LETTER

Initiation of transcription-coupled repair characterized at single-molecule resolution

Kévin Howan†, Abigail J. Smith‡, Lars F. Westblade§, Nicolas Joly¶, Wilfried Grange∥, Sylvain Zorman†, Seth A. Darst∥∥, Nigel J. Savery¶ & Terence R. Strick†

Transcription-coupled DNA repair uses components of the transcription machinery to identify DNA lesions and initiate their repair. These repair pathways are complex, so their mechanistic features remain poorly understood. Bacterial transcription-coupled repair is initiated when RNA polymerase stalled at a DNA lesion is removed by Mfd, an ATP-dependent DNA translocase. Here we use single-molecule DNA nanomanipulation to observe the dynamic interactions of Escherichia coli Mfd with RNA polymerase elongation complexes stalled by a cyclopurimidine dimer or by nucleotide starvation. We show that Mfd acts by catalysing two irreversible, ATP-dependent transitions with different structural, kinetic and mechanistic features. Mfd remains bound to the DNA in a long-lived complex that could act as a marker for sites of DNA damage, directing assembly of subsequent DNA repair factors. These results provide a framework for considering the kinetics of transcription-coupled repair in vivo, and the way to reconstruction of complete DNA repair pathways at single-molecule resolution.

In bacteria, the Mfd protein couples transcription and nucleotide excision repair (NER). Mfd is a superfamily 2 ATP-dependent DNA translocase (Supplementary Fig. 1) with two distinct functions: first, it recognizes a stalled RNA polymerase (RNAP) and uses the energy of ATP hydrolysis to dissociate the RNAP–DNA elongation complex (RDs). Second, Mfd recruits UvrA, a component of the Uvr(A)BC excinuclease machinery. As a consequence of Mfd action, bulky lesions that stall RNAP are repaired more efficiently than similar lesions repaired by the transcription-independent global NER pathway. Transcription-coupled repair (TCR) in eukaryotes is more complex, but the functional homologue of Mfd in humans (CSB, also known as ERCC6) is also a superfamily 2 DNA translocase.

Although the role of Mfd in the recognition and dissociation of a stalled RDs has been well-studied genetically, biochemically, and structurally, the mechanism by which Mfd promotes RDs dissociation has not been determined. Because Mfd can rescue backtracked RDs, it is thought that it acts by pushing RNAP downstream, in what is termed the forward-translocation mechanism. Furthermore, identification of a specific interaction between Mfd and RNAP that is required for Mfd function led to the proposal that Mfd simultaneously identifies a specific interaction between Mfd and RNAP that is required for Mfd function led to the proposal that Mfd simultaneously interacts with RNAP (by means of its RNAP-interacting domain) and with DNA (by means of its ATPase domains), allowing its translocase activity to generate positive torque on the DNA, wrenching shut the transcription bubble and thus destabilizing RDs.

To investigate the mechanism of action of Mfd on stalled RDs, we developed a single-molecule assay allowing us to monitor Mfd interactions with single stalled E. coli RNAP molecules in real time (Fig. 1a and Supplementary Fig. 2). Here the mechanical properties of a nanomanipulated DNA are used to detect a single RNAP initiating transcription, progressing to form RDs, and dissociating after reaching a terminator (Fig. 1b). By omitting cytidine triphosphate (CTP), we can stall RNAP at position +20 of the template for an indefinite amount of time (Fig. 1c and Supplementary Fig. 3).

A single-molecule time-trace obtained when Mfd is present during the RNAP stall indicates that Mfd forms a remarkably long-lived intermediate, denoted I (Fig. 1d). Analysis of changes in DNA extension during formation and resolution of I provides information on its structural properties (Supplementary Fig. 4 and refs 18, 19), revealing that, in the intermediate, at least two-thirds of the transcription bubble...
is rewound and the DNA is bent by approximately 90°. As permanganate footprinting suggests that the transcription bubble is rapidly and fully rewound by Mfd action (Supplementary Fig. 5), the DNA deformation observed in the intermediate is probably due to bending/wrapping interactions between proteins and DNA rather than residual unwinding in a partially collapsed transcription bubble.

