Structural basis of Mfd-dependent transcription termination

Jing Shi1,2,3,†, Aijia Wen1,†, Minxing Zhao4,†, Sha Jin1, Linlin You5,6, Yue Shi2, Shuling Dong7, Xiaoting Hua8,9, Yu Zhang5 and Yu Feng1,2,3,*

1Department of Biophysics, Zhejiang University School of Medicine, Hangzhou 310058, China, 2Department of Pathology, Sir Run Run Shaw Hospital, Zhejiang University School of Medicine, Hangzhou 310058, China, 3Zhejiang Provincial Key Laboratory of Immunity and Inflammatory diseases, Hangzhou 310058, China, 4Department of Emergency Medicine of the First Affiliated Hospital, Zhejiang University School of Medicine, Hangzhou 310003, China, 5Key Laboratory of Synthetic Biology, CAS Center for Excellence in Molecular Plant Sciences, Shanghai Institute of Plant Physiology and Ecology, Chinese Academy of Sciences, Shanghai 200032, China, 6University of Chinese Academy of Sciences, Beijing 100049, China, 7School of Chemistry and Life Sciences, Suzhou University of Science and Technology, Suzhou 215009, China, 8Department of Infectious Disease, Sir Run Run Shaw Hospital, Zhejiang University School of Medicine, Hangzhou 310058, China and 9Key Laboratory of Microbial Technology and Bioinformatics of Zhejiang Province, Hangzhou 310058, China

Received August 10, 2020; Revised September 26, 2020; Editorial Decision September 28, 2020; Accepted October 01, 2020

ABSTRACT

Mfd-dependent transcription termination plays an important role in transcription-coupled DNA repair, transcription-replication conflict resolution, and antimicrobial resistance development. Despite extensive studies, the molecular mechanism of Mfd-dependent transcription termination in bacteria remains unclear, with several long-standing puzzles. How Mfd is activated by stalled RNA polymerase (RNAP) and how activated Mfd translocates along the DNA are unknown. Here, we report the single-particle cryo-electron microscopy structures of T. thermophilus Mfd-RNAP complex with and without ATPγS. The structures reveal that Mfd undergoes profound conformational changes upon activation, contacts the RNAP β1 domain and its clamp, and pries open the RNAP clamp. These structures provide a foundation for future studies aimed at dissecting the precise mechanism of Mfd-dependent transcription termination and pave the way for rational drug design targeting Mfd for the purpose of tackling the antimicrobial resistance crisis.

INTRODUCTION

Mfd is a highly conserved ATP-dependent DNA translocase in bacteria. It recognizes stalled RNA polymerase (RNAP) and removes it from DNA, leading to Mfd-dependent transcription termination. Mfd-dependent transcription termination has several physiological roles.

The best-characterized function of Mfd-dependent transcription termination is transcription-coupled DNA repair (TCR), which is a sub-pathway of nucleotide excision repair (NER) (1,2). In TCR, Mfd binds to RNAP stalled at the lesion site, displaces RNAP, and recruits NER machinery to the lesion site (3,4).

Mfd-dependent transcription termination has also been suggested to play important roles other than TCR. For example, replisome and RNA polymerase translocate along the same DNA template, often in opposite directions. These processes routinely interfere with each other and lead to catastrophic effects on genome stability and cell viability (5). Mfd may resolve conflicts between DNA replication and transcription by removing RNAP stalled at the replication fork, facilitating unimpeded replication, and thus reducing possible DNA damage (6,7).

Recent studies have shown that Mfd promotes antibiotic resistance in diverse bacterial species, including Mycobacterium tuberculosis, by increasing the mutation rates (8,9). It is proposed that Mfd-dependent transcription termination leads to mutagenic DNA repair through error-prone gap filling. Compared with wild-type strains, Δmfd strains develop antibiotic resistance much slower and to a lower level. Therefore, the combination of Mfd inhibitor and antibiotics may prevent the evolution of antimicrobial resistance.

Mfd can be functionally dissected into an N-terminal region (NTR), an RNAP interacting domain (RID), a translocation module (TM), and a C-terminal domain.

*To whom correspondence should be addressed. Tel: +86 571 8898 1287; Fax: +86 571 8898 8094; Email: yufengjay@zju.edu.cn
†The authors wish it to be known that, in their opinion, the first three authors should be regarded as Joint First Authors.

