PNAS RESEARCH ARTICLE | BIOCHEMISTRY

The structure and activities of the archaeal transcription termination factor Eta detail vulnerabilities of the transcription elongation complex

Craig J. Marshall[a,1,2], M. Zuhaib Qayyum[b,1,3], Julie E. Walker[a,2], Katsuhiko S. Murakami[b,4], and Thomas J. Santangelo[a,4]

Edited by Tina Henkin, The Ohio State University, Columbus, OH; received May 5, 2022; accepted June 22, 2022

Transcription must be properly regulated to ensure dynamic gene expression underlying growth, development, and response to environmental cues. Regulation is imposed throughout the transcription cycle, and while many efforts have detailed the regulation of transcription initiation and early elongation, the termination phase of transcription also plays critical roles in regulating gene expression. Transcription termination can be driven by only a few proteins in each domain of life. Detailing the mechanism(s) employed provides insight into the vulnerabilities of transcription elongation complexes (TECs) that permit regulated termination to control expression of many genes and operons. Here, we describe the biochemical activities and crystal structure of the superfamily 2 helicase Eta, one of two known factors capable of disrupting archaeal transcription elongation complexes. Eta retains a twin-translocase core domain common to all superfamily 2 helicases and a well-conserved C terminus wherein individual amino acid substitutions can critically abrogate termination activities. Eta variants that perturb ATPase, helicase, single-stranded DNA and double-stranded DNA translocase and termination activities identify key regions of the C terminus of Eta that, when combined with modeling Eta–TEC interactions, provide a structural model of Eta-mediated termination guided in part by structures of Mfd and the bacterial TEC. The susceptibility of TECs to disruption by termination factors that target the upstream surface of RNA polymerase and potentially drive termination through forward translocation and allosteric mechanisms that favor opening of the clamp to release the encapsulated nucleic acids emerges as a common feature of transcription termination mechanisms.

RNA polymerase | transcription termination | Eta | Archaea | superfamily 2 helicase

Significance

Transcription elongation complexes (TECs) are normally extremely stable but remain vulnerable to dissociation by domain-specific transcription termination factors. We report the X-ray crystal structure of the superfamily 2 (SF2) helicase Eta (Euryarchaeal termination activity) lacking the nonessential N-terminal domain, one of only two known archaeal-encoded factors capable of disrupting TECs. Eta retains a twin-translocase core domain common to SF2 helicases and a well-conserved C terminus wherein amino acid substitutions can critically abrogate termination activities. Modeling Eta–TEC interactions and the biochemical results obtained with Eta variants add valuable insights into the functional activities of ubiquitously distributed SF2 helicases, highlight areas of susceptibility of TECs to transcription termination factors, and further support a connection between transcription regulation and DNA repair pathways.

Factors that promote or inhibit efficient transcription initiation are abundantly encoded in most genomes, with many species harboring dozens to hundreds of site-specific DNA binding proteins that can influence the assembly and activities of basal transcription factors and RNA polymerase (RNAP) at promoter sequences (1–6). The abundance of transcription factors that control initiation does not preclude regulation throughout the remainder of the transcription cycle, and an increasing number of factors have been demonstrated to influence postinitiation transcription that are often rate-limiting for gene expression (7–10). DNA-bound proteins and nucleoid or chromatin structures typically hinder the progression of RNAP along the DNA template, slowing elongation and providing regulatory pauses that can be exploited to control the rate of RNA production (11, 12). Many eukaryotic factors can posttranslationally modify RNAP or intimately associate with RNAP to control the rate of elongation and translocation (13, 14), but only a few factors encoded in all of life can reduce the elongation rate to zero and disrupt the normally extremely stable transcription elongation complex (TEC) (15–21). Portions of this work were developed from the doctoral thesis of C.J.M. (22).

The termination phase of the transcription cycle is subject to intricate control mechanisms that take advantage of vulnerabilities to the stability of the TEC. The bacterial and archaeal RNAP, as well as eukaryotic RNAP III (Pol III) and bacteriophage RNAP, are responsive to intrinsic termination sequences wherein DNA sequences encoding weak, rU:dA RNA:DNA hybrids, often in conjunction with hairpin RNA structures, can stall and disrupt TECs (23–26). While intrinsic termination sequences are often encoded downstream of genes and operons, intrinsic termination signals embedded within the 5′ untranslated region and coding sequences of transcripts form the foundation of many riboswitches and attenuation mechanisms (27). While the absence or nonessential nature of known termination factors implies intrinsic termination mechanisms

This article is a PNAS Direct Submission.

[1]C.J.M. and M.Z.Q. contributed equally to this work.

[2]Present address: Watchmaker Genomics, Boulder, CO 80301.

[3]Present address: Protein Technologies Center, Danny Thomas Research Center, Department of Structural Biology, St Jude Children's Research Hospital, Memphis, TN 38105.

[4]To whom correspondence may be addressed. Email: thomas.santangelo@colostate.edu or kum14@psu.edu.

essential nature of known termination factors implies intrinsic termination mechanisms alone may suffice for some bacterial species, most Bacteria, all Archaea, and all Eukarya are dependent on protein factors that can stimulate transcription termination to ensure proper expression of the genome (28). Only a few domain-specific transcription factors

This article contains supporting information online at http://www.pnas.org/lookup/suppl/doi:10.1073/pnas.2207581119/-/DCSupplemental.