Pulse-chase experiments show that both formation and resolution of the intermediate require ATP binding (Fig. 2a, b) and hydrolysis (Supplementary Figs 6 and 7). Moreover, when we trap the intermediate with a wash step that removes both ATP and free Mfd from solution we find that adding back only ATP is sufficient to allow resolution of the trapped intermediate (Fig. 2b). Thus, the same molecule of Mfd is responsible for formation and resolution of the intermediate. Finally, we occasionally observe that after formation of the intermediate, a second RNAP can bind to the promoter, initiate transcription, stall, and become a target for a second molecule of Mfd (Fig. 2c). No such reloading events are observed when RNAP is stalled in the absence of Mfd (Supplementary Fig. 3). Thus, as Mfd modifies the structure of RDe, it also clears the transcription start site for a new RNAP to initiate⁷,¹⁵.

Surprisingly, single-molecule experiments with Mfd alone show that in the nucleotide-bound state (that is, in the presence of ATP-γ-S but not in the presence of ATP), the protein distorts the DNA in a pattern characteristic of unwinding or wrapping (Supplementary Figs 6 and 8). In the presence of ATP, Mfd distorts the DNA provided that core RNAP is present at high concentration in solution (Supplementary Fig. 8). Because permanganate attack of DNA in the presence of Mfd and ATP-γ-S detects no unwinding (Supplementary Fig. 6), this distortion is most likely due to partial, left-handed wrapping of DNA about Mfd. However, in contrast to the remarkably stable nature of the Mfd-containing TCR intermediate in all nucleotide conditions tested, the Mfd–DNA complexes in Supplementary Figs 6 and 8 are unstable as they do not persist after free Mfd is washed out.

Real-time monitoring of the formation and resolution of the repair intermediate allows for its precise kinetic description. Although the mean lifetime $t_1$ of the stalled RDe decreases as the Mfd concentration increases (Fig. 3a), the mean lifetime $t_2$ of the intermediate complex is unchanged (Supplementary Table), confirming that the same molecule of Mfd is present throughout the reaction (Fig. 2b). Thus, formation of the intermediate is expected to obey Michaelian kinetics according to:

$$Mfd + RDe \overset{k_1}{\underset{k_2}{\rightleftharpoons}} Mfd \cdot RDe \overset{k_3}{\longrightarrow} I$$

(in which $k_1$, $k_2$, and $k_3$ denote rate constants; $I$ denotes intermediate), whereas disassembly of the intermediate is expected to depend on simple rate constants (see below). Fitting the mean lifetime data

![Figure 2](image_url)

**Figure 2** | Mfd ATP usage and displacement of stalled RNAP. Arrows indicate wash steps for trapping and release of the intermediate. For clarity, we present 200-s snapshots of the states thus obtained. a. ATP is required for formation of the intermediate. Stalled RDe, is formed, then trapped by washing out (blue arrow) free RNAP and nucleoside 5’-triphosphates (NTPs). After addition of Mfd (200 nM), DNA extension is unchanged. When supplemented with ATP (2 mM) (orange arrow), the intermediate forms. b. ATP is required for release of the intermediate. Stalled RDe, is formed and trapped as above. After addition of both Mfd and ATP (first orange arrow) the intermediate forms. After washing out free Mfd and ATP (second blue arrow) the intermediate is stable for thousands of seconds. After adding back ATP (second orange arrow) the intermediate is resolved. c. After intermediate formation, stalled RNAP is displaced from its promoter-proximal stall site. Black arrows indicate transcription initiation events; red arrows indicate intermediate formation. The first initiation event (first black arrow) is followed by intermediate formation (first red arrow), a second RNAP can then initiate transcription (second black arrow) and become stalled, forming a new substrate for another Mfd to displace (second red arrow). No reloading by a second RNAP occurs in the absence of Mfd (Supplementary Fig. 3).

![Figure 3](image_url)

**Figure 3** | Kinetic characterization of the action of Mfd on RDe. a. Mean of RDe lifetime $t_1$ versus Mfd concentration, and fit to the Michaelis–Menten model. Error bars denote s.e.m. b. Full distributions of $t_1$ for $[Mfd] = 50 \text{nM}$ ($n = 273$ events) and $[Mfd] = 750 \text{nM}$ ($n = 281$ events). Red curves denote global fits with a difference-of-two-exponentials according to the single-molecule Michaelis–Menten model (see Supplementary Materials and Supplementary Fig. 9). c. Distribution of reaction intermediate lifetime $t_2$ ($n = 146$ events). Red curve, fit with a normal distribution model with a mean of 335 ± 15 s and a standard deviation of 181 s. Experiments were performed on DNA supercoiled by +4 turns.
for $t_1$ (Fig. 3a) yields Michaelis constant ($K_m$) = 760 ± 350 nM and $1/k_3 = 19 ± 7$ s. Thus, formation of the intermediate is characterized by weak binding of Mfd to RD$_e$ followed by a remarkably slow catalytic step.