© The Author(s) 2020. Published by Oxford University Press on behalf of Nucleic Acids Research.
This is an Open Access article distributed under the terms of the Creative Commons Attribution License (http://creativecommons.org/licenses/by/4.0/), which permits unrestricted reuse, distribution, and reproduction in any medium, provided the original work is properly cited.
(CTD) (Figure 1A). Biochemical, biophysical and structural analyses have uncovered some aspects of the mechanism of Mfd-dependent transcription termination. First, the interaction between the RID and RNAP β1 domain has been identified to be essential for the function of Mfd (10–13). Second, Mfd can rescue backtracked RNAP by promoting forward translocation via ATP hydrolysis (10). Third, Mfd simultaneously interacts with RNAP via the RID and with DNA via the TM, allowing its translocase activity to generate positive torque on the DNA, thereby overwinding the transcription bubble and disrupting the transcription elongation complex (TEC) (12,14–21).

In the crystal structures of Mfd (11), the conformation of the TM is incompatible with DNA binding and ATP hydrolysis, and the determinants of NTR, which are responsible for recruiting NER machinery, are masked by the CTD. Genetic and biochemical studies suggested that conformational changes are expected upon Mfd activation (17–19,22,23). To determine the exact conformational changes upon Mfd activation, we solved the single-particle cryo-electron microscopy (cryo-EM) structures of Mfd-RNAP complexes with and without ATP γS. The structures trap the active conformation of Mfd and define the protein–protein and protein–DNA interactions that mediate Mfd-dependent transcription termination.

**MATERIALS AND METHODS**

**Protein expression and purification**

*Escherichia coli* RNAP-α70 holoenzyme was purified and assembled as previously described (24). *T. thermophilus* RNAP core enzyme and *T. thermophilus* RNAP-αA holoenzyme were purified and assembled as reported (25). NusG was purified as reported (26).

*Escherichia coli* strain BL21(DE3) (Invitrogen, Inc.) was transformed with plasmid pET28a-NH-EcoMfd encoding N-hexahistidine-tagged *E. coli* Mfd under the control of T7 promoter. Single colonies of the resulting transformants were used to inoculate 1 l LB broth containing 50 μg/ml kanamycin, cultures were incubated at 37°C with shaking until OD<sub>600</sub> = 0.6. Protein expression was induced by addition of IPTG to 1 mM, and cultures were incubated 4 h at 30°C. Then cells were harvested by centrifugation (5000 rpm; 10 min at 4°C), resuspended in 20 ml buffer A (50 mM Tris–HCl, pH 8.0, 0.1 M NaCl, 5% glycerol) and lysed using a JN-02C cell disrupter (JNBIO, Inc.). The lysate was centrifuged (12 000 rpm; 45 min at 4°C), and the supernatant was loaded onto a 2 ml column of Ni-NTA agarose (Qiagen, Inc.) equilibrated with buffer B. The column was washed with 10 ml buffer B containing 0.1 M imidazole and eluted with 10 ml buffer B containing 0.5 M imidazole. The eluate was loaded onto a 5 ml column of HiTrap Heparin HP (GE Healthcare, Inc.) equilibrated in buffer B and eluted with a 100 ml linear gradient of 0.2–1 M NaCl in buffer B. Fractions containing *T. thermophilus* Mfd were pooled and stored at −80°C. *T. thermophilus* Mfd derivatives were expressed and purified in the same way as wild type protein. Yields were ∼2 mg/l, and purities were >95%.

**ATPase activity assay**

ATPase activity assays were performed in a 96-well microplate format using a commercial kit (MAK113, Sigma-Aldrich, Inc.). Reaction mixtures contained (40 μl): 5 μM *T. thermophilus* Mfd, 0.4–2 mM ATP, 50 mM Tris–HCl (pH 7.9), 0.1 M KCl, 10 mM MgCl<sub>2</sub>, 1 mM DTT and 5% glycerol. Reaction mixtures were incubated 60 min at 37°C. 200 μl reagent (MAK113A, Sigma-Aldrich, Inc.) was added to terminate the enzyme reaction and generate the colorimetric product. The absorbance at 620 nm was measured using a SpectraMax M5 microplate reader (Molecular Devices, Inc.). Phosphate standard solution was used to calculate the extinction coefficient of the colorimetric product. Less than 2% ATP was consumed to make sure initial velocity was measured.

**RNAP displacement assays**

Template strand DNA oligonucleotide and non-template strand DNA oligonucleotide (Sangon Biotech, Inc.) were annealed at a 1:1 ratio in 10 mM Tris–HCl, pH 7.9, 0.2 M NaCl and stored at −80°C. *T. thermophilus* RNAP displacement assay was performed in reaction mixtures containing (20 μl): 2 μM RNAP holo enzyme, 0.1 μM DNA scaffold, 2 mM ATP, 0.2 mM UTP, 0.2 mM GTP, 50 mM Tris–HCl (pH 7.9), 0.1 M KCl, 10 mM MgCl<sub>2</sub>, 1 mM DTT and 5% glycerol. Reaction mixtures were incubated 10 min at 65°C, supplemented with 0.1 mg/ml heparin and 4 μM Mfd or Mfd derivative, incubated 10 min at 65°C.

