Ratcheting of RNA polymerase toward structural principles of RNA polymerase operations

Shun-ichi Sekine1,*, Yuko Murayama1, Vladimir Svetlov2, Evgeny Nudler2,3, and Shigeyuki Yokoyama4,*

1Division of Structural and Synthetic Biology; RIKEN Center for Life Science Technologies; Suehiro-cho, Tsurumi-ku, Yokohama, Japan; 2Department of Biochemistry and Molecular Pharmacology; New York University School of Medicine; New York, NY USA; 3Howard Hughes Medical Institute; New York University School of Medicine; New York, NY USA; 4RIKEN Structural Biology Laboratory; Suehiro-cho, Tsurumi-ku, Yokohama, Japan

RNA polymerase (RNAP) performs various tasks during transcription by changing its conformational states, which are gradually becoming clarified. A recent study focusing on the conformational transition of RNAP between the ratcheted and tight forms illuminated the structural principles underlying its functional operations.

Multi-subunit DNA-dependent RNA polymerase (RNAP) is a huge protein complex responsible for gene transcription. It accomplishes multiple tasks required for transcription initiation, elongation, and termination, often with assistance from various regulatory factors. The versatile nature of RNAP is supported by its conformational plasticity. Recent studies are starting to reveal the relationships between particular RNAP conformational states and functions, paving the way toward understanding the general principles of the functions and regulation of RNAP.

The bacterial RNAP core is composed of at least 5 subunits, while the eukaryotic RNAPs I, II, and III comprise 12 or more subunits, and their total masses are over 400 kDa. From bacteria to eukaryotes, RNAP adopts a similar “crab-claw” shape, which can be divided into 4 massive blocks called “modules”.1,2 The central part of RNAP is composed of the “shell” and “core” modules, which form the primary nucleic-acids-binding channel and the secondary channel, a likely path for the substrate nucleoside triphosphates (NTPs) (Fig. 1A, B). The “clamp” and “jaw-lobe” modules protrude from the shelf and core modules, respectively, to complete the primary channel. The active site is formed in the middle of the primary channel, and includes Mg2+ ions and flexible structural elements, such as the “trigger loop (TL)” and the “bridge helix (BH).”

The modules can move relative to each other around a particular rotational axis formed between them. Therefore, RNAP has a rich conformational space produced by multiple combinations of module orientations, in addition to the conformations of the flexible elements. For example, the clamp module can swing relative to the shelf module, allowing the opening and closing of the primary channel.3,4 The clamp module is closed upon DNA binding by the RNAP holoenzyme for initiation complex formation,5 while it is opened during transcriptional pausing dependent on a hairpin structure formed upstream of the nascent transcribed RNA.5 While the flexible TL assumes the “straight” or trigger-helices conformation for NTP incorporation into RNA in the transcription elongation complex (EC),6,7 it assumes a “bent” conformation in a paused or backtracked RNAP.8-10

The shelf module can also be rotated or “ratcheted” relative to the core module (Fig. 1B). We previously identified the ratcheted form of RNAP in the complex of Thermus thermophilus RNAP, bound with its inhibitor protein Gfh1 and a nucleic-acid scaffold with RNA bearing a hairpin structure.4 In contrast to the “tight” form seen in the EC,6,11 the active site of the ratcheted form exhibits extensive remodeling, including kinking of the BH, bending

Keywords: RNA polymerase, structure-function relationship, transcription, transcription factor, transcription regulation

Abbreviations: BC, backtracked complex; BH, bridge helix; CPX, Cys-pair cross-linking; EC, elongation complex; NTP, nucleoside triphosphate; RNAP, DNA-dependent RNA polymerase; TL, trigger loop.

