biotin-TATCA AAAAG AGTAT TGA CT TAAAG TC
5_primerR_L+15	Cy5-CTTCC CTCTC CCCCC AATAA AAAG
5_primerR_L+62	Cy5-GGTGG TGGAT CCTGG AGATC G
5_primerR_L+112	Cy5-GCGAG ATTAC CATTA AGTGA A
5_primerR_L+212	Cy5-GACAA CTTTA CTGGG GCGAA CTAAA AC
5_primerR_L+312	Cy5-GCTTC CATA CACCT GGCGC CATCC AT
5_primerR_L+512	Cy5-GGTCT AAGCT TGTGG TGTCT TGTGT GT
primerF_lambda	GTTTT CTGGG TTGGT
primerR_lambda	GGCGG GTTTT GTTTT

3.2.3. Single-molecule transcription termination experiments

To construct the transcription complex with fluorescent RNAP, 50 nM of L+ λ was incubated with Cy5-labeled RNAP core enzyme (40 nM), sigma factor (1 μ M), ATP, CTP, GTP (each 20 μ M, GE Healthcare), and Cy3-labeled ApU (250 μ M, TriLink) for 30 min at 37 °C in an initiation buffer (10 mM Tris-HCl, pH8.0, 20 mM NaCl, 20 mM MgCl₂ and 1 mM dithiothreitol). For the experiments measuring the 1D diffusion coefficient of the recycling RNAP, 50 nM of DNA template was incubated with E. coli RNA holoenzyme (20 nM, NEB) ATP, CTP, GTP (each 20 μ M, GE Healthcare), and Cy3-labeled ApU (250 μ M, TriLink) for 30 min at 37 °C in a transcription buffer.

Before immobilizing the stalled elongation complex on the quartz slides, quartz slides were coated with polyethylene glycol (PEG) to suppress non-specific binding (Sofia et al., 1998). At first, quartz slides were cleaned using

piranha solution to remove organic residues, incubated with (3-aminopropyl) trimethoxysilane (United Chemical Technologies), and coated with polymers by incubation in a 1:40 mixture of biotin-PEG-5000 and m-PEG-5000 (Laysan Bio.) (Roy et al., 2008). After that, the slides were treated with 0.2 mg/ml streptavidin (Invitrogen) for 5 min. The stalled elongation complexes were immobilized on polymer-coated quartz slides through biotin-streptavidin conjugation. Unbound RNAPs and un-immobilized elongation complexes were removed by washing with an initiation buffer and an imaging buffer extensively.

In order to image the single-molecule fluorescence from Cy3 and Cy5-labeled elongation complex, a home-made TIRF microscopy was used. A 532-nm green laser (EXLSR-532-50-CDRH, Spectra-physic) and a 640-nm red laser (EXLSR-640C-60-CDRH, Spectra-physic) were used for Cy3 excitation and Cy5 excitation, respectively. A single-molecule image was obtained by EM-CCD (Ixon DV897, Andor Technology) controlled by a customized C# program. Experiments to measure the retention time were performed with 0.1 s exposure time, and experiments to measure the 1D diffusion coefficient of the recycling RNAP were performed with 0.05 s exposure time in an ALEX mode. Because the experiments were performed in an ALEX mode, the actual time resolution was 0.2s to measure the retention time and 0.1 s to measure the 1D diffusion coefficient. All experiments were performed at 37 °C.

All experiments were performed in an elongation buffer (10 mM Tris-HCl, pH 8.0, various concentration of NaCl, 2 mM MgCl₂, 1 mM dithiothreitol, 5

mM 3,4-protocatechuic acid and 100 nM protocatechuate-3,4-dioxygenase) and the concentration of NaCl was changed from 20 mM to 200 mM. The results were analyzed using IDL (7.0, ITT), Matlab (R2014b, The MathWorks), and Origin (8.5, OriginLab).

3.3. Results

3.3.1. The measurement of the recycling RNAP retention time

Since it is known that RNAP preferentially binds blunt ends of DNA, the time that post-terminational RNAP retains on DNA, defined as the RNAP retention time, on short DNA is longer than those on long DNA where diffusion of RNAP takes too long to reach an end of DNA. The RNAP retention times were measured with various NaCl concentrations using L+ λ with 2,336-bp downstream part from TS and Cy5-labeled RNAP. Because tR2 termination occurs at 6.1 ± 0.9 s after NTP injection, Cy3 signal vanishing within 15 s after NTP injection was regarded as RNA release by intrinsic termination. Regardless of NaCl concentration, most of the Cy5 signals remained after Cy3 vanishing (Figure 3.2), and a few Cy5 signals vanished at the same time as Cy3. Since the percentage of Cy5 signal remaining after Cy3 vanishing refers to the probability that RNAP keeps the binding on DNA after RNA release during intrinsic termination, we defined this value as the RNAP retention probability. This value does not change significantly at NaCl concentration lower than 150mM but start to decrease at higher concentration.

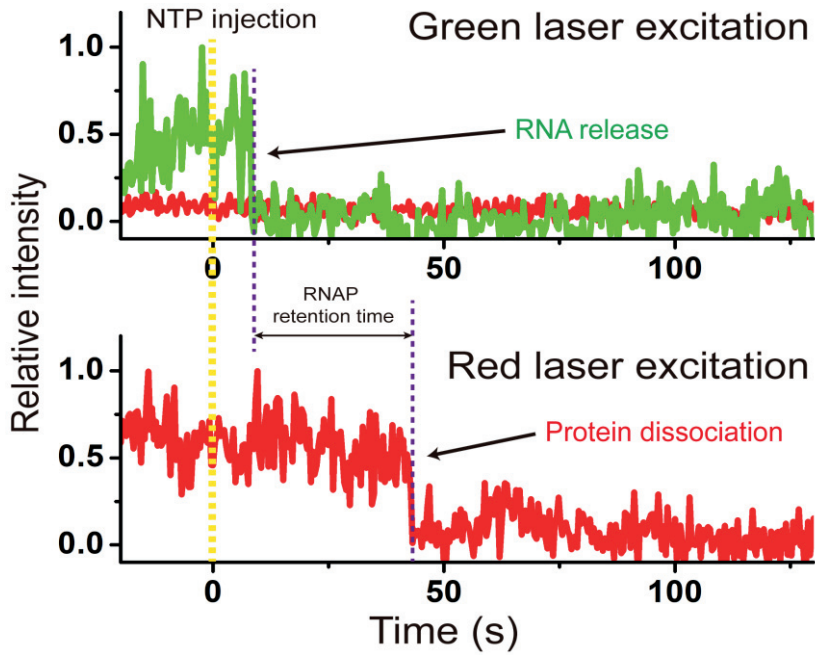


Figure 3.2. Representative fluorescence trace for the experiment using Cy5-labeled RNAP. Cy3 signal and Cy5 signal were represented as a green line and a red line, respectively. Cy3-RNA vanishing within 15 s after NTP injection (yellow line) was regarded as RNA release. In most RNA release molecules, the Cy5-RNAP signal remains after RNA release. We defined the RNAP retention time as the time difference between Cy3 vanishing and Cy5 vanishing.

Among the result where Cy5 signal remains after Cy3 vanishing, the time difference between Cy3 signal vanishing and Cy5 signal vanishing was collected, defined as the RNAP retention time (Figure 3.2). Since the dissociation mechanism of RNAP is a single-step process, the RNAP retention times were defined as the decay constant obtained from single-exponential fitting (Figure 3.3). As expected from other DNA binding proteins, the RNAP retention time decreases with rising NaCl concentration (Figure 3.4 and Table 3.3). The photobleaching time of Cy5 with the same exposure time (957s, Figure 3.5) is long enough compared to the RNAP retention time; therefore, it does not affect the RNAP retention time measurement.

Table 3.3. The RNAP retention times and probabilities at varying NaCl concentration.

[NaCl] (mM)	RNAP retention time (s)	RNAP retention probability (%)
20	124 ± 30	84
50	62 ± 19	86
100	48 ± 6	87
150	34 ± 6	78
200	32 ± 15	74

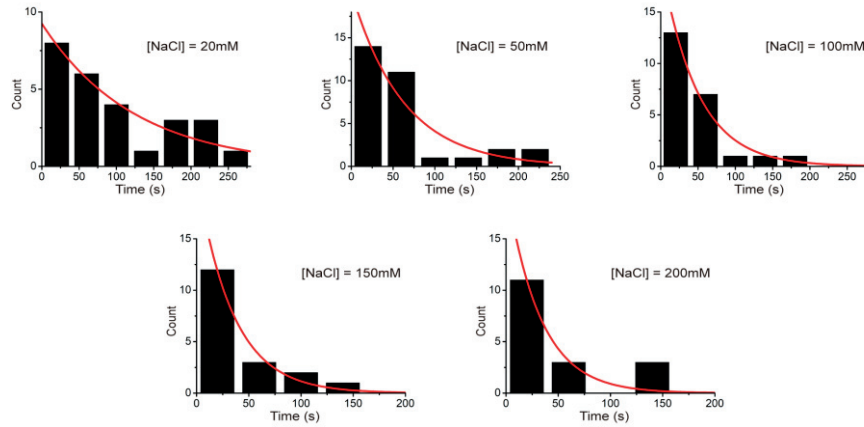


Figure 3.3. The RNAP retention times with various NaCl concentrations. We measured the RNAP retention time with 20 mM ($n = 26$), 50 mM ($n = 31$), 100 mM ($n = 26$), 150 mM ($n = 18$) and 200 mM ($n = 17$) NaCl. These distributions were fitted to a single exponential function (Red line).

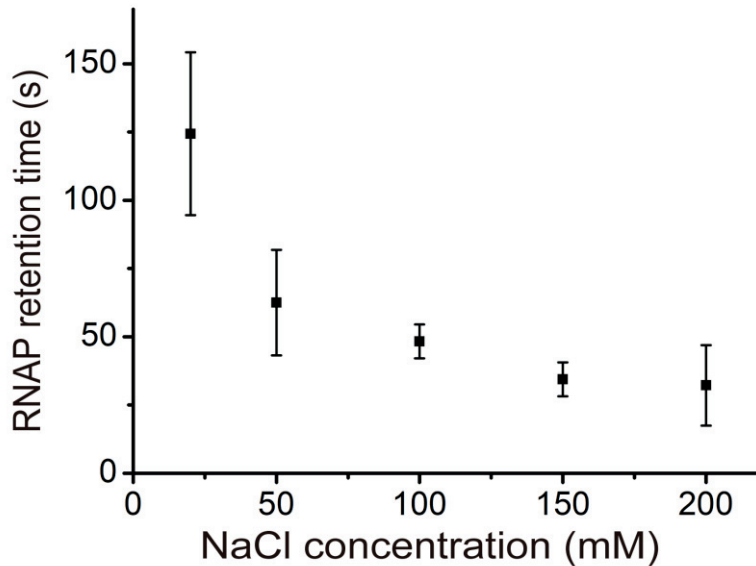


Figure 3.4. The RNAP retention times at varying NaCl concentration. Error bars mean standard error. All data are summarized in Table 3.4.

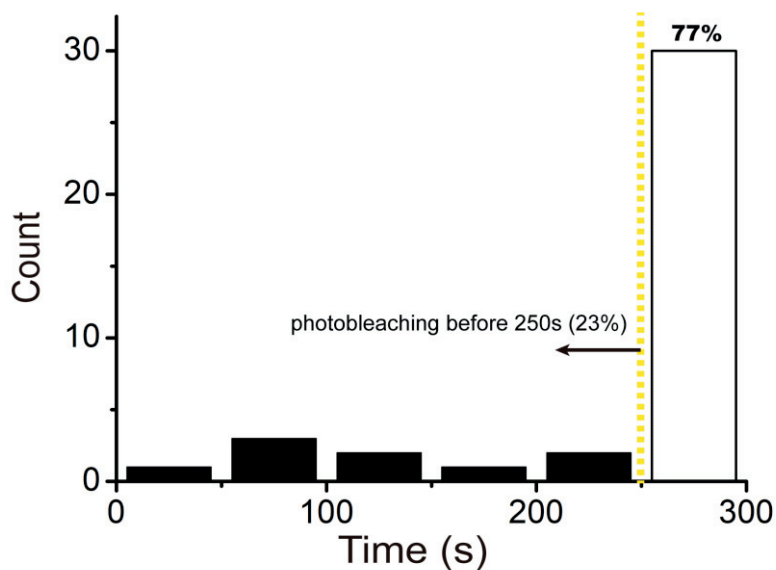


Figure 3.5. Distribution of Cy5 photobleaching time. We measured the photobleaching time with 0.2 s exposure time. From 39 molecules observed, 77 % of molecules (white bar) survived for 250 s, whereas 23 % of molecules (black bar) vanished before 250 s. By assuming that photobleaching is a single-step process, survival probability is expressed as $e^{-t/\tau}$, where τ is the photobleaching time. From these data, the photobleaching time is estimated as 957 ($= -250/\ln 0.77$) s.