*E. coli* RNAP displacement assay was performed in reaction mixtures containing (20 μl): 0.1 μM RNAP holo enzyme, 0.13 μM DNA scaffold, 2 mM ATP, 0.2 mM UTP, 0.2 mM GTP, 50 mM Tris–HCl (pH 7.9), 0.1 M KCl, 10 mM MgCl<sub>2</sub>, 1 mM DTT and 5% glycerol. Reaction mixtures were incubated 10 min at 37°C, supplemented with 0.1 mg/ml heparin and 1 μM Mfd or Mfd derivative, incubated 10 min at 37°C.
Figure 1. Cryo-EM structures of Mfd-dependent transcription termination complex. (A) The schematic representation of Mfd architecture. Cyan, NTR; blue, RID; yellow, TD1; green, TD2; purple, CTD. (B) Nucleic-acid scaffold sequence used for cryo-EM. Salmon, non-template strand DNA; red, template strand DNA; magenta, RNA. (C) The model of MTCATP\(^5\). Protein and nucleic-acid scaffold are shown as ribbon; ATP\(^5\)S is shown as spheres. Gray, RNAP; red, template strand DNA; salmon, non-template strand DNA; blue, RID; yellow, TD1; green, TD2; cyan, ATP\(^5\)S. The TM binds to the upstream dsDNA and the clamp, while the RID binds to the RNAP \(\beta\)1 domain. The cryo-EM density map and the superimposed model of ATP\(^5\)S are highlighted in a black box. (D) The model of MTCapo. View orientations and colors as in (C). The TM binds to the upstream dsDNA and the clamp, while the RID binds to the RNAP \(\beta\)1 domain. (E) The model of TEC. View orientations and colors as in (C).
The reaction mixtures were applied to 5% polyacrylamide slab gels (29:1 acrylamide/bisacylamide), electrophoresed in 90 mM Tris-borate, pH 8.0 and 0.2 mM EDTA, stained with 4S Red Plus Nucleic Acid Stain (Sangon Biotech, Inc.) according to the procedure of the manufacturer.

### Assembly of Mfd-dependent transcription termination complex

DNA oligonucleotides and RNA oligonucleotide (sequences in Figure 1B) (Sangon Biotech, Inc.) were dissolved in nuclease-free water to ~1 mM and stored at -80°C. Template strand DNA, nontemplate strand DNA, and RNA were annealed at a 1:1:1 ratio in 10 mM Tris-HCl, pH 7.9, 0.2 M NaCl and stored at -80°C. MTC\(^{ATP\gamma S}\) was prepared in reaction mixtures containing (31 μl): 14 μM T. thermophilus RNAP core enzyme, 16 μM nucleic acid scaffold, 17 μM T. thermophilus Mfd, 5 mM MgCl\(_2\) and 2 mM ATP\(^\gamma S\). MTC\(^{ATP\gamma S}\) was prepared in reaction mixtures containing (31 μl): 18 μM T. thermophilus RNAP core enzyme, 19 μM nucleic acid scaffold, 21 μM T. thermophilus Mfd, 5 mM MgCl\(_2\). T. thermophilus RNAP core enzyme was incubated with nucleic acid scaffold for 10 min at 4°C, and incubated with T. thermophilus Mfd for 10 min at 4°C.

### Cryo-EM grid preparation

Immediately before freezing, 8 mM CHAPSO was added to the sample. C-flat grids (CF-1.2/1.3-4C; Protochips, Inc.) were glow-discharged for 60 s at 15 mA prior to the application of 3 μl of the complex, then plunge-frozen in liquid ethane using a Vitrobot (FEI, Inc.) with 95% chamber humidity at 10°C.

### Cryo-EM data acquisition and processing

In the preliminary experiment, some data of E. coli Mfd complex with ATP\(^\gamma S\) were collected, but there is no density for Mfd in the final map. Then we turned to determine the structure of T. thermophilus Mfd complex. The grids were imaged using a 300 kV Titan Krios (FEI, Inc.) equipped with a K2 Summit direct electron detector (Gatan, Inc.). Images were recorded with Serial EM (27) in counting mode with a K2 Summit direct electron detector (Gatan, Inc.) with 95% chamber humidity at 10°C.