Published August 2, 2022.

have been identified that can disrupt TECs to release both RNAP and the RNA transcript from DNA, and with just one exception [e.g., FttA, also termed aCPSF1, and the eukaryotic CPSF73 (29)], there is no cross-domain conservation of any known transcription termination factor. This suggests that in each domain unique proteins have evolved that likely target vulnerabilities of the TECs to tip the energetic balance in favor of TEC disassembly versus continued elongation. Given the known structures and conserved nature of the contacts that stabilize TECs in each domain, it is possible, if not likely, that the known termination factors may be reliant on similar mechanisms to disrupt TECs and control gene expression.

Debate remains regarding the exact mechanism(s) employed by transcription factors to disrupt TECs, but the essentiality of many transcription termination factors underlies the importance of termination factors to control gene expression. Transcription termination factors can be broadly divided into two clades: "RNA-dependent" versus "DNA-dependent." RNA-dependent termination factors [e.g., Rho in bacteria (30) and FttA in Archaea (29)] are typically associated with general governance of TEC activity, terminating TECs that have transcribed to the end of a gene or operon or TECs that fail to engage the translation apparatus. In many prokaryotic species, the uncoupling of transcription and translation provides an RNA binding site and ultimate access to the TEC (29, 31–33). In eukaryotes, RNA cleavage associated with polyadenylation signals provides access to an uncapped 5′ RNA terminus that permits RNA degradation and TEC access (34). Conversely, DNA-dependent termination factors [e.g., Mfd in Bacteria (35), TTF2 in eukaryotes (36), and Eta in Archaea (16)] tend to have more specialized functions, typically acting to disrupt TECs regardless of position or coupling to the translation apparatus to recycle RNAPs irreversibly stalled due to DNA damage or to clear chromosomes prior to condensation or replication (21, 37).

Archaea encode a single multisubunit RNAP that shares substantial structural similarities to eukaryotic RNAP II (Pol II) (38, 39). The archaeal RNAP is sensitive to intrinsic termination signals, and at least in vitro the archaeal RNAP is also sensitive to bacterial rho-mediated termination (40); despite this susceptibility, rho is not encoded in any archaeal genome. Instead, archaeal species universally encode an essential transcription termination factor termed FttA [Factor that terminates transcription in Archaea; also termed aCPSF1 (29)] and most euryarchaeal species encode a second termination factor, Eta [Euryarchaeal termination activity (16)]. FttA is an β-CASP, metallo-β-lactamase RNA-dependent termination factor likely responsible for global transcription termination events that provides the regulation normally afforded by rho in Bacteria and the poly-A dependent transcription termination common in eukaryotes. Eta, in contrast, is a DNA-dependent, superfamily 2 (SF2) helicase/translocase transcription termination factor. SF2 helicases are ubiquitously distributed across the domains with varying yet often indispensable roles in nucleic acid metabolism (41–43). Eta-mediated termination shares similarities with bacterial Mfd-mediated termination, with both termination mechanisms being relatively slow and thus poorly suited to terminate rapidly transcribing TECs. Both Eta and Mfd are SF2 helicases reliant on DNA upstream of TECs to restart stalled or backtracked TECs or terminate TECs that cannot continue elongation due to damage in the template strand of DNA or nucleotide deprivation.

variants. The structure of Eta defines two highly conserved SF2 helicase/translocase domains flanked by an NTD that is dispensable for transcription termination and a globular C terminus harboring conserved residues that are critical to Eta function. Single amino acid changes in the C terminus of Eta suffice to separately disrupt factor-dependent transcription termination, ATPase, translocase, and helicase activities. The combined results provide the basis for a structural model of Eta–TEC interactions and a model of Eta-mediated transcription termination. Disrupting the TEC from the upstream edge of the transcription bubble in both intrinsic and factor-dependent termination suggests that this point of access takes advantage of a vulnerability to TEC stability by promoting forward translocation in the absence of continued RNA synthesis to undermine the strength of the TEC through successive disruptions to the RNA–DNA hybrid and collapse of the transcription bubble.

Results

Eta Is a Canonical SF2 Helicase/Translocase. Eta from *Thermococcus kodakarensis* (Fig. 1*A*) is an 832-amino-acid (aa), ∼96-kDa monomer with 11 readily identified motifs (Q, I, Ia, Ic, II, III, IV, V, Va, Vb, and VI) that define Eta as a member of the DEAH/RHA family of SF2 helicases/translocases (*SI Appendix*, Table S1) (44). Walker A and Walker B motifs that promote ATP and Mg^{2+} binding are also defined, and primary sequence alignments of >100 Eta homologs (*SI Appendix*, Fig. S1), all from euryarchaeal species, reveal a weakly conserved NTD harboring four almost universally conserved cysteines that likely coordinate a metal ion(s). Modeling of the presumptive metal-binding N terminus (aa 1 to 192) predicts an extended alpha-helical structure; the NTD of Eta is dispensable for Eta-mediated transcription termination (16). Full-length Eta is insoluble at high concentrations and deletion of the NTD permitted sufficient concentration of preparations of the remainder of Eta (aa 193 to 832, termed $Eta^{\Delta NTD}$) to establish conditions suitable for crystallization. The X-ray crystal structure was solved by single-wavelength anomalous dispersion at 4.5 A° and then the resolution of the structure was improved to 4.1 A° with a native crystal. Four almost identical (0.179- ∼ 0.219-Å root-mean-square deviations over all residues) $Eta^{\Delta NTD}$ molecules were resolved per asymmetric unit. The $Eta^{\Delta NTD}$ structure reveals a mostly alpha-helical flattened-disk structure ∼75 A° in diameter by ∼40 A° deep defining three approximately equally sized primary domains: aa 193 to 400 comprise the first helicase/translocase domain (TD1, Fig. 1*A*, cyan), aa 401 to 577 define a second helicase/translocase domain (TD2, Fig. 1*A*, yellow), and aa 578 to the C terminus form a completely alpha-helical C-terminal domain (CTD, Fig. 1*A*, green). The two helicase/translocase domains share a near-identical fold with each other and the SF2 conserved sequences (Fig. 1*A*, purple) form the bulk of the interface between TD1 and TD2. Dali-based structural alignments reveal that the twin translocase/helicase domains of Eta overlay nicely with the translocase/helicase domains of many other SF2 helicase/translocases, including the helicase domain of eukaryotic DNA polymerase theta (45), Rad25 (XPB) (46, 47), bacterial Mfd proteins (Fig. 1*C*), and many RNA and DNA helicases. Dali alignments (48) of the Eta CTD reveal some structural conservation with the archaeal Hel308 (49) and the Ski2 RNA-helicase Brr2 (50). Contacts between TD1/TD2 and the