3.3.2. 1D diffusion coefficient of the recycling RNAP

1D diffusion of the recycling RNAP can be modeled as a 1-dimensional random walk (See Appendix B). According to this model, Cy5 PIFE, which indicates that RNAP arrives at Cy5, occurs less frequently and starts later as the distance between TS and Cy5 increases. To confirm this model, fluorescent signals were measured with 150 mM NaCl using six different DNA templates having the distance between TS and Cy5 varies from 15-bp to 512-bp.

The three types of fluorescence time trace from the experiments using L+212 were observed. Readthrough cases in which Cy5 PIFE occurs without Cy3-RNA vanishing were observed in 61 % of the total results (150/246). On the other hand, termination cases with RNAP remaining on DNA, in which Cy5 PIFE occurs after Cy3 vanishing, were observed in 32 % of the total results (Figure 3.6, 79/246). The PIFE delay time (t_{del}) was defined as the time difference between Cy3 vanishing and Cy5 PIFE start (Figure 3.6). The 7 % of the results showed termination cases with RNAP dissociation, in which Cy3 vanishes without Cy5 PIFE (17/246). Even with using other DNA templates, three types were also observed.

We repeated the same experiments with various NaCl concentrations to measure salt concentration dependence. All three types were observed regardless of salt concentration and DNA templates, and termination properties were shown in Table 3.4-8.

Table 3.4. Termination properties for various DNA templates with 20 mM NaCl.

DNA template	Termination efficiency (%)	PIFE delay time (s)	PIFE occurrence (%)
L+15	55 ± 14	3.1 ± 0.9	78 ± 9
L+62	48 ± 6	6.6 ± 4.4	70 ± 1
L+112	64 ± 14	7.5 ± 7.9	68 ± 2
L+212	45 ± 1	29.7 ± 11.3	55 ± 18
L+312	46 ± 8	39.0 ± 28.4	46 ± 7
L+512	46 ± 13	29.5 ± 19.2	38 ± 8

Table 3.5. Termination properties for various DNA templates with 50 mM NaCl.

DNA template	Termination efficiency (%)	PIFE delay time (s)	PIFE occurrence (%)
L+15	55 ± 14	3.1 ± 0.9	78 ± 9
L+62	48 ± 6	6.6 ± 4.4	70 ± 1
L+112	64 ± 14	7.5 ± 7.9	68 ± 2
L+212	45 ± 1	29.7 ± 11.3	55 ± 18
L+312	46 ± 8	39.0 ± 28.4	46 ± 7
L+512	46 ± 13	29.5 ± 19.2	38 ± 8

Table 3.6. Termination properties for various DNA templates with 100 mM NaCl.

DNA template	Termination efficiency (%)	PIFE delay time (s)	PIFE occurrence (%)
L+15	50 ± 12	2.5 ± 1.2	80 ± 10
L+62	49 ± 13	1.8 ± 1.2	75 ± 5
L+112	51 ± 14	12.4 ± 3.4	72 ± 10
L+212	66 ± 19	22.5 ± 9.0	63 ± 10
L+312	45 ± 8	22.8 ± 14.7	47 ± 19
L+512	42 ± 2	21.3 ± 6.4	30 ± 9

Table 3.7. Termination properties for various DNA templates with 150 mM NaCl.

DNA template	Termination efficiency (%)	PIFE delay time (s)	PIFE occurrence (%)
L+15	39 ± 4	0.7 ± 0.4	82 ± 7
L+62	41 ± 11	0.8 ± 0.5	72 ± 16
L+112	43 ± 5	1.3 ± 0.8	71 ± 8
L+212	46 ± 6	9.0 ± 5.6	58 ± 13
L+312	44 ± 6	15.8 ± 4.1	52 ± 10
L+512	46 ± 7	25.5 ± 6.2	44 ± 11

Table 3.8. Termination properties for various DNA templates with 200 mM NaCl.

DNA template	Termination efficiency (%)	PIFE delay time (s)	PIFE occurrence (%)
L+15	48 ± 1	0.6 ± 0.2	64 ± 6
L+62	55 ± 5	1.3 ± 0.5	63 ± 4
L+112	52 ± 9	2.2 ± 1.4	59 ± 10
L+212	49 ± 5	5.4 ± 0.6	54 ± 11
L+312	51 ± 4	8.8 ± 8.3	47 ± 10
L+512	51 ± 10	11.9 ± 5.3	37 ± 3

Cy5 PIFE occurrences (Figure 3.7), which is the frequency of Cy5 PIFE among termination cases, decreased, and the PIFE delay times (Figure 3.8) increased as the TS-Cy5 distance increased, as expected from the model. The 1D diffusion coefficient of the recycling RNAP can be obtained by fitting these data using equations derived from the model. Furthermore, the RNAP retention probability can also be obtained by fitting Cy5 PIFE occurrence data. The diffusion coefficients obtained from the PIFE occurrence and the PIFE delay time increase with increasing NaCl concentration (Figure 3.9 and Table 3.9), demonstrating that the recycling RNAP hops during the 1D diffusion on DNA. On the other hand, the RNAP retention probability does not change at low NaCl concentration and decreases when NaCl concentration is higher than 150mM (Table 3.9). These results are similar to those measured using Cy5-labeled RNAP (Table 3.3).

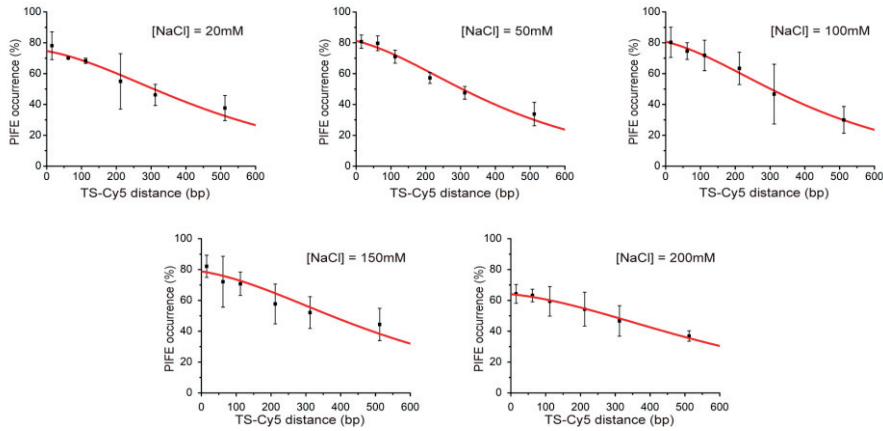


Figure 3.7. PIFE occurrences at varying NaCl concentrations. All data are mean \pm standard variation from three or more independent experiments. bp, basepair. These data were fitted (red line) to obtain the 1D diffusion coefficient and the RNAP retention probability.

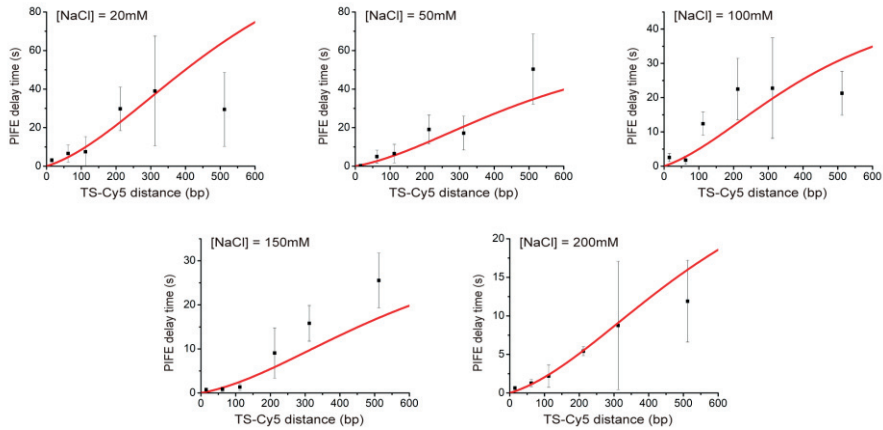


Figure 3.8. PIFE delay times at varying NaCl concentrations. All data are mean \pm standard variation from three or more independent experiments. bp, basepair. These data were fitted (red line) to obtain the 1D diffusion coefficient.

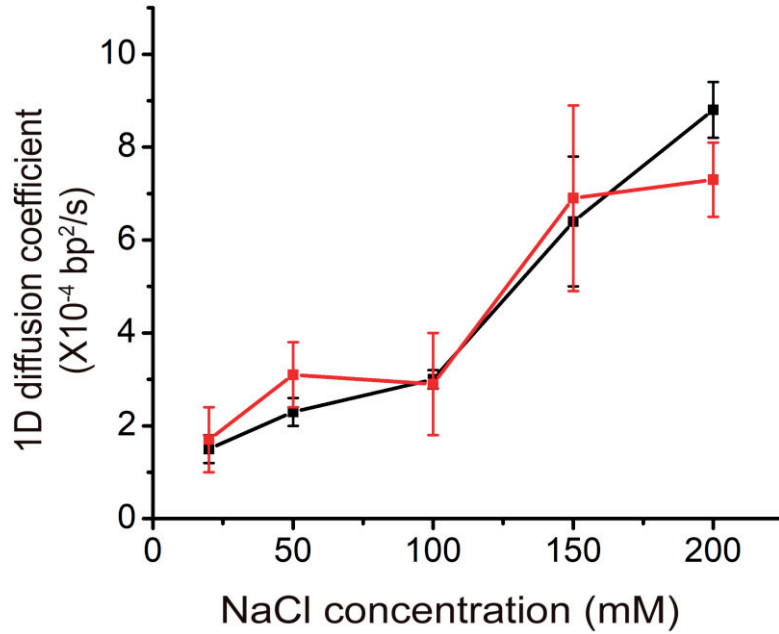


Figure 3.9. 1D diffusion coefficient of the recycling RNAP plotted against NaCl concentration. We obtained the 1D diffusion coefficient from PIFE occurrence (black line) and PIFE delay time (red line).

Table 3.9. The 1D diffusion coefficient and RNAP retention probability obtained from PIFE delay time^a and PIFE occurrence^b.

[NaCl] (mM)	Diffusion coefficient ^a (X10 ⁻⁴ bp ² /s)	Diffusion coefficient ^b (X10 ⁻⁴ bp ² /s)	RNAP retention probability ^b (%)
20	1.7 ± 0.7	1.5 ± 0.3	75 ± 3
50	3.1 ± 0.7	2.3 ± 0.3	81 ± 2
100	2.9 ± 1.1	3.0 ± 0.2	81 ± 1
150	6.9 ± 2.0	6.4 ± 1.4	79 ± 3
200	7.3 ± 0.8	8.8 ± 0.6	64 ± 1

3.4. Conclusion

In this study, we discovered that recycling RNAP after intrinsic termination diffuses on DNA using a hopping mechanism, based on the 1D diffusion coefficient of recycling RNAP. However, this finding does not exclude the sliding mechanism. It is possible that the recycling RNAP hops and slides for 1D diffusion, as previously reported for other DNA-binding proteins (Cuculis et al., 2015). In addition, it has been shown that some DNA-binding proteins can pass the roadblock through hopping (Hedglin et al., 2010; Brown et al., 2016; Cheon et al., 2019), whereas Cy5 PIFE occurrence by RNAP was reduced by EcoRI E111Q blocking. It can be explained that RNAP hops short distance. These short-range hopping or sliding is disadvantageous in overcoming the roadblock but helps search the target sequence such as a promoter. Therefore, the mechanism of the recycling RNAP is expected to optimize the search for promoters nearby the terminator.