### Cryo-EM model building and refinement

The homology model of T. thermophilus Mfd was generated on Phyre2 server (32). The model of RNAP core enzyme from the structure of T. thermophilus RPo (PDB 4G7H) (33) and the homology model of T. thermophilus Mfd were fitted into the cryo-EM density map using Chimera (34). The model of nucleic acids was built manually in Coot (35). The coordinates were real-space refined with secondary structure restraints in Phenix (36).

### Fluorescence polarization assays of Mfd-DNA interaction

5‘ 6-FAM labeled DNA oligonucleotide (5‘-AGCAAGCTTATT-3‘), Sangon Biotech, Inc.) and unmodified DNA oligonucleotide (5‘-AAAGAAGCTTTTGT-3‘, Sangon Biotech, Inc.) were annealed at a 1:1 ratio in 10 mM HEPEs, pH 7.5, 50 mM KCl. Equilibrium fluorescence polarization assays were performed in a 96-well microplate format. Reaction mixtures contained (100 μl): 0–16 μM Mfd or Mfd derivative, 0.1 μM 6-FAM-labelled DNA scaffold, 10 mM HEPEs, pH 7.5, 50 mM KCl, 5 mM MgCl\(_2\) and 2 mM ATP\(^\gamma S\). Following incubation mixtures for 10 min at 25°C, fluorescence emission intensities were measured using a SpectraMax M5 microplate reader (Molecular Devices, Inc.; excitation wavelength = 494 nm; emission wavelength = 518 nm). Fluorescence polarization was calculated using:

\[
P = \frac{(I_{VV} - I_{VH})}{(I_{VV} + I_{VH})}
\]  

where \(I_{VV}\) and \(I_{VH}\) are fluorescence intensities with the excitation polarizer at the vertical position and the emission polarizer at, respectively, the vertical position and the horizontal position.

Equilibrium dissociation constant, \(K_D\), were extracted by non-linear regression using the equation:

\[
P = P_t + ((P_0 - P_t) \times [M]/(K_D + [M]))
\]  

where \(P\) is the fluorescence polarization at a given concentration of Mfd, \(P_t\) is the fluorescence polarization for free 6-FAM-labeled DNA scaffold, \(P_0\) is the fluorescence polarization for bound 6-FAM-labeled DNA scaffold, and \([M]\) is the concentration of Mfd or Mfd derivative.

### RESULTS

Overall structures of Mfd-dependent transcription termination complex

Because the interaction between E. coli RNAP and Mfd is transient, the initial attempt to determine the structure of E. coli Mfd-dependent transcription termination complex (MTC) failed. In the preliminary experiment, some data of E. coli MTC with ATP\(^\gamma S\) were collected, but there is no density for Mfd in the final map. Then we turned to determine the structure of T. thermophilus MTC. The RID and the TM are highly conserved between E. coli Mfd and T. thermophilus Mfd (Supplementary Figure S1). ATPase activity assay verified that T. thermophilus Mfd hydrolyzed ATP, while substitution of a conserved Walker B...
Mfd undergoes profound conformational changes upon activation

The TM, composed of translocation domain 1 (TD1) and translocation domain 2 (TD2), contains the characteristic motifs that identify Mfd as a RecG-like SF2 helicase (Supplementary Figure S1). The TRG (translocation in RecG) motif from TD2 is highly conserved among RecG-like SF2 helicases. Its antiparallel helical hairpin conformation is critical for coupling nucleotide hydrolysis to duplex translocation (16,37).

The structures of Mfd from different species (E. coli, M. smegmatis, M. tuberculosis and T. thermophilus) have been solved in complex with different partners (nucleotide, DNA, and RNAP) using different methods (crystallography and cryo-EM) (Figure 2) (11,13,17,19,38,39). The most striking differences among these structures are the conformational change of the TRG motif and the repositioning of the RID. The two helices of the TRG motif are almost perpendicular to each other in the absence of DNA, while they are almost antiparallel in the presence of DNA. In MTCATP$_5^\text{S}$ and MTCapo, the RID translates $\sim70$ Å relative to its position in the absence of RNAP.

Residue E572, affected ATP hydrolysis, but not ATP binding (Supplementary Figure S2A, B) (17). Electrophoretic mobility shift assay (EMSA) confirmed that T. thermophilus Mfd displaced RNAP stalled by NTP starvation, while substitution of residue E572, which is deficient in ATP hydrolysis, failed to displace stalled RNAP (Supplementary Figure S2C, D).