To detail the mechanisms of SF2 helicase/translocase function and DNA-dependent transcription termination, we report the X-ray crystal structure of Eta (lacking the N-terminal domain [NTD]) and a structure-guided functional analysis of Eta

CTD are minor but include a globular region of the CTD (aa 577 to 630) that packs against TD1, an extended loop (residues 295 to 310) from TD1 that contacts the CTD, and the C terminus of an α-helix (aa 747 to 770) within the CTD that contacts

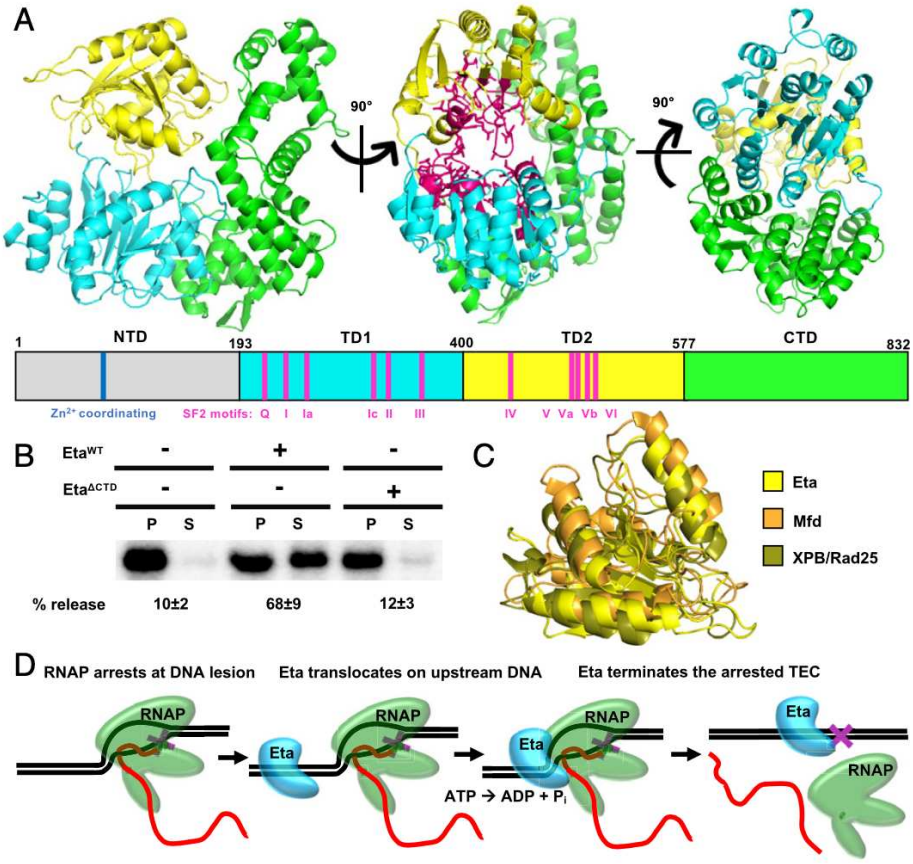

Fig. 1. The crystal structure of Eta$^{\Delta NTD}$ and proposed mechanism of Eta-mediated transcription termination. (*A*) The X-ray crystal structure of Eta$^{\Delta NTD}$ (aa 194 to 832) reveals a three-domain architecture; two RecA-like translocase domains (aa 193 to 577, cyan and yellow), and a helically structured unique CTD (aa 578 to 832, green). Eleven conserved SF2 helicase motifs identified in the primary sequence of Eta are highlighted in magenta. (*B*) The unique CTD of Eta is indispensable for Eta-mediated transcription termination. *P* = pellet, S = supernatant. Percent transcript release is reported as mean ± SD of three replicates. (*C*) Alignment of the TD2 translocase domain of Eta (yellow), the bacterial transcription termination factor Mfd (orange, PDB ID code: 6XEO), and XPB/Rad25 (olive, PDB ID code: 2FWR) reveal near-identical structures. (*D*) Eta-mediated termination necessitates ATP hydrolysis, efficient translocase activities, and correct contacts with RNAP to destabilize stalled TECs. The magenta X represents a transcription roadblock, i.e., template-strand DNA damage.