References

- Berg, O. G., & Blomberg, C. (1976). Association kinetics with coupled diffusional flows. Special application to the lac repressor--operator system. *Biophysical chemistry*, 4(4), 367–381.
- Berg, O. G., Winter, R. B., & von Hippel, P. H. (1981). Diffusion-driven mechanisms of protein translocation on nucleic acids. 1. Models and theory. *Biochemistry*, 20(24), 6929–6948.
- Blainey, P. C., van Oijen, A. M., Banerjee, A., Verdine, G. L., & Xie, X. S. (2006). A base-excision DNA-repair protein finds intrahelical lesion bases by fast sliding in contact with DNA. *Proceedings of the National Academy of Sciences of the United States of America*, 103(15), 5752–5757.
- Brown, M. W., Kim, Y., Williams, G. M., Huck, J. D., Surtees, J. A., & Finkelstein, I. J. (2016). Dynamic DNA binding licenses a repair factor to bypass roadblocks in search of DNA lesions. *Nature communications*, 7, 10607.
- Cheon, N. Y., Kim, H. S., Yeo, J. E., Schärer, O. D., & Lee, J. Y. (2019). Single-molecule visualization reveals the damage search mechanism for the human NER protein XPC-RAD23B. *Nucleic acids research*, 47(16), 8337–8347.
- Cuculis, L., Abil, Z., Zhao, H., & Schroeder, C. M. (2015). Direct observation of TALE protein dynamics reveals a two-state search mechanism. *Nature communications*, 6, 7277.
- Gorman, J., Chowdhury, A., Surtees, J. A., Shimada, J., Reichman, D. R., Alani, E., & Greene, E. C. (2007). Dynamic basis for one-dimensional DNA scanning by the mismatch repair complex Msh2-Msh6. *Molecular cell*, 28(3), 359–370.
- Gorman, J., Plys, A. J., Visnapuu, M. L., Alani, E., & Greene, E. C. (2010). Visualizing one-dimensional diffusion of eukaryotic DNA repair factors along a chromatin lattice. *Nature structural & molecular biology*, 17(8), 932–938.
- Hedglin, M., & O'Brien, P. J. (2010). Hopping enables a DNA repair glycosylase to search both strands and bypass a bound protein. *ACS chemical biology*, 5(4), 427–

436.

Lin, J., Countryman, P., Buncher, N., Kaur, P., E, L., Zhang, Y., Gibson, G., You, C., Watkins, S. C., Piehler, J., Opresko, P. L., Kad, N. M., & Wang, H. (2014). TRF1 and TRF2 use different mechanisms to find telomeric DNA but share a novel mechanism to search for protein partners at telomeres. *Nucleic acids research*, 42(4), 2493–2504.

Manning G. S. (1977). Limiting laws and counterion condensation in polyelectrolyte solutions. IV. The approach to the limit and the extraordinary stability of the charge fraction. *Biophysical chemistry*, 7(2), 95–102.

Riggs, A. D., Bourgeois, S., & Cohn, M. (1970). The lac repressor-operator interaction. 3. Kinetic studies. *Journal of molecular biology*, 53(3), 401–417.

Roy, R., Hohng, S., & Ha, T. (2008). A practical guide to single-molecule FRET. *Nature methods*, 5(6), 507–516.

Sofia, S. J., Premnath, V., V, & Merrill, E. W. (1998). Poly(ethylene oxide) Grafted to Silicon Surfaces: Grafting Density and Protein Adsorption. *Macromolecules*, 31(15), 5059–5070.

von Hippel, P. H., & Berg, O. G. (1989). Facilitated target location in biological systems. *The Journal of biological chemistry*, 264(2), 675–678.

Wang, G., Hauver, J., Thomas, Z., Darst, S. A., & Pertsinidis, A. (2016). Single-Molecule Real-Time 3D Imaging of the Transcription Cycle by Modulation Interferometry. *Cell*, 167(7), 1839–1852.e21.

Winter, R. B., Berg, O. G., & von Hippel, P. H. (1981). Diffusion-driven mechanisms of protein translocation on nucleic acids. 3. The Escherichia coli lac repressor--operator interaction: kinetic measurements and conclusions. *Biochemistry*, 20(24), 6961–6977.

Chapter 4

The recycling RNAP can initiate transcription again without dissociation from DNA

4.1. Introduction

Some studies showed that RNAP searches promoter through 1D facilitated diffusion (Singer et al., 1987; Ricchetti et al., 1988; Kabata et al., 1993; Guthold et al., 1999; Harada et al., 1999), whereas other studies contended that promoter binding rate does not exceed the 3D-diffusion limit (Roe et al., 1984; Friedman et al., 2012). Furthermore, experiments using a single-molecule fluorescence imaging contended that 1D facilitated diffusion is excluded because the binding lifetime of RNAP on DNA is too short (Wang et al., 2013; Friedman et al., 2013). The recycling RNAP is also expected to bind to a promoter when it encounters a promoter, although there is debate whether RNAP searches promoter through 1D diffusion. It is defined as reinitiation events that the recycling RNAP initiates transcription at promoter again after intrinsic termination without dissociation from DNA.

It has been reported that DNA repair glycosylase, which moves on DNA through hopping, can search for a damaged nucleotide on both DNA strands (Hedglin et al., 2010). Similarly, since the recycling RNAP diffuses on DNA via hopping, it is assumed that RNAP can search sense and antisense DNA, and reinitiation events can occur at both strands. In this chapter, it was confirmed that sense transcription and antisense transcription reinitiation occurred by using a DNA oligomer capable of binding to RNA transcript.

4.2. Materials and Methods

4.2.1. DNA preparation

A probing sequence, which is a five repeating 21-bp sequence, was inserted into the transcription unit to observe the RNA transcript through fluorescence signal directly. When this region was transcribed, it can be probed by complementary binding of a Cy5-labeled DNA oligomer (Harden et al., 2016). We designed two DNAs with two transcription units to examine reinitiation. L+2P consists of a short transcription unit containing intrinsic terminator tR2 of phage λ and a long transcription containing the probing sequence, and there is an EcoRI recognition site between two transcription units. L+2PR consists of a transcription unit containing tR2 terminator and promoter2 for antisense transcription in the downstream direction (Figure 4.1). Probing sequence is located between TS and promoter2, so both RNA transcribed from promoter1 and promoter2 can be observed using different oligomers. In addition, we also

used L+1P consisting of only one transcription unit containing probing sequence and tR2 terminator to confirm upstream reinitiation by backward diffusion of RNAP (Figure 4.1). For the control experiment, L+P2only, in which upstream of promoter2 was removed from L+2P, was used (Figure 4.1). All DNAs were prepared by PCR using AccuPower ProFi Taq PCR premix from Bioneer, Korea. All PCR products were purified using the Cleanup kit Expin PCR SV mini from GeneAll, Korea. All amplification templates and primers labeled with biotin or Cy5 were purchased from Integrated DNA Technologies, USA (Table 4.1 and Table 4.2).

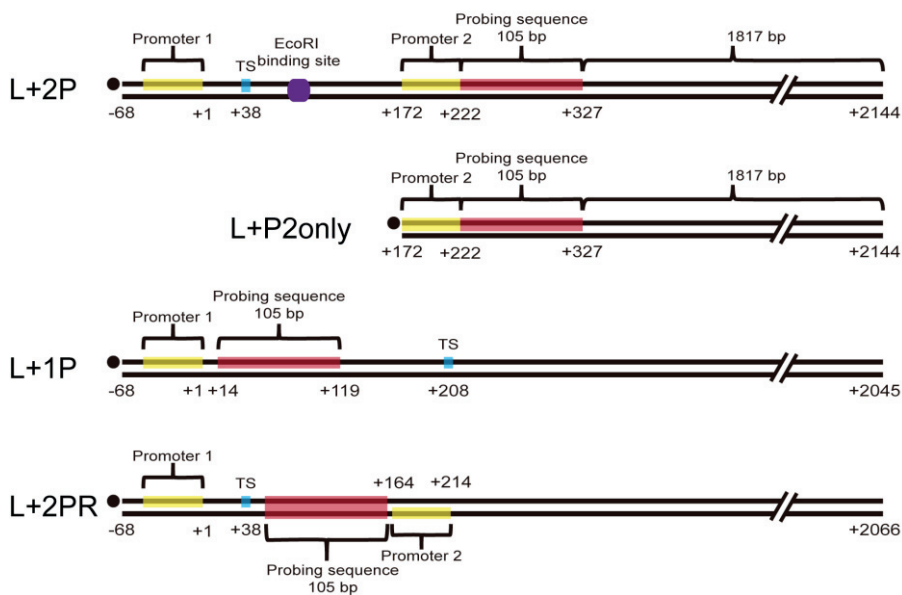


Figure 4.1. Design of various DNA templates. All DNAs were labeled with biotin (black dot) at the 5'-end of the nontemplate strand.

At first, we constructed the upstream part of each DNA. For the upstream part of L+2P and L+P2only, L+2P_partA and L+2P_partB were annealed with splint_L+2P by cooling from 90 °C to 30 °C for 120 min in an annealing buffer (10 mM Tris-HCl, pH 8.0, 50 mM NaCl), and ligated by incubating with T4 DNA ligase 2 (NEB) at 16 °C for 16 hours. These ligated products are used for amplification reactions. The upstream part of L+2P was prepared using primerF_extension and 5_primerR_2P, and the upstream part of L+P2only was prepared using B_primerF1 and 5_primerR_2P. For the upstream part of L+2PR, L+2PR_partA and L+2PR_partB were annealed with splint_L+2PR, and the same ligation was repeated. The upstream part of L+2PR was prepared using B_primerF2 and primerR_2PR. For the upstream part of L+1P, L+1P_partA and L+1P_partB were annealed with splint_L+1P, and the same ligation was repeated. The upstream part of L+1P was prepared using B_primerF2 and 5_primerR_2P. Next, we prepared a long_tail DNA with HindIII recognition site using lambda DNA (NEB), L_primerF, and L_primerR. Then, long_tail DNA and all upstream parts were each digested with HindIII (NEB) for 1 hour at 37 °C in the Cutsmart™ buffer (NEB). Afterward, HindIII was deactivated for 20 min at 80 °C. Each cleaved upstream part and long_tail DNA were annealed and ligated by the same protocol as above and used for amplification reactions. L+2P, L+2PR, and L+1P were prepared using B_primerF2 and L_primerF. L+P2only was prepared using B_primerF and L_primerF.

Table 4.1. Sequences of amplification templates used for PCR.

Oligo name	Sequence (5' → 3')
L+2P_partA	GCGAG ATTAC CATTAGTGA ATTCG AAAAA AGCAC GCTAC CGCCC CAGGC GGTGG TGGAT CCTGG AGATC GAATT TCAAA GGTTA GTTTT TTCAT GACTT CCCTC TCCCC CAAAT AAAAA GGCCT GCGAT TACCA GCAGG CCTGT TCGAT GGCTG TAAGT ATCCT ATAGG TTAGA CTTTA AGTCA ATACT CTTTT TGATA
L+2P_partB	pTAATA TCACA TCATT AGACA CTTAT CAAAA AGAGT ATTGA CTTAA AGTCT AACCT ATAGG ATACT TACAG CCTGC AGACA CCACA GACCA CACAC AAGAC ACCAC AGACC ACACA CAAGA CACCA CAGAC CACAC ACAAG ACACC ACAGA CCACA CACAA GACAC CACAG ACCAC ACACA AGACA CCACA AGCTT AGACC
splint_L+2P	AGTGT CTAAT GATGT GATAT TAGCG AGATT ACCAT TAAGT GAATT CGAAA AA
L+2PR_partA	ACTAT CTATT CTCCC ATCTA TCAAA AAGAG TATTG ACTTA AAGTC TAACC TATAG GATAC TTACA GCCAT CGAAC AGGCC TGCTG
L+2PR_partB	pGTGGA ATTCA CTAA TGTGT GTGGT CTGTG GTGTC TTGTG TGTGG TCTGT GGTGT CTTGT GTGTG GTCTG TGGTG TCTTG TGTGT GGTCT GTGGT GTCTT GTGTG TGGTC TGTGG TGTCT GCAGG CTGTA AGTAT CCTAT AGGTT AGACT TTAAG TCAAT ACTCT TTTTG ATACA CTGCG CGATA CATAA GCTTC GACGT
Splint_L+2PR	ACACA TTAAG TGAAT TCCAC GCGAG ATTAC CATTAGTGA A

Chapter 4

L+1P_partA	TATCA AAAAG AGTAT TGACT TAAAG TCTAA CCTAT AGGAT ACTTA CAGCC ATCGA ACAGG CCTAG ACACC ACAGA CCACA CACAA GACAC CACAG ACCAC ACACA AGACA CCACA GACCA CACAC AAGAC ACCAC AGACC ACACA CAAGA CACCA CAGAC CACAC ACAGC AGGAT TAAGA AGCCA ATACA AAGGC TAC
L+1P_partB	pATCCT CACTC GGCAG AUAUG ACAAU ACAGA GGCCT GCTGG TAATC GCAGG CCTTT TTATT ACACA CAAGA CACCA CAAGC TTAGA CC
Splint_L+1P	TCTGT ATTGT CATAT CTGCC GAGTG AGGAT GTAGC CTTTG TATTG GCTTC TTAAT CCTGC

Phosphate is indicated as p.

Table 4.2. Sequences of primers.