The statistical distributions of dwell-times $t_1$ preceding formation of the intermediate (Fig. 3b) allow further resolution of $K_m$ into the diffusion-controlled on-rate and the thermodynamically controlled off-rate of Mfd from RD$_e$. The shape of the distribution—an exponential rise followed by an exponential decay—is characteristic of a reaction with at least two intermediates, namely diffusion/docking of Mfd to RD$_e$ and a forward-catalytic event (see Supplementary Materials and ref. 21). Taking $k_3 = 0.05$ s$^{-1}$ from our previous analysis of $t_1$, we obtain $k_1 = 1.9 × 10^5$ M$^{-1}$ s$^{-1}$ and $k_2 = 0.11$ s$^{-1}$. By performing a global fit to the distributions obtained at four different Mfd concentrations (Supplementary Fig. 9), the constraint on $k_3$ can be relaxed, yielding $k_1 = 1.6 × 10^5 ± 0.6 × 10^5$ M$^{-1}$ s$^{-1}$, $k_2 = 0.08 ± 0.04$ s$^{-1}$, and $k_3 = 0.059 ± 0.01$ s$^{-1}$, with a reduced $\chi^2 = 0.88$, supporting the internal consistency of the data and the validity of the kinetic model.

Surprisingly, the statistical distribution of intermediate lifetimes, $t_2$ (corresponding to resolution of the intermediate), is non-exponential (Fig. 3c). It tends instead to a normal distribution characteristic of a reaction consisting of several irreversible steps with equivalent rate constants. The data thus give $t_2 = 335 ± 15$ s and a standard deviation of 181 s. Variance analysis of this distribution suggests that a minimum of three irreversible events takes place during this time, each thus having a rate constant $k_4 = 3/t_2$. Pulse-chase analysis of ATP usage in this reaction indicates that at least one of these steps requires ATP hydrolysis.

The data analysed here were collected using a substrate bearing five successive C residues on the non-template strand at the stall site, but the process leading to disassembly of the intermediate is not sequence-dependent as similar results were obtained when only one C residue was present (Supplementary Figs 10 and 11). Similarly, disassembly of the intermediate is only weakly torque-dependent, as a long-lived complex was observed using negatively supercoiled DNA: $t_2 = 180 ± 40$ s ($n = 16$; data collected as described in Supplementary Fig. 4). Finally, we analysed the interactions between Mfd and RNAP stalled on a cyclopyrimidine dimer located on the template strand (Supplementary Fig. 12). The kinetics of the entire pathway are essentially identical to those observed when RNAP is stalled by withholding a nucleotide. Indeed, a crystal structure of RNAP II stalled on a cyclopyrimidine dimer showed no conformational changes compared with elongation complexes stalled by nucleotide deprivation.

On the basis of these data, we propose a model to describe RD$_e$ displacement during TCR (Fig. 4). From a structural standpoint, the mechanism by which Mfd initiates displacement of the stalled RNAP involves sufficient collapse of the transcription bubble to destabilize the stalled RD$_e$. Concurrently, promoter-proximal DNA is cleared for another RNAP to initiate (Fig. 2c), indicating either forward translocation of RNAP or full dissociation of RNAP. Mfd remains associated with the DNA long after the bubble collapse. The source of the DNA deformation observed in the intermediate cannot be fully ascribed from our current data. On the basis of the observation that free Mfd is capable of distorting DNA, we favour a model in which the transcription bubble is fully collapsed after formation of the intermediate, and the residual DNA deformation observed in the long-lived intermediate is caused by Mfd (Fig. 4). Because Mfd alone does not form long-lived complexes on DNA, we propose that the intermediate is stabilized by a combination of Mfd–DNA interactions, Mfd–RNAP interactions (mediated by the RNAP-interacting domain) and, potentially, interaction of the Mfd-tethered core RNAP with DNA (Fig. 4). Binding of Mfd to RNAP derepresses the DNA translocation domains of the helicase, and so the Mfd–DNA interactions in this activated ternary complex probably differ from those made by the isolated protein.