TD2 around aa 473. Minimal contact points between TD1/TD2 and the CTD suggests potential rearrangements of the domains upon RNAP/TEC and/or DNA binding during the RNAP displacement process, as observed in other RNAP binding helicases such as HelD and Mfd (15, 51). The primary sequences alignments of Eta to other SF2 members predicted a winged-helix domain (WHD) resolved within the CTD (starting at aa 626), whose characterization in other SF2 members was predicted to be important for proper activity (41).

Eta-Mediated Transcription Termination Requires the CTD of Eta. Eta-mediated transcription termination is energy-dependent (dATP and ATP suffice equivalently in vitro) and requires access to DNA upstream of the TEC (16). These activities, and retention of other hallmarks of Mfd-mediated termination in Bacteria, including rescue of backtracked complexes and the ability to terminate slowly elongating or static TECs, suggested a model of Eta-mediated termination wherein Eta binds DNA upstream of the TEC then translocates along double-stranded DNA (dsDNA), collapsing the transcription bubble and forcing RNAP to translocate forward in the absence of continued RNA synthesis (Fig. 1*D*). Retention or deletion of the NTD did not in contrast, is essential for Eta-mediated termination (Fig. 1*B*), implying that the CTD may be the interaction surface with the archaeal RNAP during factor-dependent transcription termination. Deletion of the Eta CTD did not impact ATPase activities (*SI Appendix*, Fig. S4), further defining the Walker motifs and translocase/helicase activities to the twin TDs.

Conserved, Solvent-Exposed Residues in the CTD of Eta Are Critical for Helicase, ATPase, Single-Stranded DNA (ssDNA) and dsDNA Translocase, and Transcription Termination Activities. We further investigated the function of the CTD for Eta-mediated transcription termination and its helicase/translocase activities. We aligned the ~100 closest homologs of *T. kodakarensis* Eta to identify residues that were highly (>90%) conserved within the CTD that were solvent-exposed based on the Eta$^{\Delta NTD}$ crystal structure determined in this study (*SI Appendix*, Fig. S1). Thirty-seven highly conserved residues were identified within the CTD, with one-third (Q588, G644, T641, F651, F705, C741, H774, R804, and E826) solvent-exposed (Fig. 2*A*) and likely involved in RNAP and/or DNA interactions during transcription termination. To determine the functional impacts of these conserved residues on Eta activity, we generated single-

impact the efficiency of Eta-mediated transcription termination, suggesting that the NTD may play a role in coordinating Eta-mediated transcription termination with other proteins and processes, perhaps linked to DNA repair (16, 52). The Eta CTD, amino-acid variants of Eta wherein nonglycine residues were substituted with alanine and G644 was substituted for an aspartic acid (Fig. 2*B*). We also generated a variant wherein the WHD was deleted (ΔWHD) and purified each variant to

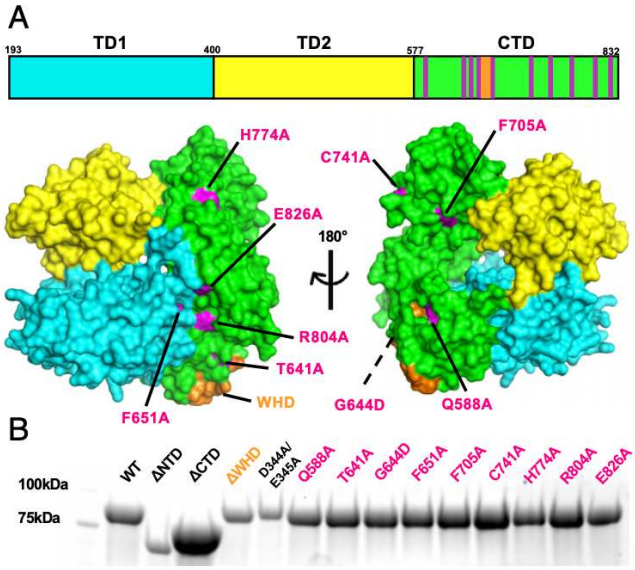

Fig. 2. Conserved and likely solvent-exposed residues and subdomains of Eta targeted for mutagenesis. (*A*) Surface modeling of the Eta$^{\Delta NTD}$ crystal structure allows identification of conserved and likely solvent-exposed residues in the CTD of Eta (green). Identified residues (magenta), a putative WHD (orange), and the Walker B domain (D344, E345) were perturbed via site-directed mutagenesis, recombinantly expressed, and purified (*B*). Molecular weight standards are indicated to the left.

homogeneity for use in in vitro helicase, translocase, ATPase, and transcription termination assays. Each variant was generated in full-length Eta, remained stably folded at 85°C, and was recombinantly purified under identical procedures to Eta^{WT}.