© Shun-ichi Sekine, Yuko Murayama, Vladimir Svetlov, Evgeny Nudler, and Shigeyuki Yokoyama
*Correspondence to: Shun-ichi Sekine; Email: shunichi.sekine@riken.jp, Shigeyuki Yokoyama; Email: yokoyama@riken.jp
Submitted: 04/14/2015
Revised: 06/01/2015
Accepted: 06/01/2015
http://dx.doi.org/10.1080/21541264.2015.1059922
This is an Open Access article distributed under the terms of the Creative Commons Attribution-NonCommercial License (http://creativecommons.org/licenses/by-nc/3.0/), which permits unrestricted non-commercial use, distribution, and reproduction in any medium, provided the original work is properly cited. The moral rights of the named author(s) have been asserted.
of the TL, and expansion of the DNA/RNA binding site. The reorganization of the active site led us to hypothesize that the ratcheted form is involved in various transcriptional functions other than nucleotide addition, beyond the transcription inhibition by Gfh1. Our recent study based on Cys-pair crosslinking (CPX) analyses combined with X-ray crystallography has revealed that the ratcheted form indeed participates in many essential functions of RNAP. Thus, the tight-ratched transition added a new axis to the conformational space of RNAP, and provided deep insights into the structurally-embedded principles underlying its operations. Here we discuss this ratcheted form and its roles in RNAP regulation, with future perspectives.

The ratcheted form in RNAP backtracking and transcript cleavage

Transcription is unidirectional (5’ to 3’), but the EC does not always move forward along the DNA. The EC sometimes pauses on the DNA, and even moves backward (or backtracks) along it, due to a transcriptional error (incorporation of a mismatched nucleotide into RNA), a lesion in the DNA, sequence contexts, etc. We demonstrated that T. thermophilus RNAP adopts the tight form in a one-nucleotide backtracked complex (BC), while it assumes the ratcheted form in a longer, 8-nucleotide BC. This is presumably because the enlarged secondary channel in the ratcheted form is more suitable to accommodate a long backtracked stretch of the RNA 3’ end, as compared to the narrower channel in the tight form. Thus, the extent of backtracking is critical for the tight-ratched transition.

In the BC, the RNA elongation function is inactive, and instead RNAP exhibits hydrolytic activity to cleave the extruded part of the RNA 3’ end at the same active site. The RNA cleavage contributes to transcription proofreading, in the case of a transcriptional error. The transcription factor GreA dramatically enhances the RNA cleavage activity by a 1–2 nucleotide backtracked RNAP. T. thermophilus GreA was revealed to induce the ratcheted form in T. thermophilus RNAP, in a one-nucleotide backtracked state. This is because the bulky coiled-coil domain of the Gre protein is compatible with the enlarged secondary channel in the ratcheted form, but not with the narrower channel in the tight form. These data, together with activity measurements of an S-S crosslinked CPX variant (fixed in the ratcheted form), suggested that the GreA-dependent RNA cleavage occurs in the ratcheted form of RNAP, while the GreA-independent, intrinsic RNA cleavage occurs in the tight form of RNAP. Consistently, in the structure of RNAP bound with a Gre protein (a chimera of GreA and Gfh1), the RNAP assumed the ratcheted form (Fig. 1C). This structure also exhibited an open clamp conformation, which is probably because the DNA/RNA scaffold included an RNA hairpin, as described below. The structure of the ratcheted form with a closed clamp has not been solved. The ratcheting would loosen the grip on the DNA/RNA, and facilitate the clamp opening. However, the clamp module can swing independently of the shelf module. Therefore, in the long backtracked or Gre-bound complex (without a hairpin), the clamp may not necessarily be opened, as proposed previously.

In addition to GreA, E. coli possesses another Gre factor homolog, GreB, which can act on a longer backtracked state of RNAP. As mentioned above, the long backtracked RNAP tends to assume the ratcheted form, which can accommodate GreB (Fig. 1C). DksA is another coiled-coil protein in E. coli, and is considered to interact with the secondary channel, although it lacks the cleavage stimulating function. It cooperates with the bacterial alarmone guanosine tetraphosphate (ppGpp) to suppress transcription initiation from many genes, as well as to increase fidelity during transcription elongation. Although the mechanism underlying the DksA activity is elusive, the ratcheted form may support it. It was previously hypothesized that ppGpp could shift E. coli RNAP to the ratcheted form.