From a kinetic standpoint, the rate constants $k_1$, $k_2$, and $k_3$ indicate that Mfd attempts binding to RD$_e$ several times before engaging productively, and the rate-limiting step itself is slow. This suggests that Mfd kinetically discriminates stalled RD$_e$ from paused RD$_e$, as only stalled RD$_e$ is sufficiently long-lived to act as an Mfd target. This model is supported by our observation that the kinetics of displacement are the same for RNAP stalled at a CPD lesion or by nucleotide starvation. The slow catalytic event could be related to derepression of Mfd, which probably involves large conformational changes$^{9,24}$. The intermediate $I$ is formed in an ATP hydrolysis-dependent step, and Mfd then performs at least three irreversible steps before releasing the DNA. As ATP hydrolysis is required to proceed out of this intermediate state, we propose that Mfd hydrolyses a second ATP molecule, and possibly more, before dissociating.

The lifetime of the intermediate is surprisingly long for an organism that repairs its DNA in much shorter times and which can divide every 20 min; in vivo, downstream repair proteins may capture and process the intermediate, reducing this period. Nevertheless, the fact that the intermediate is reliably long-lived could ensure that repair components have ample opportunity to assemble.

The recruitment of UvrA by Mfd to the damage site is not well understood. The interaction between UvrA and Mfd is inhibited by inter-domain contacts within free Mfd, and the interaction between Mfd and RNAP probably overcomes this inhibition by domain repositioning (Supplementary Fig. 1 and refs 8, 11, 24 and 26–28). As Mfd has a central role in the recruitment of the NER machinery, it will be interesting to see how the addition of subsequent players (such as UvrA, UvrB and UvrC) affects the temporal properties of the system,
and from there to pursue bottom-up reconstruction of the full DNA repair pathway at single-molecule resolution.

**METHODS SUMMARY**

See legend to Fig. 1 for the Methods Summary; full Methods can be found in the Supplementary Information.

Received 3 January; accepted 20 July 2012. Published online 9 September 2012.