To obtain the structure of T. thermophilus MTC, we modified the scaffold, which has been used for structure determination of a TEC (31), by extending the upstream dsDNA from 6 bp to 40 bp (Figure 1B), which is necessary and sufficient for Mfd to function (10,17). The cryo-EM structures of MTC with and without ATP$_7^\text{S}$ (MTCATP$_5^\text{S}$ and MTCapo) were determined at 4.1 and 5.0 Å, respectively (Figure 1C and D, Supplementary Figures S3–S9, Supplementary Table S1). As expected, a clear density feature for ATP$_7^\text{S}$ is observed in MTCATP$_5^\text{S}$ (Figure 1C). The conformations of RNAP in both structures are similar with a root-mean-square deviation (RMSD) of 0.85 Å (2954 Cα atoms aligned). Although full-length Mfd was used, cryo-EM densities for only the RID and the TM were observed (Supplementary Figure S7). The RID binds to the RNAP β1 domain, while the TM binds to the upstream dsDNA and the clamp. Another class from the dataset without ATP$_7^\text{S}$ was determined at 3.1 Å, lacked the density for Mfd, and turned out to be a regular TEC (Figure 1E, Supplementary Figures S5, S7 and S8, Supplementary Table S1).
Figure 3. The upstream dsDNA binds to a positively charged groove of TM. (A) The upstream dsDNA (salmon, non-template strand DNA; red, template strand DNA) bends $\sim 40^\circ$ relative to its direction in TEC (gray). Left, cryo-EM structure of MTC$^{\text{ATP$S$}}$; right, cryo-EM structure of MTC$^{\text{apo}}$. (B) The upstream dsDNA binds to a positively charged groove between TD1 and TD2. Left, cryo-EM structure of MTC$^{\text{ATP$S$}}$; right, cryo-EM structure of MTC$^{\text{apo}}$. Protein is shown as surface colored according to the electrostatic surface potential (red, $-5$ kT; blue, $+5$ kT); nucleic-acid scaffold is shown as ribbon (salmon, non-template strand DNA; red, template strand DNA). Mfd residues are numbered as in E. coli Mfd. (C) Effect on DNA binding affinity of substituting E. coli Mfd residue K739. Fluorescence polarization assay in the presence of 2 mM ATP$\gamma$S. Error bars represent mean $\pm$ SD out of $n = 3$ experiments. (D) Effect on RNAP displacement of alanine substitution of E. coli Mfd residue K739. (E) TD2 undergoes a rotation relative to TD1 upon ATP$\gamma$S binding. The structures of MTC with and without ATP$\gamma$S are superimposed on TD1. Protein and nucleic-acid scaffold are shown as ribbon; ATP$\gamma$S is shown as spheres. The structure of MTC$^{\text{ATP$S$}}$ is colored as follows: salmon, non-template strand DNA; red, template strand DNA; yellow, TD1; green, TD2; cyan, ATP$\gamma$S, orange, TRG motif; magenta, motif VI. The structure of MTC$^{\text{apo}}$ is colored gray.
and ADP dissociation reset Mfd in an open conformation at the new DNA register. Superposition of TD1 with and without ATPγS results in different registers of TD2, which are offset along the DNA by ~1 bp in the direction of translocation (Figure 3E), suggesting that the step size of Mfd is 1 bp.

Mfd contacts the RNAP β1 domain and the clamp

The interactions between Mfd and RNAP are essentially the same in MTC with and without ATPγS. The interactions in MTC<sub>ATPγS</sub> will be discussed in the following sections due to its superior resolution.

The RID binds to the RNAP β1 domain with a buried surface area of ~579 Å<sup>2</sup> (Figure 4A and Supplementary Figure S9C). The interface has been genetically, biochemically, and structurally characterized (11–13). The structure of the RID and the RNAP β1 domain in MTC<sub>ATPγS</sub> is superimposable on the crystal structure of RID complexed with the RNAP β1 domain (Figure 4B) (13), indicating that the RID makes a similar set of interactions in both structures. Consistently, substitutions of interface residues disrupt the Mfd-RNAP interaction and cause defects in the RNAP release activity of Mfd (11,12).

TD2 binds to the clamp with a buried surface area of ~580 Å<sup>2</sup> (Figure 4C and Supplementary Figure S9D). Specifically, it interacts with the evolutionarily conserved clamp helices and rudder. Transcription initiation factor σ binds to the clamp helices with high affinity (Figure 4D) (33), which would exclude the accessibility of the clamp helices. Therefore, transcript release by Mfd is inhibited by σ (10). NusG binds to the clamp helices, as well (41). Due to its lower affinity, NusG does not interfere with Mfd in RNAP displacement assay (Supplementary Figure S11), which is expected considering the clamp helices of most elongating RNAP are pre-occupied by NusG in vivo (42).