Oligo name	Sequence (5' → 3')
primerF_extension	ACTAT CTATT CTCCC ATCTA TCAAA AAGAG TATTG ACTTA AAGTC
B_primerF1	Biotin-TATCA AAAAG AGTAT TGACT TAAAG TC
B_primerF2	Biotin-ACTAT CTATT CTCCC ATC
5_primerR_2P	Cy5-GCGAG ATTAC CATT A AGTGA A
PrimerR_2PR	ACGTC GAAGC TTATG TATCG CGCAG TG
L_primerF	GTTTT CTGGG TTGGT
L_primerR	GGCGG GTTTT GTTTT

4.2.2. Single-molecule experiments for detecting reinitiation

To construct the fluorescent transcription complex, 50 nM of DNA template was incubated with RNAP holoenzyme (20 nM, NEB), ATP, CTP, GTP (each 20 μ M, GE Healthcare), and Cy3-labeled ApU (250 μ M, TriLink) for 30 min at 37 °C in an initiation buffer (10 mM Tris-HCl, pH 8.0, 20 mM NaCl, 20 mM $MgCl_2$ and 1 mM dithiothreitol). When L+2P was used, transcription did not initiate at promoter2 well under the above condition because the transcription unit from promoter2 starts with uracil. In the experiments using L+P2only containing the only promoter2, fewer Cy3-RNA spots were observed than in other experiments using DNA containing promoter1 (Figure 4.2).

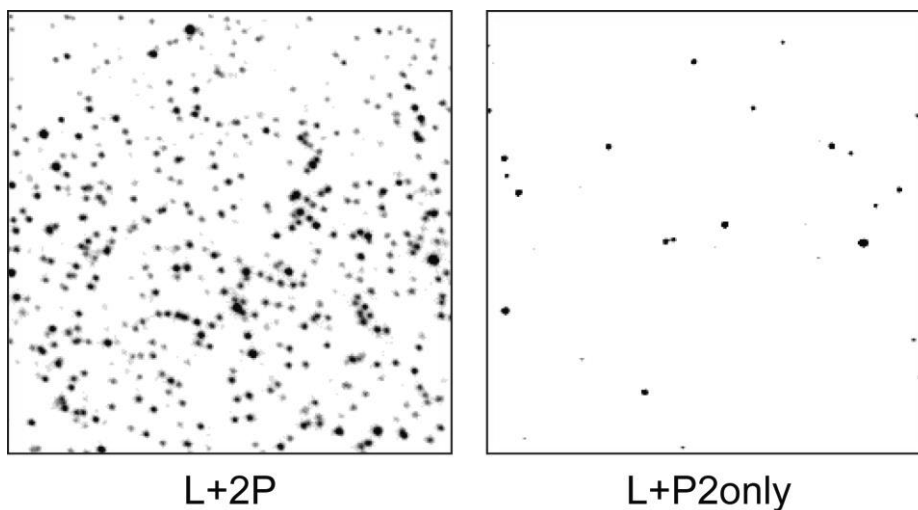


Figure 4.2. Cy3 fluorescence image at green laser excitation using different DNA templates.

Before immobilizing the stalled elongation complex on the quartz slides, quartz slides were coated with polyethylene glycol (PEG) to suppress non-specific binding (Sofia et al., 1998). At first, quartz slides were cleaned using piranha solution to remove organic residues, incubated with (3-aminopropyl) trimethoxysilane (United Chemical Technologies), and coated with polymers by incubation in a 1:40 mixture of biotin-PEG-5000 and m-PEG-5000 (Laysan Bio.) (Roy et al., 2008). After that, the slides were treated with 0.2mg/ml streptavidin (Invitrogen) for 5 min. The stalled elongation complexes were immobilized on polymer-coated quartz slides through biotin-streptavidin conjugation. Unbound RNAPs and un-immobilized elongation complexes were removed by washing with an initiation buffer and an imaging buffer extensively.

To imaging the single-molecule fluorescence from Cy3 and Cy5-labeled elongation complex, a home-made TIRF microscopy was used. A 532-nm green laser (EXLSR-532-50-CDRH, Spectra-physic) and a 640-nm red laser (EXLSR-640C-60-CDRH, Spectra-physic) were used for Cy3 excitation and Cy5 excitation, respectively. A single-molecule image was obtained by EM-CCD (Ixon DV897, Andor Technology) controlled by a customized C# program. All experiments were performed with 0.1 s exposure time in an ALEX mode at 37 °C so that the actual time resolution was 0.2 s.

All experiments were performed in a transcription buffer (10 mM Tris-HCl, pH 8.0, 20 mM NaCl, 20 mM MgCl₂, 1 mM dithiothreitol, 5 mM 3,4-protocatechuic acid and 100 nM protocatechuate-3,4-dioxygenase) or an

elongation buffer (10 mM Tris-HCl, pH 8.0, 150 mM NaCl, 2 mM MgCl₂, 1 mM dithiothreitol, 5 mM 3,4-protocatechuic acid and 100nM protocatechuate-3,4-dioxygenase). Also, 50 nM Cy5-labeled DNA oligomer was additionally added to the buffer (Table 4.3). In the experiments using L+2P and L+1P, only oligomer1 was used. On the other hand, in the experiment using L+2PR, oligomer1 was used to probe the transcript started from promote2, and oligomer2 was used to probe the transcription started from promoter1 and passed through TS. The same concentration of DNA oligomer was contained in both buffers before and after NTP injection to maintain the same background level. The results were analyzed using IDL (7.0, ITT), Matlab (R2014b, The MathWorks), and Origin (8.5, OriginLab).

Table 4.3. Sequences of Cy5-labeled DNA oligomers

Oligo name	Sequence (5' → 3')
DNA oligomer1	Cy5-TGTGT GTGGT CTGTG GTGTC T
DNA oligomer2	Cy5-AGACA CCACA GACCA CACAC A

4.3. Results

4.3.1. Fluorescent detection of upstream and downstream reinitiation

Three processes can be expected in the experiment using L+2P. First, the recycling RNAP initiates transcription again, defined as reinitiation, after RNA release at TS. We named it the reinitiation process. In this case, Cy3-

labeled RNA dissociates from DNA and unlabeled RNA, including probing sequence, which Cy5-labeled DNA oligomer can complementarily bind to, is produced by reinitiation. Second, RNA release at TS and RNAP also dissociate from DNA without reinitiation, which we named the no-reinitiation process. Third, the elongation complex ignores the intrinsic terminator, and RNAP arrives at the probing sequence without RNA release so that Cy5-labeled oligomer can complementarily bind to Cy3-labeled RNA transcript. We named it the readthrough process.

Fluorescence time traces observed from the experiment using L+2P in the transcription buffer with Cy5-labeled DNA oligomer1 were classified into three categories, which corresponds to the three processes described above:

1. Cy3-RNA signal vanished within 15s after NTP injection, and then Cy5-DNA signal appeared within 240s after NTP injection, representing the reinitiation process (Figure 4.3).
2. Cy3-RNA signal also vanished within 15s after NTP injection, but Cy5-DNA signal did not appear, representing the no-reinitiation process (Figure 4.4).
3. Cy3-RNA signal survived longer than 180s, and Cy5-DNA signal appears within 240s after NTP injection, representing the readthrough process (Figure 4.5).

In our analysis, only Cy3-RNA signal vanishing within 15 s after NTP injection was regarded as termination because RNA release by intrinsic terminator occurs at 6.1 ± 0.9 s after NTP injection on average. The traces

showing Cy3-RNA signal vanishing between 15 s and 180 s after NTP injection were excluded because there was a possibility of the termination or photobleaching. The portion of exception was minor (11 %). Also, the traces showing Cy5-DNA signal before NTP injection were excluded as non-specific probing.

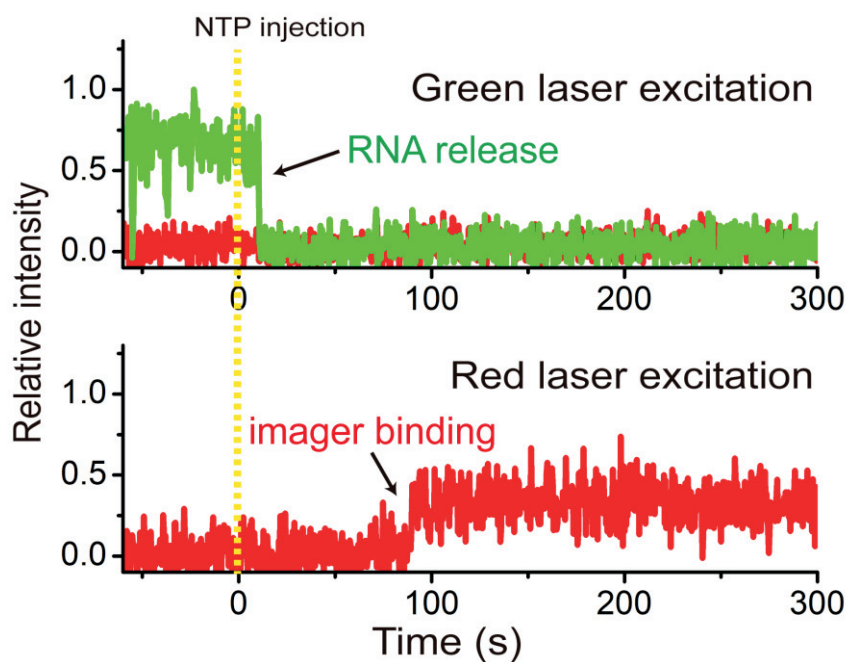


Figure 4.3. Representative fluorescence traces of the reinitiation process. Cy3 signal and Cy5 signal were represented as a green line and a red line, respectively. Cy3-RNA signal disappeared within 30 seconds after NTP injection (yellow line), and Cy5-imager signal appeared after RNA release.

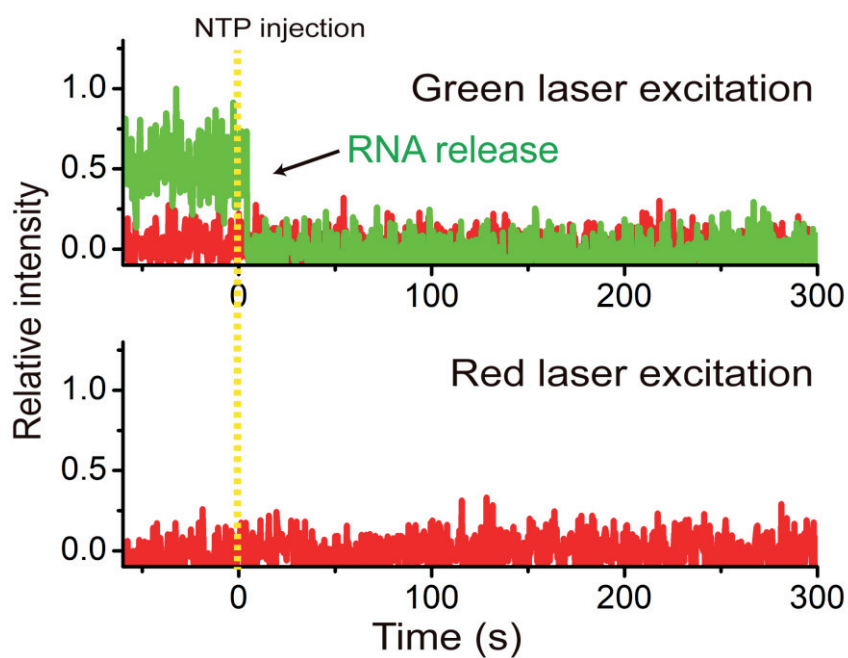


Figure 4.4. Representative fluorescence traces of the no-reinitiation process. Cy3 signal and Cy5 signal were represented as a green line and a red line, respectively. Cy3-RNA signal disappeared within 30 seconds after NTP injection (yellow line) without Cy5-probing.

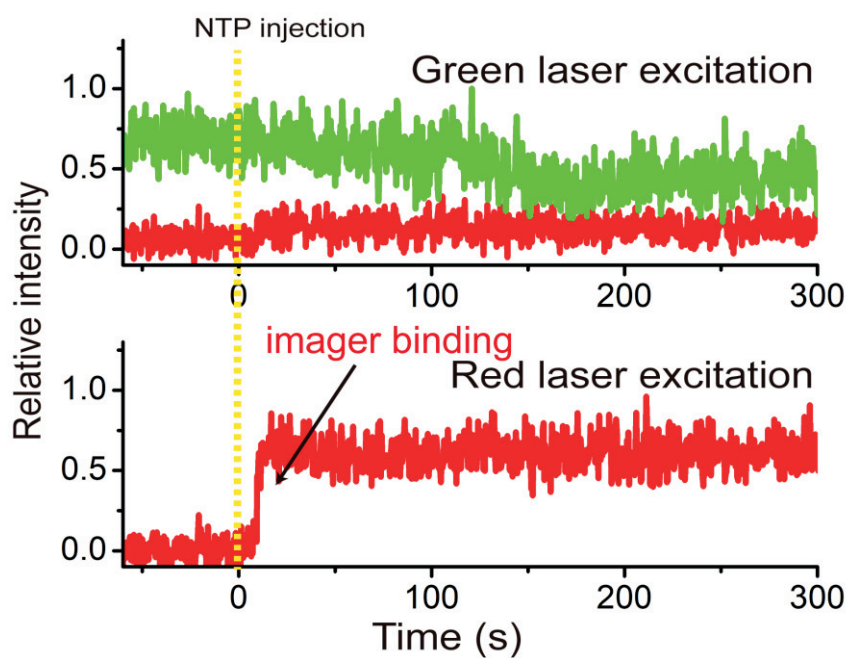


Figure 4.5. Representative fluorescence traces of the readthrough process. Cy3 signal and Cy5 signal were represented as a green line and a red line, respectively. Cy3-RNA signal did not disappear, and Cy5-imager signal appeared after NTP injection (yellow line).