Most Eta CTD variants retain translocase activities on both ssDNA and dsDNA. Eta-mediated termination relies on DNA translocase activity to terminate stalled TECs (16). While the structure of Eta predicts translocase/helicase activities to be contained within the TDs, it remained plausible that specific amino acid substitutions within the CTD might impact helicase or translocase activities through domain cross-talk or local changes to folding. To ensure our purified Eta variants retained translocase activity, the ability of Eta and Eta variants to displace streptavidin bound to an internally biotinylated and radiolabeled DNA template was investigated in both an ssDNA and a dsDNA context (Fig. 3*A*). Eta^{WT} translocating on ssDNA displaced ∼90% of streptavidin bound to an internal biotinylated base during the 8-min time course (Fig. 3*B*). To ensure that once streptavidin was displaced it would not simply rebind the biotinylated DNA substrate, reactions were carried out in a vast excess of free biotin. Biotinylated ssDNA bound streptavidin displacement curves generated using Eta^{WT} and Eta variants were fit to a Michaelis-Menten-like curve (average R^2 = 0.986), and maximum reaction rates were determined for each Eta variant (Fig. 3*D*, gray bars). Most Eta variants (Eta^{Q588A}, Eta^{T641A}, Eta^{F705A}, Eta^{C741A}, Eta^{R804A}, and Eta^{E826A}) retained wild-type (WT)- or near-WT-like activity with normalized (WT = 1.0) maximum reaction rates >0.8. Eta^{F651A} and Eta^{H774A} displayed modest disruptions to ssDNA translocase activity with normalized maximum reaction rates of 0.63 and 0.73, respectively. Eta$^{\Delta WHD}$, Eta$^{\Delta CTD}$, and Eta^{G644D} had severely hindered ssDNA translocase activity displaying maximum rates of streptavidin displacement less than a transcription termination, and it is plausible dsDNA translocation is achieved through an alternative mechanism. Thus, we performed streptavidin displacement assays for each Eta variant on dsDNA substrates with one radiolabeled strand to examine streptavidin release via native polyacrylamide gel electrophoresis (PAGE) (Fig. 3*C*). Maximum reaction rates were also generated after fitting streptavidin-displacement data to a Michaelis–Menten-like curve (average R^2 = 0.988) and normalized to Eta^{WT} (Fig. 3*D*, orange bars). Most variants (Eta^{Q588A}, Eta^{T641A}, Eta^{F705A}, Eta^{C741A}, Eta^{H774A}, Eta^{R804A}, Eta^{E826A}, and Eta$^{\Delta WHD}$) had comparable translocase activity rates on dsDNA and ssDNA. Interestingly, Eta^{G644D} and Eta$^{\Delta CTD}$ displayed significantly improved translocase activity on the dsDNA substrate, indicating some regions of the CTD of Eta are instrumental in establishing the correct substrate selection for Eta function. The observed dsDNA translocase activity of Eta$^{\Delta CTD}$, which is unable to collapse stalled TECs (Fig. 1*B*), reinforces that translocase activity is required but not sufficient for Eta-mediated transcription termination. Overall, the ssDNA and dsDNA translocase assays suggest it is likely that Eta CTD variants, with the exception of Eta^{H774A} and Eta$^{\Delta WHD}$, retain motor activities required during Eta-mediated termination.

Most Eta CTD variants retain helicase activities. Eta retains 3′ to 5′ helicase in vitro and it is plausible that helicase activity is required for Eta-mediated termination (53). Translocase assays alone do not assess strand-separation activity and thus, to ensure our purified Eta variants retained helicase activity, we tested the ability of each variant to unwind dsDNAs with a 3′ overhang (Fig. 4*A*). Eta^{WT} directed the complete, energy-dependent release of a small, radiolabeled oligonucleotide from dsDNA substrates (Fig. 4*B*) within just a few minutes. A DNA oligo completely complementary to the radiolabeled short oligonucleotide was added to ensure the radiolabeled strand did not reanneal to reform the original substrate; Eta cannot unwind blunt-ended dsDNA substrates (53). With few exceptions (Eta^{H774A}, Eta^{R804A}, and Eta$^{\Delta WHD}$), the purified Eta CTD variants (Fig. 4 *B* and *C* and *SI Appendix*, Fig. S3) revealed no deficiencies in helicase activities. As anticipated, disruption of the Walker B motif completely abrogates motor activity (Fig. 4). Eta^{R804A} and Eta$^{\Delta WHD}$ display slower helicase activities, only unwinding ∼60 to 80% of dsDNA substrates in the same time required for Eta^{WT} to completely unwind the substrates; the kinetic trajectory of Eta^{R804A} and Eta$^{\Delta WHD}$ helicase activities suggests these enzymes are simply slower. In contrast, the activity of Eta^{H774A} suggests that this amino acid substitution either perturbs local folding dynamics congruent with proper helicase activity or lacks the proper prerequisite ATPase activity. Importantly, most Eta mutants retained motor activities, suggesting any defects in transcription termination activities are not due to immobility of the enzyme (Table 1).

Eta CTD variants retain ATPase activities. Retained helicase and translocase activities in the bulk of the Eta variants suggested that each variant also retained robust ATPase activities; however, we nonetheless wanted to test if amino acid substitutions introduced into our Eta variants altered ATPase activity. ATPase activity of Eta and Eta variants is dependent on a nucleic acid substrate (*SI Appendix*, Fig. S4) and disrupting the Walker B motif (D344A/E345A) abrogates ATPase activity, as expected. Most Eta CTD substitutions have no significant impact on ATPase activity. Despite slower helicase activity, Eta^{R804A} consistently displays better than Eta^{WT} ATPase activities, whereas the Eta$^{\Delta WHD}$ variant displays just

quarter of Eta^{WT}. The broad range of disruptions to ssDNA translocase activity of Eta CTD variants suggests cross-talk between TD1 and TD2 of Eta with the CTD. However, it is likely that translocation on a dsDNA substrate is required for Eta-mediated ~60% of Eta^{WT} activity. Congruent with the weak helicase activity, Eta^{H774A} displays just ~25% of the Eta^{WT} ATPase activities. Further, while variations to some individual amino acids within the CTD of Eta had severe consequences for ATPase activity, $Eta^{\Delta CTD}$