The clamp is open in Mfd-dependent transcription termination complex

The RNAP is like a crab claw with two pincers (43). The clamp, a mobile structural module that makes up much of one pincer, undergoes swing motions that open the active center cleft to allow entry of the nucleic acid scaffold during initiation or that close the cleft around the nucleic acid scaffold to enable processive elongation (44–47). Compared with the structure of TEC, the clamp in MTC rotates open by ~14° and the DNA–RNA hybrid becomes disordered due to the loss of interaction between the hybrid and the active center cleft (Figure 5A).

Can Mfd bind to the clamp and the upstream dsDNA in the same way as in MTC if the clamp is closed? To answer this question, the structure of TEC was used as a reference to superimpose the structures of MTC via the RNAP 1 domain (Figure 4B) (13), indicating that the RID makes a similar set of interactions in both structures. Consistently, substitutions of interface residues disrupt the Mfd-RNAP interaction and cause defects in the RNAP release activity of Mfd (11,12).

TD2 binds to the clamp with a buried surface area of ~580 Å<sup>2</sup> (Figure 4C and Supplementary Figure S9D). Specifically, it interacts with the evolutionarily conserved clamp helices and rudder. Transcription initiation factor σ binds to the clamp helices with high affinity (Figure 4D) (33), which would exclude the accessibility of the clamp helices. Therefore, transcript release by Mfd is inhibited by σ (10). NusG binds to the clamp helices, as well (41). Due to its lower affinity, NusG does not interfere with Mfd in RNAP displacement assay (Supplementary Figure S11), which is expected considering the clamp helices of most elongating RNAP are pre-occupied by NusG in vivo (42).

The clamp is open in Mfd-dependent transcription termination complex

The RNAP is like a crab claw with two pincers (43). The clamp, a mobile structural module that makes up much of one pincer, undergoes swing motions that open the active center cleft to allow entry of the nucleic acid scaffold during initiation or that close the cleft around the nucleic acid scaffold to enable processive elongation (44–47). Compared with the structure of TEC, the clamp in MTC rotates open by ~14° and the DNA–RNA hybrid becomes disordered due to the loss of interaction between the hybrid and the active center cleft (Figure 5A).

Can Mfd bind to the clamp and the upstream dsDNA in the same way as in MTC if the clamp is closed? To answer this question, the structure of TEC was used as a reference to superimpose the structures of MTC via the RNAP 1 domain (Figure 4B) (13), indicating that the RID makes a similar set of interactions in both structures. Consistently, substitutions of interface residues disrupt the Mfd-RNAP interaction and cause defects in the RNAP release activity of Mfd (11,12).

Figure 4. Mfd contacts the RNAP β1 domain and the clamp. (A) The RID binds to the RNAP β1 domain. Gray, the RNAP β1 domain; blue, RID. The α-carbon atoms of residues, which have been confirmed to impair Mfd’s function, are shown as spheres. Residues are numbered as in E. coli. (B) Superposition of MTC<sub>ATPγS</sub> (colored as in (A)) on the crystal structure of RID-β1 (PDB 3MLQ, pink). (C) TD2 binds to the RNAP clamp. Gray, the clamp helices and the rudder; green, TD2. (D) σ binds to the clamp helices and excludes the accessibility of the clamp helices. The structure of RNAP holoenzyme (PDB 4G7H) and the structure of MTC<sub>ATPγS</sub> are superimposed on the clamp. Gray, the clamp helices and the rudder; green, TD2; yellow, σ conserved region σR2.

The upstream dsDNA binds to a positively charged groove of TM

The upstream dsDNA binds ~40° relative to its orientation in TEC and binds to a positively charged groove between TD1 and TD2 (Figure 3A, B). There is a kink of 29° in MTC<sub>ATPγS</sub> and 36° in MTC<sub>apo</sub> at position −16 (Supplementary Figure S10), which is consistent with the single molecule observation (15). Alanine substitution of a conserved basic residue (K739 in E. coli Mfd), which is positioned near the DNA backbone, shows defects in DNA binding assay and RNAP displacement assay (Figure 3C, D), verifying that the cryo-EM structures are biologically relevant. Furthermore, alanine substitutions of R685 and N817 in E. coli Mfd, which are positioned near the DNA backbones as well, showed severely impaired binding affinity to DNA in a previous report (40).