The relative frequencies of reinitiation, no-reinitiation and readthrough processes were 0.09 ± 0.03 , 0.40 ± 0.11 , and 0.51 ± 0.09 , respectively. However, since there is a possibility that Cy5-labeled DNA does not bind to RNA transcript during observation time, the readthrough process was underestimated than the prediction value from termination efficiency. The sum of reinitiation and no-reinitiation processes represents total termination events. Termination efficiency of tR2 terminator is 33.4 %, therefore, the relative frequency of total readthrough events is 0.98 ($= 0.49 / 0.334 - 0.49$). Among total readthrough events, Cy5-DNA signal was observed only at 0.51, so probing efficiency, defined as the probability of probing transcript by DNA-oligomer, was 51.8 % ($= 0.51 / 0.98$). Due to incomplete probing, some reinitiation events may have been shown as the no-reinitiation process instead of the initiation process. Assuming that the probing efficiencies were the same in readthrough and termination events, we estimated the modified relative frequencies of reinitiation, no-reinitiation, and readthrough processes were 0.12, 0.21, and 0.67. Thus, the reinitiation portion among the total termination events was 36.4 %.

In the process of *E. coli* transcription initiation, σ^{70} is essential for recognizing promoter (Gross et al., 1998). It is also known that σ^{70} is dissociated from the elongation complex after transcription initiation, but some of them maintain binding to the elongation complex (Harden et al., 2016). In the experiment using Cy5-labeled σ^{70} with L+ λ , it is estimated that only 76 % of the stalled elongation complex contains σ^{70} (See Appendix C).

This ratio is higher than the result in the previous paper (Harden et al., 2016), and it is assumed that the stalled elongation complex does not entirely escape the transcription initiation process in our experiment design. It is presumed that the recycling RNAP has the same ratio of σ^{70} and reinitiation does not occur by RNAP core enzyme not containing σ^{70} . With the supplement of 3 μM σ^{70} , the holoenzyme population was increased; therefore, the reinitiation efficiency increased to 55 % (Table 4.4). On the other hand, the reinitiation efficiency decreased to 18.6 % with EcoRI E111Q binding, suggesting that reinitiation is generated by 1D diffusion of RNAP (Table 4.4).

In order to see whether reinitiation occurs in the upstream promoter or not, the experiment was conducted using L+1P instead of L+2P. In this case, the reinitiation process was still observed (Figure 4.6). With the supplement of 3 μM σ^{70} , the relative frequencies of reinitiation and no-reinitiation processes were 0.05 ± 0.01 and 0.38 ± 0.03 , respectively. Therefore, the upstream reinitiation efficiency was estimated as 43 % by modifying using probing efficiency (Table 4.4).

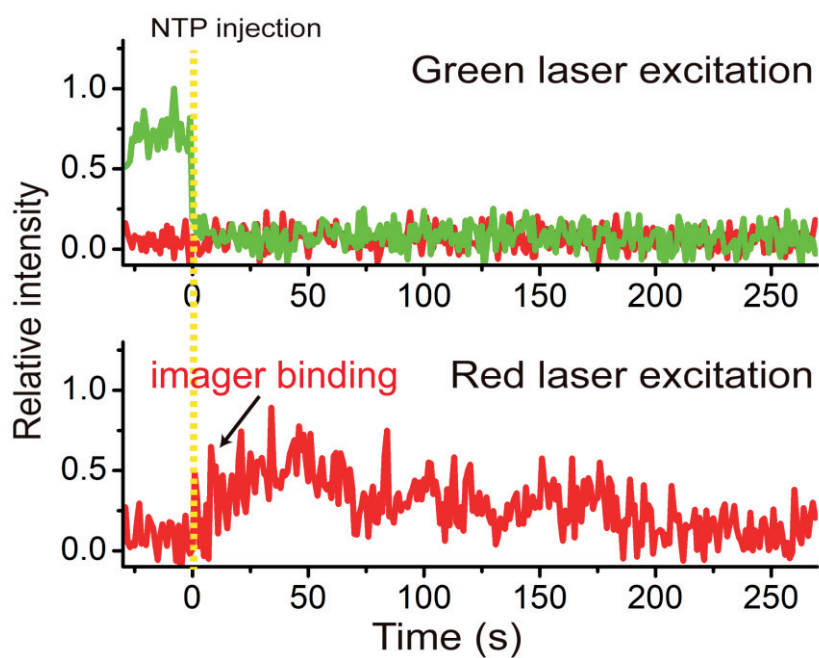


Figure 4.6. Representative fluorescence traces of the reinitiation event at upstream promoter using L+1P. Cy3 signal and Cy5 signal were represented as a green line and a red line, respectively. Cy3-RNA signal disappeared within 30 seconds after NTP injection (yellow line), and Cy5-imager signal appeared after RNA release.

Table 4.4. Relative Frequencies of three processes and reinitiation efficiency for various DNA templates with various conditions.

DNA template	Relative Frequency of reinitiation/ no-reinitiation/ readthrough (%)	Reinitiation efficiency (%)
L+2P	$9.5 \pm 2.7 / 39.7 \pm 10.6 / 50.8 \pm 9.4$	37.2
L+2P with 3uM σ^{70}	$14.1 \pm 7.3 / 35.0 \pm 11.8 / 50.9 \pm 10.1$	55.1
L+2P with EcoRI roadblock	$5.6 \pm 4.9 / 52.5 \pm 6.5 / 41.9 \pm 8.8$	18.6
L+1P with 3uM σ^{70}	$4.9 \pm 1.4 / 37.8 \pm 2.7 / 57.3 \pm 3.9$	42.9

4.3.2. Fluorescent detection of antisense reinitiation

All three processes were observed (Figure 4.7-9) in an elongation buffer with 3 μM σ^{70} and 50 nM DNA oligomer 2. The relative frequencies of reinitiation, no-reinitiation and readthrough processes were 0.07 ± 0.02 , 0.53 ± 0.03 and 0.41 ± 0.03 , respectively. On the other hand, only reinitiation and no-reinitiation processes were observed (Figure 4.10-11) with 3uM sigma70 and 50nM DNA oligomer1, and the relative frequencies of reinitiation and no-reinitiation processes were 0.10 ± 0.03 and 0.90 ± 0.03 , respectively. We assumed that the probing efficiencies of DNA oligomer1 and DNA oligomer2 were the same because these DNA oligomers have a complementary sequence. Using probing efficiency as 51.8%, the reinitiation efficiency proved by DNA oligomer1 (P₁) and DNA oligomer2 (P₂) was 22.7 % and 26.2 %, respectively (Table 4.5).

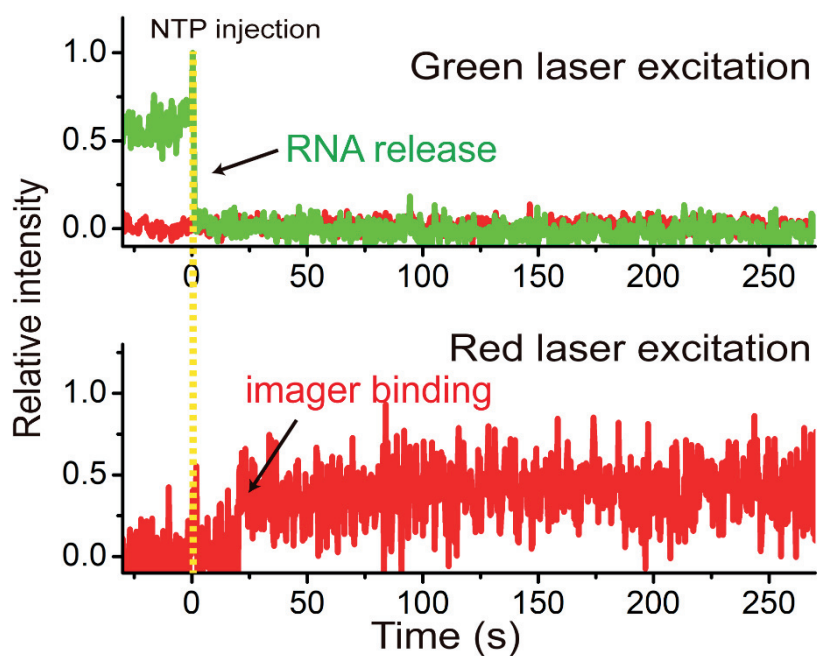


Figure 4.7. Representative fluorescence traces of the reinitiation process at promoter 1 of L+2PR using DNA oligomer2. Cy3 signal and Cy5 signal were represented as a green line and a red line, respectively. Cy3-RNA signal disappeared within 30 seconds after NTP injection (yellow line), and Cy5-imager signal appeared after RNA release.

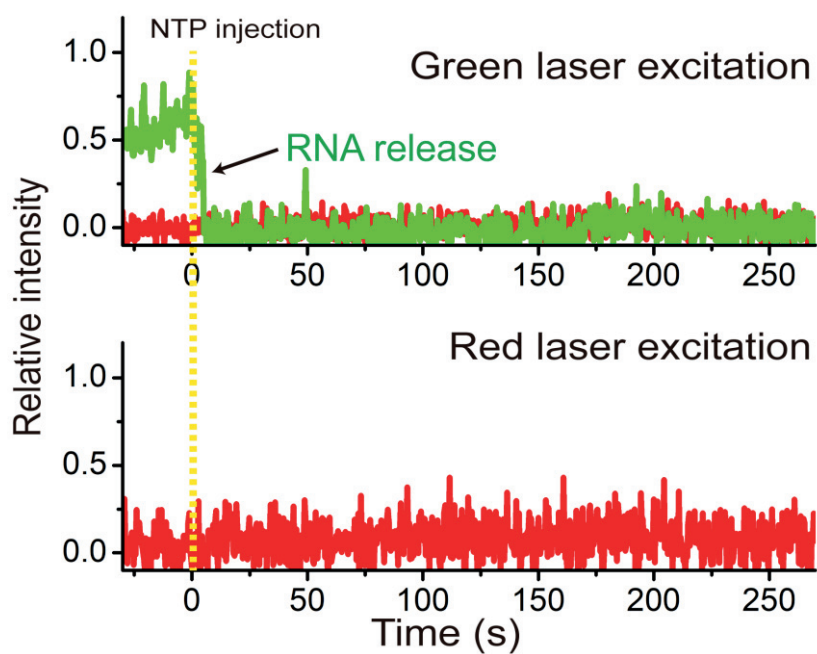


Figure 4.8. Representative fluorescence traces of the no-reinitiation process on L+2PR using DNA oligomer2. Cy3 signal and Cy5 signal were represented as a green line and a red line, respectively. Cy3-RNA signal disappeared within 30 seconds after NTP injection (yellow line) without Cy5-probing.

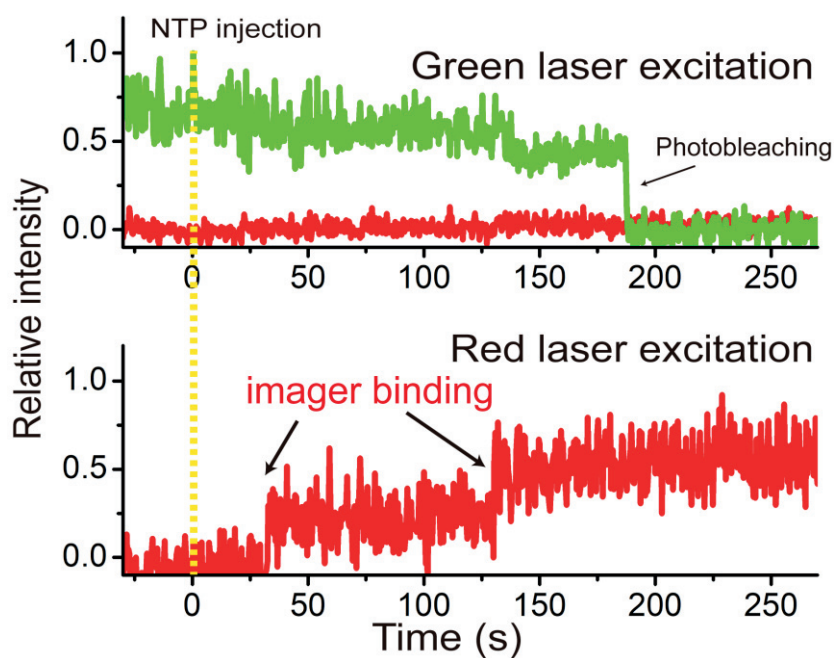


Figure 4.9. Representative fluorescence traces of the readthrough process on L+2PR using DNA oligomer2. Cy3 signal and Cy5 signal were represented as a green line and a red line, respectively. Cy3-RNA signal did not disappear, and Cy5-imager signal appeared after NTP injection (yellow line).