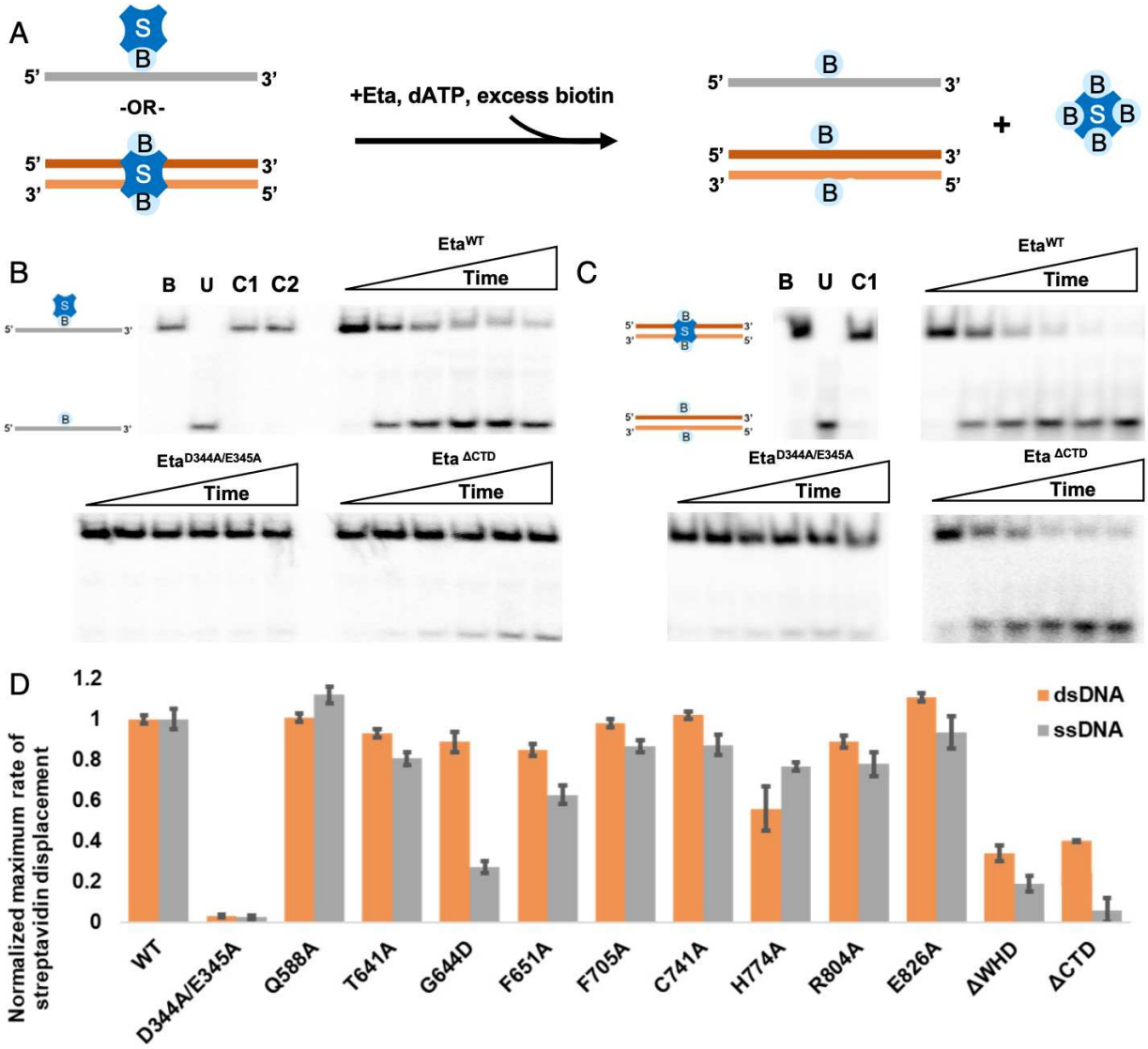

Fig. 3. ATP hydrolysis and the CTD of Eta are necessary for effective translocase activity. (*A*) ssDNA (gray) with a single, internal biotinylated base (light blue) or dsDNA (orange and red) wherein each strand contains a single, internal biotinylated base (light blue) are bound by streptavidin (dark blue) to produce substrates for translocase assays. ATP-dependent translocase activity displaces streptavidin from the DNA substrate(s) where excess free biotin traps streptavidin displaced due to Eta translocation from rebinding the DNA. (*B*) Populations of radiolabeled streptavidin bound (B) and unbound (U) ssDNA resolved by native PAGE. Reactions lacking dATP (C1) or Eta (C2) display no intrinsic release of streptavidin, whereas enzymatic release of streptavidin by Eta^{WT} and Eta variant translocation along ssDNA can be tracked over time. Substitutions withing the Walker B motif or deletion of the CTD of Eta block and significantly reduce, respectively, ssDNA translocation. (*C*) Populations of radiolabeled streptavidin bound (B) and unbound (U) dsDNA resolved by native PAGE. Reactions lacking dATP (C1) display no intrinsic release of streptavidin from dsDNA, whereas enzymatic release of streptavidin by Eta^{WT} and Eta variant translocation along dsDNA can be tracked over time. Substitutions within the Walker B motif or deletion of the CTD of Eta block and significantly reduce ssDNA translocation. (*D*) Quantification of maximum reaction rates of Eta and Eta variant translocase activity on ssDNA (gray) and dsDNA (orange). Error bars represent one SD from the mean ($n = 3$).

displayed only a very slight perturbation to ATPase activity (~90% of Eta^{WT}). This is suggestive of a role of the CTD in coordinating TD1 and TD2 into correct conformations for DNA-dependent ATP hydrolysis.