The ratchetable tight form in RNAP backtracking and pausing

Interestingly, the one-nucleotide BC readily transitioned to the ratcheted form with the aid of GreA, whereas ECs exhibited resistance to the conformational shift. This suggested that GreA selectively acts on the BC to help the stalled RNAP cleave the RNA and thus resume transcription, without interfering with the transcribing ECs. This is quite reasonable for efficient and accurate transcription. The difference between the EC and the BC probably correlates with their TL conformations. In contrast to the straight conformation of the TL in the EC, it assumes a “bent” conformation in the BC, due to the extrusion of the RNA 3’ end. The bent TL partly opens the path from the secondary channel through the active site, and thus probably facilitates the docking of the GreA coiled-coil domain. Similarly to the BC, a paused complex due to a mismatch-frayed RNA 3’ end is also readily transitioned to the ratcheted form. Thus, the tight form can adopt either the “ratchetable” or “unratchetable” state, according to the status of the RNA 3’ end. We speculate that the ratchetable state is an unfixed state in equilibrium between the tight and partly ratcheted forms, and thereby it could serve as the precedent state for switching to the ratcheted form underlying various RNAP functions. In this regard, the ratchetable state may correspond to the so-called “elemental pause” state.

The ratcheted form in transcriptional pausing and termination

The hairpin structure formed in the nascent RNA is another important element that governs the fate of transcription. In the hairpin-dependent transcriptional pause, such as the well characterized his pause, an RNA hairpin formed upstream of the DNA/RNA hybrid stabilizes the paused state of RNAP. In the hairpin-independent, intrinsic transcription termination, the U-rich tract at the RNA 3’ end causes transcriptional pausing, and a hairpin structure formed just upstream of the
U-rich tract facilitates a conformational change of RNAP for dissociation from DNA/RNA.\textsuperscript{22} We found that the RNA hairpin induces the ratcheted form of RNAP, which may underlie the transcriptional pause and termination.\textsuperscript{10} Recent studies have revealed that the presence of the RNA hairpin in the his-paused complex is coupled with the clamp opening and the TL bending.\textsuperscript{5,23} These observations have clarified how the signal from the hairpin formation at the periphery of RNAP is transmitted to the active site. The bulky RNA hairpin in the RNA exit channel causes not only the clamp opening, but also the shelf/core ratcheting, which is inevitably accompanied by the TL bending and transcriptional pausing or complete inactivation (trapping) during termination\textsuperscript{22} (Fig. 1C). Consistently, in the structures of RNAP bound with Gre proteins, RNAP assumes the ratcheted form with an opened clamp, probably due to the accommodation of the RNA hairpin, although it was not visible.\textsuperscript{4,10}

NusA is an essential transcription factor involved in the hairpin-dependent transcriptional pausing, termination, and anti-termination. It binds near the RNA exit site including the β-flap domain, facilitates RNA duplex formation in the RNA exit channel, and slows RNAP translocation.\textsuperscript{5,24,25} Therefore, NusA may maintain the open-clamp ratcheted form induced by the hairpin. NusA is also involved in transcription-coupled DNA repair, where it helps UvrD helicase to backtrack RNAP and expose the DNA lesion for access by repair enzymes.\textsuperscript{26} Here, the ratcheted form of RNAP could support the RNAP backtracking and DNA repair.

Besides the hairpin-dependent, intrinsic termination, Rho-dependent termination is another major mechanism of bacterial transcription termination. Rho is a hexameric ATPase/helicase that binds to the nascent RNA at a specific Rho-utilization site, where it accesses the transcribing EC and causes termination, probably by inducing a conformational change in the RNAP. Although the exact structural basis of the conformational change is unknown, the Rho-dependent termination was suggested to proceed in a similar manner to the intrinsic termination.\textsuperscript{27} Therefore, we hypothesize that the ratcheted form with the opened clamp is the common intermediate for both the Rho-dependent and independent terminations. Similarly to the Rho-dependent termination by a bacterial RNAP, the Rat1 and Sen1-dependent terminations by eukaryotic RNAPs I and II are also proposed to be accompanied by an allosteric change in the EC.\textsuperscript{28} Again, the open-clamp ratcheted form could be the intermediate for the eukaryotic transcription termination mechanism.