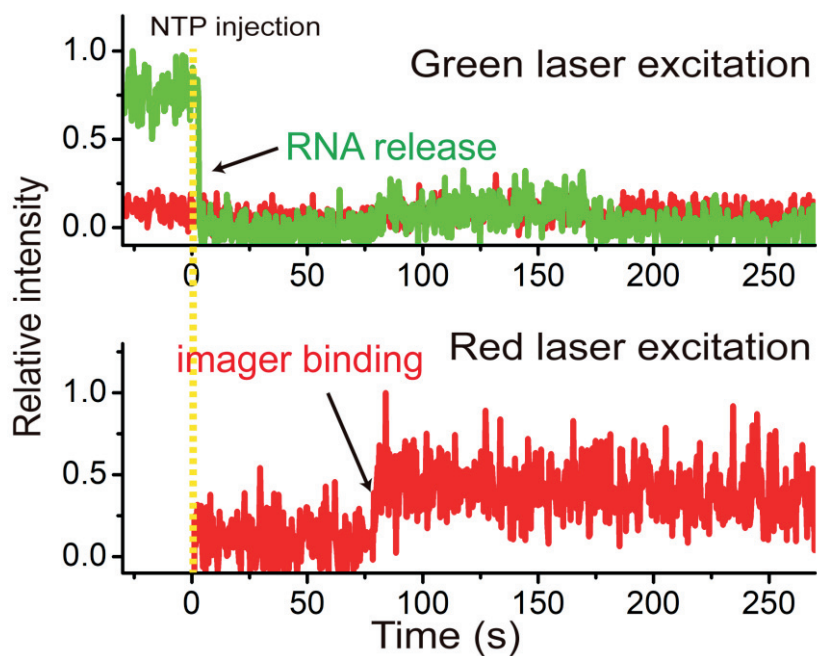


Figure 4.10. Representative fluorescence traces of the reinitiation process at promoter 2 of L+2PR using DNA oligomer1. Cy3 signal and Cy5 signal were represented as a green line and a red line, respectively. Cy3-RNA signal disappeared within 30 seconds after NTP injection (yellow line), and Cy5-imager signal appeared after RNA release.

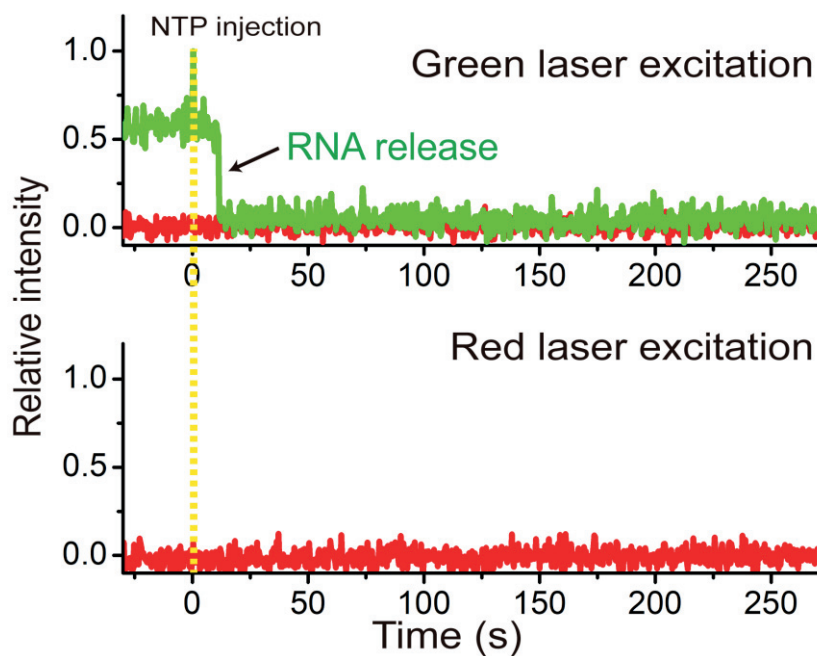


Figure 4.11. Representative fluorescence traces of the no-reinitiation process on L+2PR using DNA oligomer1. Cy3 signal and Cy5 signal were represented as a green line and a red line, respectively. Cy3-RNA signal disappeared within 30 seconds after NTP injection (yellow line) without Cy5-probing.

Table 4.5. Relative Frequencies of three processes and reinitiation efficiency for L+2PR with various DNA oligomers

DNA templates	Relative Frequency of reinitiation/ no-reinitiation/ readthrough (%)	Reinitiation efficiency (%)
L+2PR + 3uM σ^{70} + DNA oligomer1	9.9 \pm 3.2/ 90.1 \pm 3.2/ not available	22.7
L+2PR + 3uM σ^{70} + DNA oligomer2	6.8 \pm 1.6/52.6 \pm 2.8/40.6 \pm 2.7	26.2

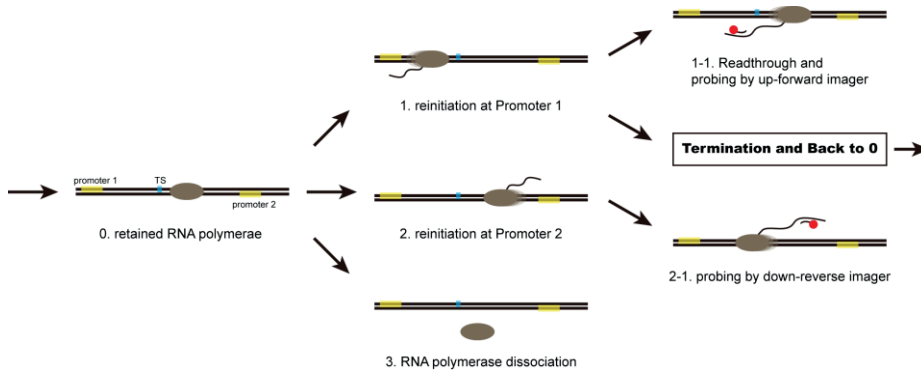


Figure 4.12. Reinitiation pathway for L+2PR. Because elongation complex reinitiated at promoter1 encounters terminator again, all reinitiation events at promoter1 are not probed by DNA oligomer1. Therefore, reinitiation efficiencies at promoter1 and promoter2 are not equal to reinitiation efficiency probed by DNA oligomer1 and oligomer2.

However, since multiple rounds of termination and reinitiation can occur, the actual reinitiation efficiencies at promoter1 (P_{sense}) and promoter2 ($P_{\text{antisense}}$) are not equal to P_1 and P_2 . Based on the reinitiation pathway (Figure 4.12), P_1 and P_2 are expressed as equations (4.1) and (4.2), where ε_T is termination efficiency.

$$\text{Equation (4.1)} \quad P_1 = \sum_{n=0}^{\infty} P_{\text{sense}} (1 - \varepsilon_T) (P_{\text{sense}} \varepsilon_T)^n$$

$$\text{Equation (4.2)} \quad P_2 = \sum_{n=0}^{\infty} P_{\text{antisense}} (P_{\text{sense}} \varepsilon_T)^n$$

We can convert these equations for P_{sense} and $P_{\text{antisense}}$ as equations (4.3) and (4.4).

$$\text{Equation (4.3)} \quad P_{\text{sense}} = P_1 / (1 - \varepsilon_T + P_1 \varepsilon_T)$$

$$\text{Equation (4.4)} \quad P_{\text{antisense}} = P_2 \left(1 - \varepsilon_T P_1 / (1 - \varepsilon_T + P_1 \varepsilon_T) \right)$$

In the elongation buffer, the termination efficiency of the tR2 terminator is measured as 39 % (See Table 3.3). The reinitiation efficiency is estimated to be 36.9 ± 3.2 % for sense transcription and 19.4 ± 6.3 % for antisense transcription. Since promoter2 is positioned farther away from TS than promoter1, we can approximate that reinitiation efficiencies are comparable for sense and antisense transcriptions.

4.4. Conclusion

Until this study, the transcription process had been understood in three stages. However, this study revealed that RNA was first released during intrinsic termination and that RNAP remained on DNA through fluorescence signals. The retention of RNAP on DNA constitutes a previously unidentified stage, and this fourth stage of transcription after termination is termed as the recycling stage. During the recycling stage, the recycling RNAP is moved on DNA via a hopping mechanism and initiate transcription again at promoter with σ^{70} .

The Recycling stage has several functional implications, although they have not been yet observed in vivo. First, they would promote transcriptional burst, which has been observed for bacterial and eukaryotic transcriptions (So et al., 2011; Chong et al., 2014; Fujita et al., 2016). The recycling stage may increase the local density of RNAP near the termination site. This accumulation of RNAP may increase the chance of reinitiation, which would cause many transcriptions in a specific region.

The hopping mechanism also makes it possible to change the recycling RNAP's orientation so that accumulated recycling RNAP may facilitate reinitiation at oppositely oriented nearby promoters, which so-called antisense transcripts. Antisense transcriptions are common in bacteria, archaea, and eukaryotic (Georg et al., 2011), particularly in the *E. coli* genome; about 30% of transcription units are antisense against other overlapping or nearby genes and operons (Thomason et al., 2015). Antisense transcription participates in

gene regulation by several mechanisms, and previous studies showed that it is related to the threshold of the genetic switch (Lenstra et al., 2015; Brophy et al., 2016). Therefore, the recycling stage would affect gene regulation by antisense transcription.

Finally, recycling and reinitiation could affect the simultaneous transcriptional regulation of multiple genes or operons. In bacterial genomes, functionally related genes and operons are often clustered and regulated together as a regulon (Rocha 2008; De et al. 2010; Meyer et al. 2014; Pannier et al. 2017). One can hypothesize that a single RNAP continually transcribes a regulon through reinitiation.

In summary, we introduced a single-molecule fluorescence assay to study bacterial transcription termination, and this assay was able to divide the transcription termination into a termination stage in which RNA is released and a recycling stage in which RNAP is recycled. As a result, a new four-stage transcription cycle was asserted (Figure 4.13).

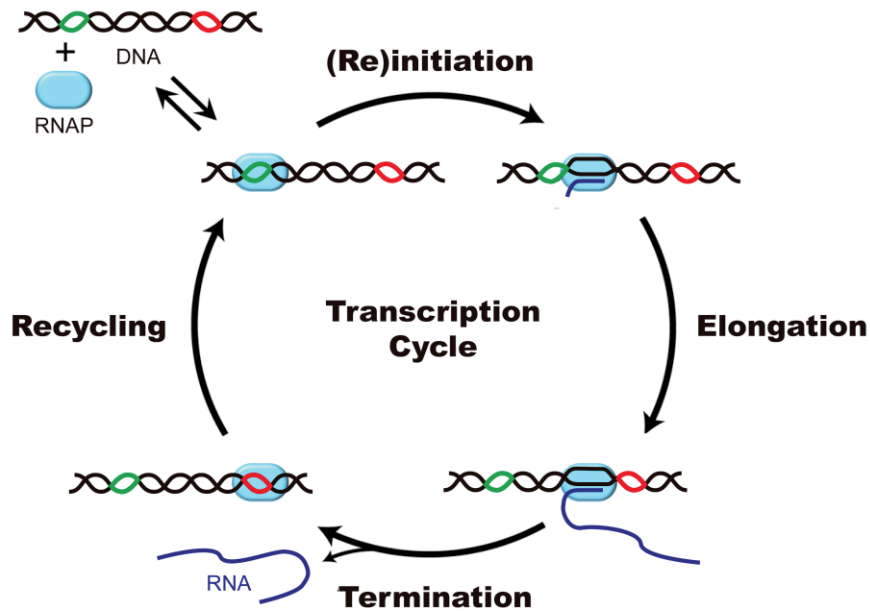


Figure 4.13. Four stages of the transcription cycle. The first stage is initiation, where RNAP (cyan oval) binds DNA (black line) at a promoter (green line). After initiation, RNAP incorporates several NTP into RNA (blue line) and escapes promoter. The second stage is elongation, where RNAP advances downward and extends RNA. The third stage is termination, where RNAP pauses at a terminator (red line) and RNA from DNA. RNAP remains on DNA after termination. The last stage is recycling, where RNAP diffuses on DNA via a hopping mechanism. The recycling RNAP can initiate at a promoter again (reinitiation) or dissociate from DNA.