Substitutions to conserved and solvent-exposed residues in the CTD of Eta abrogate transcription termination activity. Stalled TECs are easily formed via promoter-directed initiation and selective nucleotide deprivation on solid supports to provide an ideal substrate to monitor the rate and efficiency of Eta-mediated transcription termination (Fig. 5*A*) (16). The archaeal basal transcription allows formation of promoter proximal stalled TECs. Magnetic separations collect intact TECs with 58 nucleotide radiolabeled transcripts ($TECs_{+58}$) within a pellet (P) fraction, and in the absence of any protein additions the extreme stability of TECs results in only minor amounts of nascent transcript release to solution (S) even after extended incubation at 85 °C. While addition of Eta^{WT} to stalled TECs results in ~80% efficient transcription termination (Fig. 5*B*), many Eta CTD variants nearly completely lose the ability to terminate transcription or display reduced termination efficiencies (Fig. 5 *C* and *D* and *SI*

tion factors TFB and TBP suffice to permit the archaeal RNAP to recognize BRE and TATA promoter elements, respectively, and initiate transcription with just three of the four NTPs (ATP, UTP, and CTP); elongation on templates with a G-less cassette

Appendix, Fig. S5). As anticipated, Eta variants with poor motor and ATPase activities (Eta^{H774A} and $Eta^{\Delta WHD}$) also display poor transcription termination activity. Substitutions at the far ends of the CTD (Eta^{Q588A} and Eta^{E826A}) did not

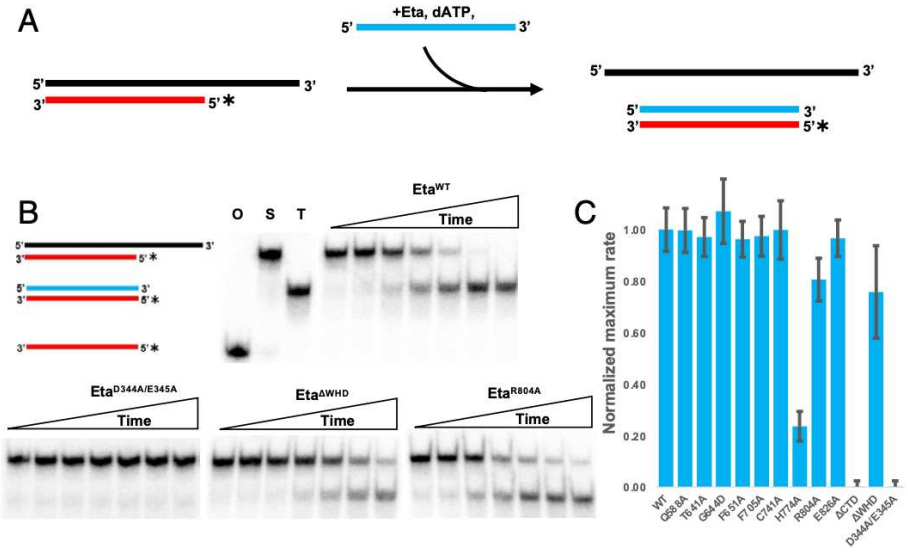

Fig. 4. The CTD of Eta and ATP hydrolysis are necessary for 3′ to 5′ helicase activities of Eta. (*A*) A short, radiolabeled DNA strand (red; asterisk denotes 5′-^{32}P labeling), initially paired to a longer complementary strand (black) generating a 3′-overhung dsDNA substrate, is separated by Eta in an ATP-dependent manner and anneals to a complementary trap DNA (blue) provided in excess. (*B*) The radiolabeled short oligo (O) can be electrophoretically resolved from the substrate DNA complex (S) and the trap DNA complex (T) to reveal strand separation by Eta or Eta variants to monitor helicase activity efficiency. Deletion of the CTD or substitutions of the Walker B motif abolish 3′–5′ helicase activity. (*C*) Quantification of Eta and Eta variant helicase activity. Error bars represent one SD from the mean (n = 3).

affect transcription termination efficiency; in the case of Eta^{E826A} a reproducibly modest increase in transcription termination activity is observed. Importantly, many substitutions that have minimal impact on helicase, dsDNA translocase, or ATPase activities were identified that detrimentally affect transcription termination activity. Two variants, Eta^{G644D} and Eta^{R804A}, are of particular interest as both result in almost no transcription termination activity while retaining high levels of dsDNA translocase, helicase, and ATPase activity. Eta^{T641A}, Eta^{F651A}, Eta^{F705A}, and Eta^{C741A} are also of interest as termination efficiencies decreased to just ∼50 to 70% of Eta^{WT}. Decreases in ssDNA translocation activities often correlated with decreases in termination efficiency (e.g., Eta^{T641A}, Eta^{G644D}, Eta^{F651A}, and Eta^{C741A}), implying that ssDNA translocase activities are likely necessary for Eta-mediated termination; however, Eta^{R804A} and Eta^{F705A} display more dramatic decreases in termination efficiency than losses in ssDNA translocase activities, suggesting that while necessary ssDNA translocase activity cannot suffice to drive transcription termination without retention of specific Eta–TEC interactions that may be impaired by select Eta variants.