**Insights into transcriptional regulation**

Our study revealed that many transcription functions, such as nucleotide addition, RNA cleavage, RNAP backtracking, pausing, termination, and inhibition, are accomplished by one of the 2 alternative forms of RNAP: the tight and ratcheted forms. As the active site of RNAP is formed between the shelf and core modules, switching between the tight
and ratcheted forms could be the major regulation point of the functions and activities of RNAP. Taking into account other changes, including the clamp swinging and the conformational changes in the TL and the BH, the operational code of RNAP in terms of its conformational space has emerged (Table 1). This code may also represent the basis for the regulation of RNAP functions by various elements and transcription factors.

As described above, the nascent RNA (the 3' end and hairpin) plays a key role in the tight-ratcheted transition and/or clamp swinging, and determines the fate of transcription (Fig. 1C, Table 1). GreA/B, Gfh1, and possibly DksA, which bind to the secondary channel, stabilize the ratcheted form for their functions. The functions of NusA and Rho could be based on the ratcheted form with an opened clamp. These RNA elements and transcription factors are all related to the secondary channel or the RNA exit channel of RNAP. In contrast, other changes, including the module movement, limited the RNAP conformational changes, including the module movements. An estimation of the RNA

| Shelf/Core | Clamp     | Trigger loop | Bridge helix | Functions/states | TFs                      | References |
|------------|-----------|--------------|--------------|-----------------|-------------------------|------------|
| Tight      | Closed    | Straight     | Straight     | Nucleotide addition | NusG/RfaH Ribosome    | (6)        |
| Tight      | Closed    | Mobile       | Straight     | Pre-translocation state | NusG/RfaH Ribosome  | (11)       |
| Tight      | Closed    | Bent         | Straight     | Paused (frayed) | Entry point for the actions by Gre, Rho and RNA elements | (10)       |
| Tight      | Open      | Straight/bent| Straight/kinked| ?                  | ?                        |            |
| Ratcheted  | Closed?   | Bent         | Kinked       | Gre-dependent RNA cleavage | GreA/B DksA + ppGpp? UvrD + NusA? Gfh1 | (4, 10)    |
| Ratcheted  | Open      | Bent         | Kinked       | Free RNA          | Hairpin + NusA Rho?   | (3, 5, 10, 26, 27) |
| Tight or ratcheted? | Closed or open? | Bent | Kinked       | Translocation intermediate | NusG/RfaH Ribosome |            |

Table 1. Structure-function relationships of bacterial RNAP (“conformational code”)

RNAP ratcheting and translocation

In every nucleotide addition cycle of transcription elongation, RNAP moves or translocates by a one-nucleotide step along DNA/RNA. After nucleotide addition to the nascent RNA 3' end, the pre-translocation state (the RNA 3' end residing in the i + 1 position relative to the active site) is shifted to the post-translocation state (the RNA 3' end residing in the i position). In previous studies, we hypothesized that the ratcheted form of RNAP could also correlate with an intermediary state for RNAP translocation.4,12,13 The ratcheted form represents a kinked BH, which sterically prevents the incorporation of the i + 1 DNA template nucleotide into the active site. The kinking of the BH has been suggested to correlate with RNAP translocation, as it could bias the translocation state to the post-translocation state.31-33 Therefore, it seems reasonable that the ratcheted form could mediate the shift from the pre-translocation state to the post-translocation state, in the Brownian-ratchet equilibrium.34 The ratcheted form also possesses the enlarged DNA/RNA binding site, which would have a looser grip on the DNA/RNA and thus could facilitate RNAP translocation. However, it is not clear whether this hypothesis is correct.

The question could be answered, for example, by restricting the RNAP conformation to the tight form by cross-linking or other techniques. If the RNAP translocation actually accompanies the ratcheting motion, then the restriction of the RNAP conformation would impair the RNA elongation rate, without affecting the rate of the chemical step. It was recently reported that RNAP elongates RNA in crystals,35 in which the spatial constraints should limit the RNAP conformational changes, including the module movements. An estimation of the RNA
elaboration rate in the crystal and a comparison of the rate with that in solution might provide new insights into the translocation mechanism.

Disclosure of Potential Conflicts of Interest
No potential conflicts of interest were disclosed.