References

- Brophy, J. A., & Voigt, C. A. (2016). Antisense transcription as a tool to tune gene expression. *Molecular systems biology*, 12(1), 854.
- Chong, S., Chen, C., Ge, H., & Xie, X. S. (2014). Mechanism of transcriptional bursting in bacteria. *Cell*, 158(2), 314–326.
- De, S., & Babu, M. M. (2010). Genomic neighbourhood and the regulation of gene expression. *Current opinion in cell biology*, 22(3), 326–333.
- Friedman, L. J., & Gelles, J. (2012). Mechanism of transcription initiation at an activator-dependent promoter defined by single-molecule observation. *Cell*, 148(4), 679–689.
- Friedman, L. J., Mumm, J. P., & Gelles, J. (2013). RNA polymerase approaches its promoter without long-range sliding along DNA. *Proceedings of the National Academy of Sciences of the United States of America*, 110(24), 9740–9745.
- Fujita, K., Iwaki, M., & Yanagida, T. (2016). Transcriptional bursting is intrinsically caused by interplay between RNA polymerases on DNA. *Nature communications*, 7, 13788.
- Georg, J., & Hess, W. R. (2011). cis-antisense RNA, another level of gene regulation in bacteria. *Microbiology and molecular biology reviews : MMBR*, 75(2), 286–300.
- Gross, C. A., Chan, C., Dombroski, A., Gruber, T., Sharp, M., Tupy, J., & Young, B. (1998). The functional and regulatory roles of sigma factors in transcription. *Cold Spring Harbor symposia on quantitative biology*, 63, 141–155.
- Guthold, M., Zhu, X., Rivetti, C., Yang, G., Thomson, N. H., Kasas, S., Hansma, H. G., Smith, B., Hansma, P. K., & Bustamante, C. (1999). Direct observation of one-dimensional diffusion and transcription by Escherichia coli RNA polymerase. *Biophysical journal*, 77(4), 2284–2294.

Harada Y, Funatsu T, Murakami K, Nonoyama Y, Ishihama A, Yanagida T. Single-molecule imaging of RNA polymerase-DNA interactions in real time. *Biophys J*. 1999 Feb;76(2):709-15.

Harden, T. T., Wells, C. D., Friedman, L. J., Landick, R., Hochschild, A., Kondev, J., & Gelles, J. (2016). Bacterial RNA polymerase can retain $\sigma 70$ throughout transcription. *Proceedings of the National Academy of Sciences of the United States of America*, 113(3), 602–607.

Hedglin, M., & O'Brien, P. J. (2010). Hopping enables a DNA repair glycosylase to search both strands and bypass a bound protein. *ACS chemical biology*, 5(4), 427–436.

Kabata, H., Kurosawa, O., Arai, I., Washizu, M., Margaron, S. A., Glass, R. E., & Shimamoto, N. (1993). Visualization of single molecules of RNA polymerase sliding along DNA. *Science (New York, N.Y.)*, 262(5139), 1561–1563.

Lenstra, T. L., Coulon, A., Chow, C. C., & Larson, D. R. (2015). Single-Molecule Imaging Reveals a Switch between Spurious and Functional ncRNA Transcription. *Molecular cell*, 60(4), 597–610.

Meyer, S., & Beslon, G. (2014). Torsion-mediated interaction between adjacent genes. *PLoS computational biology*, 10(9), e1003785.

Pannier, L., Merino, E., Marchal, K., & Collado-Vides, J. (2017). Effect of genomic distance on coexpression of coregulated genes in *E. coli*. *PloS one*, 12(4), e0174887.

Ricchetti, M., Metzger, W., & Heumann, H. (1988). Directed one-dimensional diffusion of *Escherichia coli* RNA-polymerase, a mechanism to facilitate promoter location. *Biochemical pharmacology*, 37(9), 1805–1806.

Rocha E. P. (2008). The organization of the bacterial genome. *Annual review of genetics*, 42, 211–233.

Roe, J. H., Burgess, R. R., & Record, M. T., Jr (1984). Kinetics and mechanism of the interaction of Escherichia coli RNA polymerase with the lambda PR promoter. *Journal of molecular biology*, 176(4), 495–522.

Roy, R., Hohng, S., & Ha, T. (2008). A practical guide to single-molecule FRET. *Nature methods*, 5(6), 507–516.

Singer, P., & Wu, C. W. (1987). Promoter search by Escherichia coli RNA polymerase on a circular DNA template. *The Journal of biological chemistry*, 262(29), 14178–14189.

So, L. H., Ghosh, A., Zong, C., Sepúlveda, L. A., Segev, R., & Golding, I. (2011). General properties of transcriptional time series in Escherichia coli. *Nature genetics*, 43(6), 554–560.

Sofia, S. J., Premnath, V., V., & Merrill, E. W. (1998). Poly(ethylene oxide) Grafted to Silicon Surfaces: Grafting Density and Protein Adsorption. *Macromolecules*, 31(15), 5059–5070.

Thomason, M. K., Bischler, T., Eisenbart, S. K., Förstner, K. U., Zhang, A., Herbig, A., Nieselt, K., Sharma, C. M., & Storz, G. (2015). Global transcriptional start site mapping using differential RNA sequencing reveals novel antisense RNAs in Escherichia coli. *Journal of bacteriology*, 197(1), 18–28.

Wang, F., Redding, S., Finkelstein, I. J., Gorman, J., Reichman, D. R., & Greene, E. C. (2013). The promoter-search mechanism of Escherichia coli RNA polymerase is dominated by three-dimensional diffusion. *Nature structural & molecular biology*, 20(2), 174–181.

Appendix

A. The measurement of EcoRI E111Q binding efficiency.

When EcoRI E111Q binds to the EcoRI recognition site, it will prevent the DNA cleavage by EcoRI. Using this, we can estimate the EcoRI E111Q binding efficiency by measuring the probability of cleavage by EcoRI depending on the binding of EcoRI E111Q. First, biotin and Cy5 double-labeled DNA containing the EcoRI recognition site was immobilized on a quartz slide, and the number of Cy5-labeled DNA was counted through a single-molecule fluorescence image (Figure A.1, bottom left). Next, DNA was incubated with 1 nM EcoRI E111Q in a transcription buffer (10 mM Tris-HCl, pH 8.0, 20 mM NaCl, 20 mM MgCl₂, 1 mM dithiothreitol, 5 mM 3,4-protocatechuic acid and, 100 nM protocatechuate-3,4-dioxygenase) for 5min to attach EcoRI E111Q to the DNA. After that, the number of Cy5-labeled DNA was counted again in the same area 10 min after wild type EcoRI

Appendix

injection (Figure A.1, bottom right). We counted Cy5-labeled DNA again before and 10 minutes after wild type EcoRI injection except for EcoRI E111Q incubation (Figure A.1, top).

The number of Cy5-labeled DNA was decreased by $86 \pm 2 \%$ without EcoRI E111Q binding, whereas with EcoRI E111Q binding, the number was decreased by $26 \pm 4 \%$. Since the binding of EcoRI E111Q reduced the efficiency of EcoRI cleavage by $71 \pm 5 \%$, we estimated the EcoRI E111Q binding efficiency as this value.

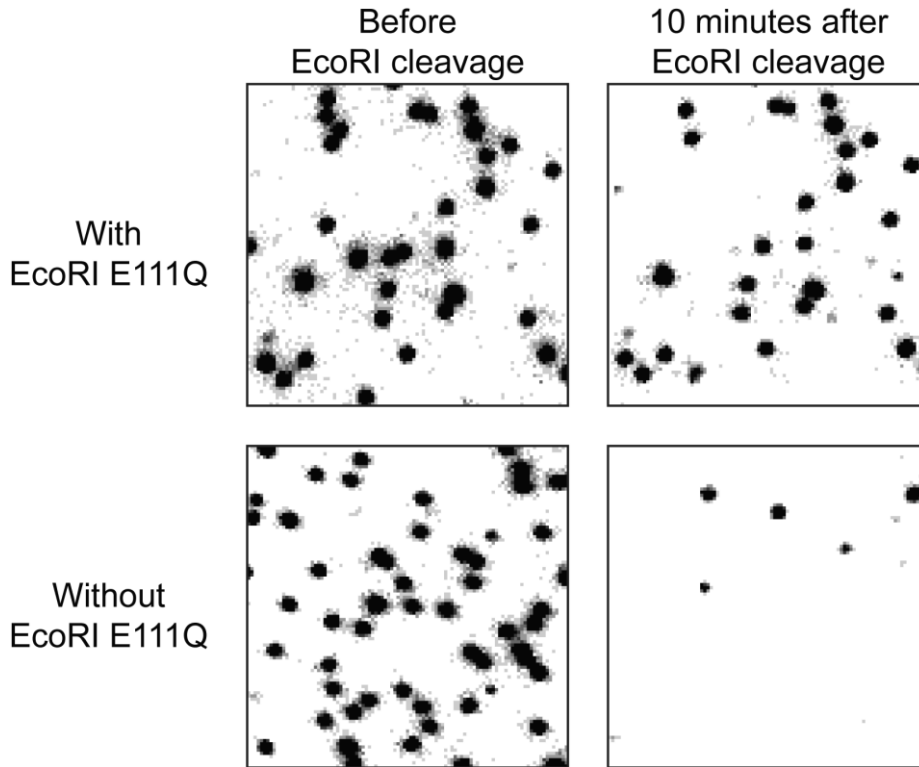


Figure A.1. Cy5 fluorescence image at red laser excitation before and 10min after EcoRI cleavage. With EcoRI E111Q, the number of Cy5 spots was reduced by 26 ± 4 %, whereas the number of Cy5 spots was reduced by 86 ± 2 % without EcoRI E111Q. From these data, we estimated the EcoRI E111Q binding efficiency as 71 ± 5 %.

B. The modeling of the recycling RNAP diffusing on DNA.

Every DNA template used in experiments has biotin and Cy5 at each end. Biotin-end binds streptavidin for immobilizing the DNA on the quartz slide, and this binding prevents RNAP from arriving at the biotin-end. On the other hand, Cy5-end is freely exposed so that RNAP can arrive at Cy5-end. Because RNAP preferentially binds DNA blunt ends, RNAP arrived at Cy5-end stays in that position for a long time. This can be confirmed that the duration of Cy5 PIFE (536 s, Figure B.1) in the experiment using L+15 with a transcription buffer is much longer than the RNAP retention time (See Table 3.3). Based on this, the biotin-end and the Cy5-end were regarded as the reflection end and the absorbing end, respectively. The recycling RNAP was modeled as a point particle that performs 1D diffusion on the line of length L (Figure B.2).