1. Selby, C. P. & Sancar, A. Molecular mechanism of transcription-repair coupling. Science 260, 53–58 (1993).
2. Witkin, E. M. Time, temperature, and protein synthesis: a study of ultraviolet-induced mutation in bacteria. Cold Spring Harb. Symp. Quant. Biol. 21, 123–140 (1956).
3. Mellon, I. & Hanawalt, P. C. Induction of the Escherichia coli lactose operon selectively increases repair of its transcribed DNA strand. Nature 342, 95–98 (1989).
4. Hanawalt, P. C. & Spivak, G. Transcription-coupled DNA repair: two decades of progress and surprises. Nature Rev. Mol. Cell Biol. 9, 958–970 (2008).
5. Selby, C. P. & Sancar, A. Gene- and strand-specific repair in vitro: partial purification of a transcription-repair coupling factor. Proc. Natl Acad. Sci. USA 88, 8232–8236 (1991).
6. Selby, C. P. & Sancar, A. Transcription-repair coupling and mutation frequency decline. J. Bacteriol. 175, 7509–7514 (1993).
7. Park, J. S., Marr, M. T. & Roberts, J. W. E. coli transcription repair coupling factor: autoinhibition and the role of RNA polymerase. Nucleic Acids Res. 31, 1802–1811 (2003).
8. Smith, A. J., Szczelukin, M. D. & Savery, N. J. Controlling the motor activity of a transcription-repair coupling factor: autoinhibition and the role of RNA polymerase. Nucloic Acids Res. 35, 1802–1811 (2007).
9. Smith, A. J. & Savery, N. J. Effects of the bacterial transcription-repair coupling factor during transcription of DNA containing non-bulky lesions. DNA Repair (Amst.) 7, 1670–1679 (2008).
10. Chambers, A. L., Smith, A. J. & Savery, N. J. A DNA translocation motif in the transcription-repair coupling factor. Mol. Cell. 9, 6409–6418 (2003).
11. Deaconescu, A. M. et al. Structural basis for bacterial transcription-coupled DNA repair. Cell 124, 507–520 (2006).
12. Deaconescu, A. M. & Darst, S. A. Crystallization and preliminary structure determination of Escherichia coli Mfd, the transcription-repair coupling factor. Acta Crystallogr. F 61, 1062–1069 (2005).
13. Westblad, L. F. et al. Structural basis for the bacterial transcription-repair coupling factor/RNA polymerase interaction. Nucleic Acids Res. 38, 8357–8369 (2010).
14. Deaconescu, A. M., Savery, N. J. & Darst, S. A. The bacterial transcription repair coupling factor. Curr. Opin. Struct. Biol. 17, 96–102 (2007).
15. Park, J. S. & Roberts, J. W. Role of DNA bubble rewinding in enzymatic transcription termination. Proc. Natl Acad. Sci. USA 103, 4870–4875 (2006).
16. Savery, N. J. The molecular mechanism of transcription-coupled DNA repair. Trends Microbiol. 15, 326–333 (2007).
17. Reyakin, A., Liu, C. Y., Ebright, R. H. & Strick, T. R. Abortive initiation and productive initiation by RNA polymerase involve DNA scrunching. Science 314, 1139–1143 (2006).
18. Reyakin, A., Ebright, R. H. & Strick, T. R. Promoter unwinding and promoter clearance by RNA polymerase: detection by single-molecule DNA nanomanipulation. Proc. Natl Acad. Sci. USA 101, 4776–4780 (2004).
19. Reyakin, A. et al. Single-molecule DNA nanomanipulation: detection of promoter-unwinding events by RNA polymerase. Methods Enzymol. 370, 577–598 (2003).
20. Smith, A. J. & Savery, N. J. RNA polymerase mutants defective in the initiation of transcription-coupled DNA repair. Nucleic Acids Res. 33, 755–764 (2005).
21. Kou, C. et al. Single-molecule Michaelis–Menten equations. J. Phys. Chem. B 109, 19068–19081 (2005).
22. Brueckner, F., Hennecke, U., Carell, T. & Cramer, P. CPD damage recognition by transcribing RNA polymerase II. Science 315, 859–862 (2007).
23. Komissarova, N. et al. Shortening of RNA:DNA hybrid in the elongation complex of RNA polymerase is a prerequisite for transcription termination. Mol. Cell 10, 1151–1162 (2002).
24. Srivastava, D. B. & Darst, S. A. Derepression of bacterial transcription-repair coupling factor is associated with a profound conformational change. J. Mol. Biol. 406, 275–284 (2011).
25. Landick, R. The regulatory roles and mechanism of transcriptional pausing. Biochem. Soc. Trans. 34, 1062–1066 (2006).
26. Manelyte, L. et al. Regulation and rate enhancement during transcription-coupled DNA repair. Mol. Cell 40, 714–724 (2010).
27. Selby, C. P. & Sancar, A. Structure and function of transcription-repair coupling factor. I. Structural domains and binding properties. J. Biol. Chem. 270, 4882–4889 (1995).
28. Deaconescu, A. M., Sevostyanova, A., Artsimovitch, I. & Grigorieff, N. Nucleotide excision repair (NER) machinery recruitment by the transcription-repair coupling factor involves unmasking of a conserved intramolecular interface. Proc. Natl Acad. Sci. USA 109, 3353–3358 (2012).

Supplementary Information is available in the online version of the paper.

Acknowledgements We thank K. Neumann for showing us how to perform global fitting with the Igor software package. K.H. was supported by a PhD scholarship from the Frontières Interdisciplinaires du Vivant Doctoral Program and the Fondation pour la Recherche Médicale. Work in the laboratory of N.J.S. was supported by BBSRC grant BB/003142/1, and work in the laboratory of S.A.D. was supported by National Institutes of Health grant GM073829. This work was also made possible by a EURYI grant, in addition to CNRS and University of Paris Diderot core funding, to T.R.S. Author Contributions K.H., N.J.S. and T.R.S. devised and carried out experiments; S.A.D., K.H., N.J., N.J.S., A.J.S., T.R.S. and S.Z. analysed data; and A.J.S., N.D. and T.R.S. wrote the paper.

Author Information Reprints and permissions information is available at www.nature.com/reprints. The authors declare no competing financial interests. Readers are welcome to comment on the online version of the paper. Correspondence and requests for materials should be addressed to T.R.S. (strick.terence@ym.univ-paris-diderot.fr)