Funding
This work was supported in part by Japan Society for the Promotion of Science KAKENHI Grant Number 15H04344 (to SS) and the Platform for Drug Discovery, Informatics, and Structural Life Science from the Ministry of Education, Culture, Sports, Science and Technology of Japan (to SY), and grants from the National Institutes of Health and the Howard Hughes Medical Institute (to EN).

References
1. Zhang G, Campbell EA, Minakhin L, Richter C, Severinov K, Darst SA. Crystal structure of Thermus aquaticus core RNA polymerase at 3.3 Å resolution, Cell 1999; 98, 811-824; PMID:10499798; http://dx.doi.org/10.1016/S0092-8674(00)81515-9.
2. Cramer P, Bushnell DA, Kornberg RD. Structural basis of transcription: backtracked RNA polymerase II at 3.4 Å resolution, Science 2009; 324, 1206-66; PMID:19305801; http://dx.doi.org/10.1126/science.1168729.
3. Sekine S, Murayama Y, Svetlov V, Nudler E, Yokoyama S. The ratcheted and ratchetable structural States of RNA polymerase reveal insights into multiple transcribe mechanisms, Mol. Cell 2015; 57, 408-21; PMID:25601758; http://dx.doi.org/10.1016/j.molcel.2014.12.014.
4. Vasylyev DG, Vasylyeva MN, Perederina A, Tahirov T, Aruoma IO, Kornberg RD. Structural basis for transcription elongation by bacterial RNA polymerase, Nature 2007; 448, 157-62; PMID:17581590; http://dx.doi.org/10.1038/nature05932.
5. Tagami S, Segers S, Yokoyama S. A novel form of RNA polymerase sheds light on the mechanism of transcription, Nature 2011; 472, 162-7; PMID:21922057; http://dx.doi.org/10.1038/nature10416.
6. Sekine S, Tagami S, Yokoyama S. Structural basis of transcription by bacterial and eukaryotic RNA polymerases, Curr. Opin. Struct. Biol. 2011; 22, 110-18; PMID:21255178; http://dx.doi.org/10.1016/j.sbi.2011.08.001.
7. Nudler E. RNA polymerase backtracking in gene regulation and genome instability, Cell 2012; 149, 1438-45; PMID:22726433; http://dx.doi.org/10.1016/j.cell.2012.06.003.
8. Borukhov S, Lee J, Lappenko O. Bacterial elongation factor DksA/ppGpp increase fidelity of transcription, Nat. Struct. Mol. Biol. 2007; 14, 1876-82; PMID:17581590; http://dx.doi.org/10.1038/ng.591.
9. Lappenko O, Kim SS, Lee J, Starodubtseva M, Cava F, Berenguer J, Kong XP, Borukhov S. Ph-dependent conformational switch activates the inhibitor of transcription elongation, EMBO J 2006; 25, 2151-41; PMID:16628221; http://dx.doi.org/10.1038/jem.2006.288.
10. Weisbaumer A, Leon K, Lindquist R, Darst SA. Structural basis of transcriptional pausing in bacteria, Cell 2013; 152, 431-48; PMID:23743340; http://dx.doi.org/10.1016/j.cell.2012.12.020.
11. Paul BJ, Barker MM, Ross W, Schneider DA, Webb C, Fosler JW, Gourse RL. DksA: a critical regulator of the bacterial transcription initiation machinery that potentiates the regulation of RNA promoters by ppGpp and the initiating NTP, Mol. Cell 2004; 11, 311-22; PMID:15294157.
12. Roghanian M, Zemkin N, Yuzenkova Y. Bacterial global transcription regulators DksA/ppGpp increase fidelity of transcription, Nucleic Acids Res 2015; 43, 1529-36; PMID:25605801; http://dx.doi.org/10.1093/nar/gku039.
13. Zuo Y, Wang Y, Steitz TA. The mechanism of E. coli RNA polymerase regulation by ppGpp is suggested by the structure of its complex, Mol. Cell 2013; 50, 430-6; PMID:23625685; http://dx.doi.org/10.1016/j.molcel.2013.03.020.