After starting from TS, RNAP performs a 1D random walk that moves δ to the right or left with the same probability every τ s. However, since RNAP does not permanently bind to DNA, it dissociates from DNA with a probability of $\tau/t_{\text{retention}}$ every τ s (Figure B.2), where $t_{\text{retention}}$ is the RNAP retention time. The mean time that it takes for RNAP to arrive at the absorbing end and the probability that RNAP arrived at the absorbing end can be defined as a function of the distance between TS and reflection end, and denoted by $W(x)$ and $P(x)$, respectively, where x is the distance between TS and reflection end. Then, $W(x)$ and $P(x)$ satisfy equations (B.1) and (B.2) (Berg 1983).

$$\text{equation (B.1)} \quad W(x) = \tau + (W(x + \delta) + W(x - \delta))\left(\frac{1}{2} - \frac{\tau}{t_{\text{retention}}}\right)$$

$$\text{equation (B.2)} \quad P(x) = (P(x + \delta) + P(x - \delta))\left(\frac{1}{2} - \frac{\tau}{t_{\text{retention}}}\right)$$

These equations are expressed as second-order differential equations such as equations (B.3) and (B.4) in the limit δ and τ goes to zero, where D is a 1D diffusion coefficient of RNAP, defined as $\delta^2/2\tau$.

$$\text{equation (B.3)} \quad W(x) - t_{\text{retention}} = t_{\text{retention}} D \frac{\partial^2}{\partial x^2} W(x)$$

$$\text{equation (B.4)} \quad P(x) = t_{\text{retention}} D \frac{\partial^2}{\partial x^2} P(x)$$

When TS is located at the reflection end ($x = 0$), the mean time and the probability do not change with x . The boundary conditions (B.5) and (B.7) were obtained from this. Also, when TS is located at the absorbing end ($x = L$), RNAP arrives at the absorbing end right after intrinsic termination, and the probability is equal to the RNAP retention probability during intrinsic termination. These two conditions generate boundary conditions (B.6) and (B.8).

$$\text{boundary condition (B.5)} \quad \frac{dW(x)}{dx} \Big|_{x=0} = 0$$

$$\text{boundary condition (B.6)} \quad W(L) = 0$$

$$\text{boundary condition (B.7)} \quad \frac{dP(x)}{dx} \Big|_{x=0} = 0$$

$$\text{boundary condition (B.8)} \quad P(L) = y_0$$

With these boundary conditions, equations (B.3) and (B.4) are solved as functions of x and the total length L and the solutions are as in equations (B.9)

and (B.10), where l_D^2 is equal to $t_{\text{retention}}D$.

$$\text{equation (B.9)} \quad W(x) = t_{\text{retention}} \left(1 - \frac{\cosh(x/l_D)}{\cosh(L/l_D)} \right)$$

$$\text{equation (B.10)} \quad P(x) = y_0 \frac{\cosh(x/l_D)}{\cosh(L/l_D)}$$

All DNA we used in the experiments to measure the diffusion coefficient (L+15, L+62, L+112, L+212, L+312, and L+512) has a constant distance of 88-bp between reflection end, which is biotin-end, and TS. Therefore, $W(x)$ and $P(x)$ can be expressed as a function of the distance between the absorbing end, which is Cy5-end, and TS. $W(b)$ and $P(b)$, where defined as $L - x$, are equations (B.11) and (B.12), respectively.

$$\text{equation (B.11)} \quad W(b) = t_{\text{retention}} \left(1 - \frac{\cosh(88/l_D)}{\cosh((88+b)/l_D)} \right)$$

$$\text{equation (B.12)} \quad P(b) = y_0 \frac{\cosh(88/l_D)}{\cosh(b+88/l_D)}$$

The 1D diffusion coefficient of RNAP can be estimated by fitting PIFE delay time data to equation (B.11), and the diffusion coefficient of RNAP and RNAP retention probability during intrinsic termination at TS can be estimated by fitting PIFE occurrence data to equation (B.12).

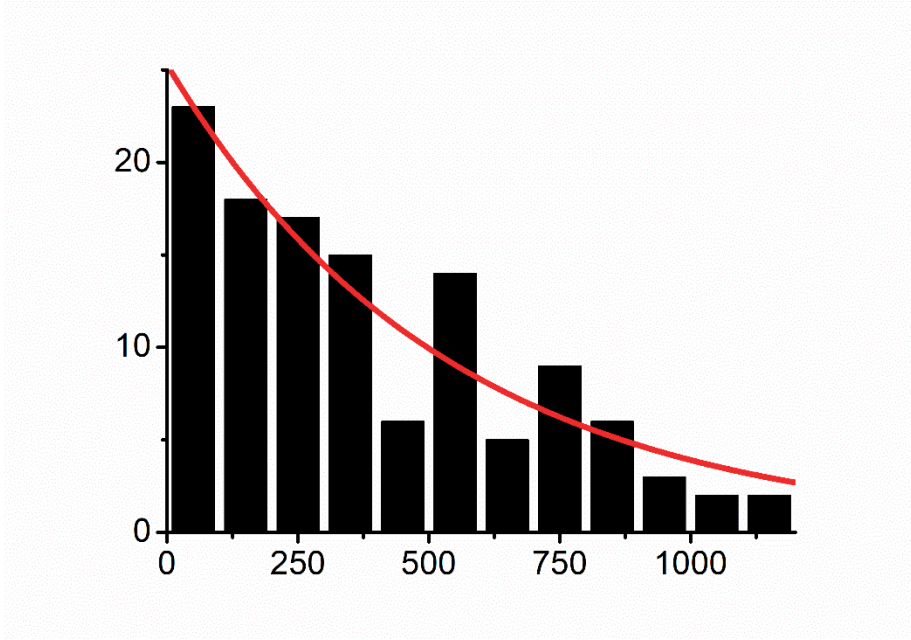


Figure B.1. The duration time of the recycling RNAP at the DNA end after RNA release. We measured Cy5 PIFE duration after RNA release ($n = 120$). The distribution was fitted to a single exponential function to obtain the duration of Cy5 PIFE of 536 ± 83 s.

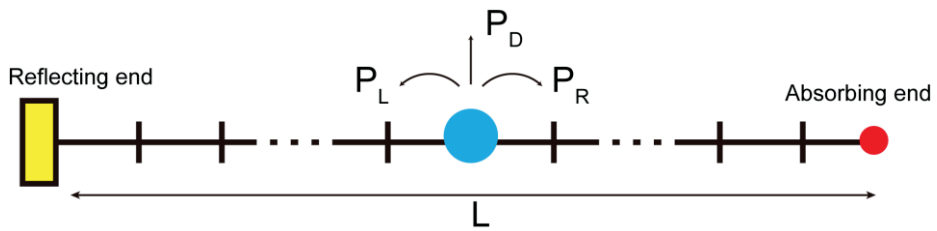


Figure B.2. The model for 1D diffusion of the recycling RNAP on DNA. Cy5-labeled DNA end was regarded as absorbing end and biotin-labeled DNA end was regarded reflecting end. RNAP, modeled as a point particle, moved left or right with same probability or dissociated from DNA.

C. The measurement of σ^{70} retention efficiency.

To measure σ^{70} retention efficiency, we used Cy5-labeled sigma70 and L+ λ . We prepared E. coli σ^{70} , where 132nd, 291st, and 295th Cys were substituted for Ser, and 366th Ser was substituted for Cys (Callaci et al., 1998). A single-Cys derivative of σ^{70} was expressed with an N-terminal His6 tag from pRPODS3666C in E. coli BL21 (DE3) at 25 °C and purified previously described (Wang et al., 2016). The single-Cys σ^{70} was incubated with Cy5-maleimide mono-reactive dye (1 mM, GE Healthcare) in the storage buffer (50 mM Tris-HCl, pH 8.0, 100 mM NaCl, and 0.1 mM EDTA) overnight at room temperature. Unreacted Cy5-maleimide molecules were removed using an Amicon ultra centrifugal filter (Merck).

To construct Cy3/Cy5-double labeled elongation complex, 50nM L+ λ was incubated with RNAP core enzyme (20 nM, NEB), Cy5-labeled σ^{70} (1 μ M), ATP, CTP, GTP (each 20 μ M, GE Healthcare), and Cy3-labeled ApU (250 μ M, TriLink) for 30 min at 37 °C in a transcription buffer (10 mM Tris-HCl, pH 8.0, 20 mM NaCl, 20 mM MgCl₂ and 1 mM dithiothreitol). Also, 50 nM Cy5-labeled DNA (L+112) was incubated with RNAP holoenzyme (20 nM, NEB), ATP, CTP, GTP (each 20 μ M, GE Healthcare), and Cy3-labeled ApU (250 μ M, TriLink) as same protocol to construct Cy3-labeled elongation complex with Cy5-labeled DNA.

Each of these elongation complexes was immobilized on a polymer-coated quartz slide and imaged by EM-CCD (Ixon DV897, Andor Technology). We measured the Cy3-Cy5 colocalized spot ratio to the Cy3

spot from these images (Figure C.1). When Cy5 was labeled at the DNA end, 46 ± 10 % of Cy3 was colocalized with Cy5. When Cy5 was labeled on σ^{70} , 35 ± 6 % of Cy3 was colocalized with Cy5. From these results, we estimated that 76 % ($35 / 46$) of the elongation complex has σ^{70} .

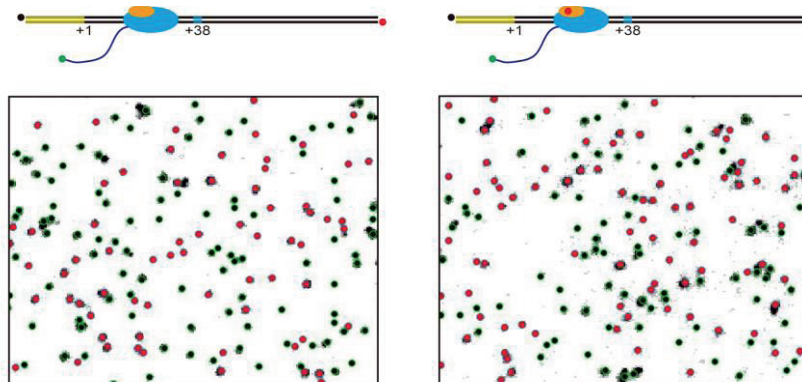


Figure C.1. σ^{70} retention efficiency. We estimated the percentage of elongation complex containing σ^{70} by comparing the number of Cy3-Cy5 colocalized spots between the cases of Cy5-end of DNA and those of Cy5- σ^{70} . In the case of Cy5-DNA, 46 ± 10 % of the Cy3 signal was colocalized with Cy5. In the case of Cy5- σ^{70} , 35 ± 6 % of the Cy3 signal was colocalized with Cy5. In the figure, the green circles, red dots, and black dots indicate all Cy3 spots identified, Cy3 spots colocalized with Cy5 and Cy3 spots non-colocalized with Cy5, respectively. From these data, we estimated that 76 % of the elongation complex contains σ^{70} .

References

Berg, H. C. (1983). Random walks in Biology (pp 37-47) Princeton Univ. Press.

Callaci, S., Heyduk, E., & Heyduk, T. (1998). Conformational changes of Escherichia coli RNA polymerase sigma70 factor induced by binding to the core enzyme. The Journal of biological chemistry, 273(49), 32995–33001.

Guanshi Wang et al 2016 Wang, G., Hauver, J., Thomas, Z., Darst, S. A., & Pertsinidis, A. (2016). Single-Molecule Real-Time 3D Imaging of the Transcription Cycle by Modulation Interferometry. Cell, 167(7), 1839–1852.e21.

Abstract in Korean (국문초록)

내재 종결 이후 RNA 중합효소의 일차원확산을
발견한 단일 분자 형광 연구

서울대학교 물리천문학부

물리학 전공

강 우 영

전사과정은 RNA 중합효소에 의해서 DNA에서 RNA를 합성하는
과정으로 유전자 발현의 조절에서 중요하다. 기존에는 전사과정이
개시, 연장, 종결의 세단계로 구성되어 있어 있다고 이해되어 왔다.
전사 개시과정에 대한 많은 연구들이 이루어졌고, 그 기작에
대해서도 많이 알려져 있다. 반면에 전사 종결에 대한 연구는 그
수도 적고 이해 역시 여전히 부족하다. 특히 DNA에서 RNA와 RNA
중합 효소 중에서 무엇이 먼저 떨어지는지, 혹은 동시에 떨어지는지
알려져 있지 않다. 우리는 원핵생물의 내재종결자(intrinsic

terminator)를 E.coli RNA 중합효소를 이용하여 단분자 형광실험을 통해서 관찰 하였다.

이 결과는 RNA 전사체가 DNA에서 먼저 떨어져 나가며 대부분의 RNA 중합효소는 DNA위에 남아있다는 것을 보여주었다. 남아있는 RNA 중합효소는 DNA를 따라 1차원 확산을 하며, RNA가 떨어져 나간 뒤 수십초 정도 후에 DNA에서 떨어져 나가게 된다. 그리고 내재종결 이후에 남아있는 RNA 중합효소의 1차원 확산 계수가 염의 농도가 증가함에 따라 증가하는 것으로부터 남아있는 RNA 중합효소가 홉핑(hopping)을 통해 DNA위를 1차원 확산을 한다는 것을 알 수 있었다.

전사 종결이 DNA에서 RNA 전사체가 떨어져 나가는 것을 의미하기 때문에, RNA 중합효소가 머무름은 전사의 새로운 네번째 단계를 구성하며 이를 여기서 재생 (recycling) 단계라고 명명 하였다. 또한 남아있는 RNAP를 재생 RNAP라고 명명하였다. 그리고 재생단계동안, RNA 중합효소가 촉진자(promoter)를 만나면 다시 전사를 시작할 수 있다는 것이 형광신호를 통해서 관찰 되었으며,

이를 우리는 전사 재개시(reinitiation)로 정의하였다. 전사 재개시는 전사 종결 위치 상류(upstream)와 하류(downstream) 모두에서 발생하며, 역류 재개시 (antisense reinitiation) 또한 발생하는 것을 확인하였다. 이러한 새로운 전사 개시의 기작은 원핵생물에서 인접한 유전자의 전사에 효율을 높여주며, 진핵생물에서도 한 데 모여있는 유전자들의 전사 효율을 높여 줄 것으로 기대한다.

핵심어: 단일분자 형광 분광학, 단백질 유도 형광 증대 (PIFE), E. coli

RNA 중합효소, 전사 종결

학번: 2014-21368