TD2 undergoes a rotation relative to TD1 upon ATPγS binding

Compared to MTC<sub>apo</sub>, TD2 rotates by ~33° toward TD1 in MTC<sub>ATPγS</sub> (Figure 3E and Supplementary Figure S9B). Structural comparison of MTC with and without ATPγS suggests a model for how ATP binding and hydrolysis result in Mfd translocation (Supplementary Movie S1). The TM first binds DNA in an open conformation. ATP binding then leads to the closure of TD2. Finally, ATP hydrolysis...
Figure 5. The clamp is open in Mfd-dependent transcription termination complex. (A) RNAP clamp conformational change for the three cryo-EM structures determined in this work. TEC structure is used as a reference to superimpose MTC structures via α-carbon atoms of the RNAP core module, revealing a common RNAP core module but with a rotation of the clamp. Gray, the RNAP of TEC; yellow, the open clamp of MTCATP; red, the open clamp of MTCapo. (B) Mfd cannot bind to the clamp and the upstream dsDNA in the same way as in MTC if the clamp is closed. TEC structure is used as a reference to superimpose MTC structures via α-carbon atoms of the clamp, revealing severe clashes between the upstream dsDNA and the RNAP β1 domain. Gray, the clamp and the RNAP β1 domain of TEC; green, TD2 of MTC; red, template strand DNA of MTC; salmon, non-template strand DNA of MTC.

DISCUSSION

In this work, we determined the cryo-EM structures of Mfd-dependent transcription termination complex with and without ATPγS, revealing the precise mechanism of Mfd activation and translocation. Dots are connected based on this work and previous studies (Figure 6). In the absence of DNA, the two helices of the TRG motif adopt the perpendicular configuration, and the UvrA binding determinant is sequestered by the CTD (11,17,18). After binding to DNA, the two helices of the TRG motif turn into the antiparallel configuration and couple ATP hydrolysis to duplex translocation (11,39).

If Mfd encounters a backtracked RNAP, the RID binds to the RNAP β1 domain (Figure 6A). In the meantime, Mfd translocates on the upstream dsDNA and pushes RNAP forward via ATP hydrolysis. As soon as the 3’ end of the RNA aligns with the active center, transcription resumes. Because the rate of transcription elongation (∼14 bp/s) is faster than the rate of Mfd translocation (∼7 bp/s) (40), RNAP moves forward and leaves Mfd behind. During this process, the UvrA binding determinant of Mfd remains sequestered by the CTD, so the NER machinery will not be recruited.

If Mfd encounters an RNAP stalled by DNA damage, the RID binds to the RNAP β1 domain (Figure 6B). In the meantime, the TM translocates on the upstream dsDNA via ATP hydrolysis. However, RNAP cannot move forward due to the DNA damage. The RID cannot move forward either, due to its interaction with the RNAP. Therefore, the RID translates a long distance relative to the TM after several cycles of ATP hydrolysis. When the TM steps into the clamp of RNAP, it pushes against and pries open the clamp. During this process, ATP hydrolysis drives profound conformational changes of Mfd, including the translation of the RID and the exposure of the UvrA binding determinant. Therefore, although Mfd is capable of binding both paused RNAP and RNAP stalled by DNA damage, only in the case of DNA damage can Mfd complete the conformational change to expose the UvrA binding determinant and recruit NER machinery. This model is consistent with the observation that ATP hydrolysis is required for Mfd activation (15). This model is also consistent with the proposal that Mfd kinetically discriminates stalled RNAP from backtracked RNAP (15,40,48).

A closed clamp is critical for the processivity of transcription elongation. Even when RNAP is paused by backtracking or an RNA hairpin in the RNA exit channel, the clamp remains closed (49–51). On the contrary, the clamp is open in MTC structures. The open clamp loses its interaction with the DNA-RNA hybrid and probably aids the dissociation of the DNA–RNA hybrid.

Rad26, a Swi2/Snf2 family helicase, is among the first proteins to be recruited to Pol II during the initiation of S. cerevisiae TCR. The cryo-EM structure of Rad26 in complex with Pol II shows that Rad26 binds upstream of Pol II and translocates toward Pol II, suggesting Rad26 may play a role similar to that of Mfd (52). However, structural analysis reveals the divergences between Rad26 and Mfd (Supplementary Figure S12). First, the clamp is closed and the transcription bubble is ordered in Rad26-Pol II complex, while Mfd pries open the clamp and disrupts the contacts between the transcription bubble and RNAP. Second, Rad26 bends the upstream dsDNA by ∼80°, while Mfd has a less dramatic impact on the conformation of the upstream dsDNA. Third, the second-largest subunit of Pol II is the major contact site of Rad26, while Mfd contacts both the largest and the second-largest subunits of the polymerase.