14. Landick R. The regulatory roles and mechanism of transcriptional pausing, Biochem. Soc. Trans 2006; 34, 1062-66; PMID:17073751.
15. Epstein V, Cardinaux CJ, Ruckenstein AE, Borukhov S, Nudler E. An allosteric path to transcription termination, Mol. Cell 2007; 28, 991-1001; PMID:18158897.
16. Nakay D, Voss M, Windgassen T, Mooney RA, Landick R. Cys-pair reporters detect a constrained trigger loop in a paused RNA polymerase, Mol. Cell 2013; 50, 882-93; PMID:23760974; http://dx.doi.org/10.1016/j.molcel.2013.05.015.
17. Yang X, Molimau S, Doherty GP, Johnston EB, Maris MA, Wright J, Rothnagel R, Hankamer B, Lewis RJ, Lewis PJ. The structure of bacterial RNA polymerase in complex with the essential transcription elongation factor NusA, EMBO Rep 2009; 10, 997-1002; PMID:19460289; http://dx.doi.org/10.1038/embor.2009.128.
18. Gusarov I, Nudler E. Control of intrinsic transcription termination by N and NusA: the basic mechanisms, Cell 2001; 107, 457-49; PMID:11719185.
19. Epstein V, Kamarchapu V, McGary K, Svetlov V, Urbich S, Proshkin S, Mizrov A, Nudler E. UvrD facilitates DNA repair by pulling RNA polymerase backtracks, Nature 2014; 505, 372-77; PMID:24402227; http://dx.doi.org/10.1038/ nature13128.
20. Epstein V, Duta D, Wade J, Nudler E. An allosteric mechanism of Rho-dependent transcription termination, Nature 2010; 463, 245-9; PMID:20075920; http://dx.doi.org/10.1038/nature08669.
21. Borukhov S, Lister D. Transcription termination and the control of the transcriptome: why, where and how to stop, Nat. Rev. Mol. Cell Biol 2015; 16, 190-202; PMID:25650800; http://dx.doi.org/10.1038/nrm4394.
22. Mandyam-Murumo M, Arusimovich I. Ubiquitous transcription factors display structural plasticity and diverse functions: NusG proteins - Shifting shapes and paradigms, BioEssays 2015; 37, 324-34; PMID:25640595; http://dx.doi.org/10.1002/bies.201400177.
23. Burnazza RM, Frasnau SH, Svesyovanaei A, Schweimer K, Mooney RA, Landick R, Arusimovich I, Rosch P. An α helix to β barrel domain switch transforms the transcription factor RhaH into a translation factor, Cell 2012; 150, 391-303; PMID:22841789; http://dx.doi.org/10.1016/j.cell.2012.05.012.
24. Gnart AL, Cramer P, Fujii J, Bushnell DA, Kornberg RD. Structural basis of transcription: an RNA polymerase II elongation complex at 3.5 Å resolution, Science 2001; 292, 1876-82; PMID:11313499.
25. Vasylyev DG, Sekine S, Lappenko O, Lee J, Vasylyeva MN, Borukhov S, Yokoyama S. Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6 Å resolution, Nature 2002; 417, 712-9; PMID:12000971; http://dx.doi.org/10.1038/nature00518.
26. Epshtein V, Cardinaux CJ, Ruckenstein AE, Borukhov S, Nudler E. A ratchet mechanism of transcription elongation, Mol. Cell 2007; 28, 991-1001; PMID:18158897.
27. Bar-Nahum G, Epshtein V, Ruckenstein AE, Ralikov R, Mustaev A, Ruckenstein J. The structure of bacterial RNA polymerase and genome instability, Cell 2012; 149, 1438-49; PMID:22769674; http://dx.doi.org/10.1016/j.molcel.2013.05.015.
28. Basu RS, Warner BA, Molodtsov V, Pupov D, Eysunina D, Fernandez-Tornero C, Kulbachinskiy A, Murakami KS. Structural basis of transcription initiation by bacterial RNA polymerase holoenzyme, J. Biol. Chem. 2014; 289, 24549-59; PMID:24973216; http://dx.doi.org/10.1016/j.jbc.2014.11.045.