Besides its role in TCR, Mfd is proposed to be an ‘evolvability factor’ that promotes mutagenesis and is required for rapid resistance development to antibiotics (8,9). Therefore, Mfd may be an ideal target for ‘anti-evolution’ drugs that inhibit antimicrobial resistance development. The structures of Mfd in action provide a basis for rational drug design targeting Mfd. For example, because the conformational
Figure 6. Proposed model of Mfd-dependent rescue and Mfd-dependent transcription termination. (A) If Mfd encounters a backtracked RNAP, the RID binds to the RNAP and the TM binds to the upstream dsDNA. Mfd and RNAP translocate forward together via ATP hydrolysis. When the 3' end of the RNA aligns with the active center, transcription resumes as long as NTPs are available. (B) If Mfd encounters an RNAP stalled by DNA damage, the RID binds to the RNAP and the TM binds to the upstream dsDNA. The TM translocates forward via ATP hydrolysis, but the RNAP and the RID stay stationary due to the DNA damage. Eventually, the TM steps into and pries open the clamp, which will facilitate the dissociation of DNA-RNA hybrid. During this process, the UvrA binding determinant of Mfd gets exposed.

change of the TRG motif is critical for the function of Mfd, inhibitors might be designed to lock the conformation of the TRG motif.

DATA AVAILABILITY

The accession numbers for the cryo-EM density map reported in this paper is Electron Microscopy Data Bank: EMD-30117 (MTCapo), EMD-30118 (MTCATP/ATP) and EMD-30119 (TEC). The accession numbers for the atomic coordinates reported in this paper are Protein Data Bank: 6M6A (MTCapo), 6M6B (MTCATP/ATP), and 6M6C (TEC).

SUPPLEMENTARY DATA

Supplementary Data are available at NAR Online.

ACKNOWLEDGEMENTS

We thank Shenghai Chang at the Center of Cryo Electron Microscopy in Zhejiang University School of Medicine for help with cryo-EM data collection. We thank for the technical support by the Core Facilities, Zhejiang University School of Medicine.

Authors contributions: J.S., A.W., M.Z., S.J., L.Y., Y.S. and S.D. performed the experiments. X.H., Y.Z. and Y.F. supervised the experiments. All authors contributed to the analysis of the data and the interpretation of the results. Y.F. wrote the manuscript with contributions from the other authors.

FUNDING

National Key R&D Program of China [2018YFA0507800 to Y. F.]; National Natural Science Foundation of China [31970040 to Y. F.]. Funding for open access charge: National Natural Science Foundation of China [31970040 to Y. F.].

Conflict of interest statement. None declared.
44. Gnatt, A.L., Cramer, P., Fu, J., Bushnell, D.A. and Kornberg, R.D. (2001) Structural basis of transcription: an RNA polymerase II elongation complex at 3.3 Å resolution. *Science*, **292**, 1876–1882.

45. Chakraborty, A., Wang, D., Ebright, Y.W., Kollann, Y., Kortkhonejia, E., Kim, T., Chowdhury, S., Wigneshweraraj, S., Irshik, H., Jansen, R. *et al.* (2012) Opening and closing of the bacterial RNA polymerase clamp. *Science*, **337**, 591–595.

46. Feklistov, A., Bae, B., Hauver, J., Lass-Napiorkowska, A., Kalesse, M., Glaus, F., Altmann, K.H., Heyduk, T., Landick, R. and Darst, S.A. (2017) RNA polymerase motions during promoter melting. *Science*, **356**, 863–866.

47. Boyaci, H., Chen, J., Jansen, R., Darst, S.A. and Campbell, E.A. (2019) Structures of an RNA polymerase promoter melting intermediate elucidate DNA unwinding. *Nature*, **565**, 382–385.

48. Fan, J., Leroux-Coyau, M., Savery, N.J. and Strick, T.R. (2016) Reconstruction of bacterial transcription-coupled repair at single-molecule resolution. *Nature*, **536**, 234–237.

49. Abdelkareem, M., Saint-Andre, C., Takacs, M., Papai, G., Crucifix, C., Guo, X., Ortiz, J. and Weixlbaumer, A. (2019) Structural basis of transcription: RNA polymerase backtracking and its reactivation. *Mol. Cell*, **75**, 298–309.

50. Kang, J.Y., Mishanina, T.V., Bellecourt, M.J., Mooney, R.A., Darst, S.A. and Landick, R. (2018) RNA polymerase accommodates a pause RNA hairpin by global conformational rearrangements that prolong pausing. *Mol. Cell*, **69**, 802–815.

51. Guo, X., Myasnikov, A.G., Chen, J., Crucifix, C., Papai, G., Takacs, M., Schultz, P. and Weixlbaumer, A. (2018) Structural basis for NusA stabilized transcriptional pausing. *Mol. Cell*, **69**, 816–827.

52. Xu, J., Lahiri, I., Wang, W., Wier, A., Cianfrocco, M.A., Chong, J., Hare, A.A., Dervan, P.B., DiMaio, F., Leschziner, A.E. *et al.* (2017) Structural basis for the initiation of eukaryotic transcription-coupled DNA repair. *Nature*, **551**, 653–657.