Direct interactions between Eta and the TEC are anticipated based on our modeling and the ability of Eta to disrupt TECs; however, quantifying and comparing the stability of interactions between Eta or Eta variants with TECs has not been experimentally possible despite exhaustive attempts at pull-down or direct binding assays. Given these limitations, we used the more-readily detected movements and activities of TECs to monitor the impact of amino acid substitutions within the Eta CTD on putative Eta–TEC interactions (Fig. 6). TECs prepared by nucleotide deprivation cannot accurately transcribe further along the DNA template than substrate availability permits, but TECs are free to undergo retrograde movement, often coincident with endonucleolytic cleavage of RNA phosphodiester linkages that shorten nascent transcripts. Retrograde movement, termed backtracking, is stimulated from https://www.pnas.org by COLORADO STATE UNIVERSITY LIBRARIES; SERIALS SECTION on June 9, 2025 from IP address 129.82.95.95.

proficient TECs$_{+58}$, backtracked TECs$_{+58}$ with internal phosphodiester linkages occupying the active center, and TECs$_{+\sim50–57}$ that could either be backtracked or catalytically proficient but with shortened transcripts. Eta, like Mfd, is known not only to terminate TECs but also to influence the propensity of TECs to backtrack, with Eta and Mfd both capable of rescuing backtracked complexes to catalytically competent conformations.

Table 1. Summary of ATPase, helicase, translocase, and transcription termination activities of Eta and Eta variants normalized to Eta^{WT} = 1.0

Variant	Termination activity	Helicase activity	ATPase activity
WT	59.6%	100%	0.72
	1.00X	1.00X	1.00X
Q588A	60.1%	100%	0.87
	1.01X	1.00X	1.28X
T641A	46.2%	100%	0.80
	0.78X	1.00X	1.15X
G644D	14.7%	97.4%	0.79
	0.25X	0.97X	1.12X
F651A	41.8%	100%	0.72
	0.70X	1.00X	1.01X
F705A	31.6%	100%	0.76
	0.53X	1.00X	1.08X
C741A	45.4%	97.8%	0.74
	0.76X	0.98X[a]	1.04X
H774A	7.3%	27.6%	0.33
	0.12X	0.28X	0.31X
R804A	8.1%	78.9%	1.01
	0.14X	0.79X	1.53X
E826A	69.1%	100%	0.73

lated by extended incubations at physiological temperatures. When limited NTP subsets are available, a population of $TECs_{+58}$ is not static but is most accurately described as a dynamic and interchanging population of forward translocated, catalytically

	1.16X	1.00X	1.02X
ΔWHD	34.7%	65.0%	0.54
	0.58X	0.65X	0.67X

[a]Color gradients highlight deviations in protein activities compared to Eta^{WT}.

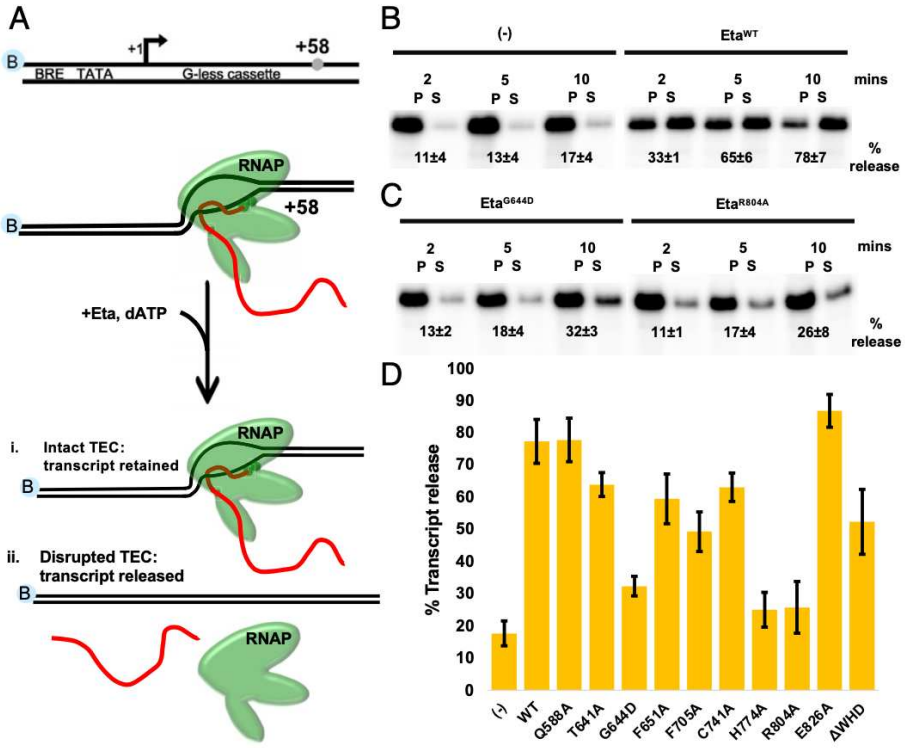

Fig. 5. Conserved CTD residues are required for efficient Eta-mediated transcription termination. (*A*) A terminally biotinylated (light blue) dsDNA (black) encoding a strong archaeal promoter (BRE, TATA box, and transcription initiation site +1 are shown) followed by a G-less cassette (shown) permits transcription initiation with TFB, TBP, and RNAP. Selective NTP availability stalls promoter-initiated TECs after incorporation of 58 nucleotides generating $TECs_{+58}$. Addition of Eta, followed by separation of intact versus disrupted TECs through capture of the biotinylated dsDNA on a streptavidin-based solid support permits quantification of nascent transcript release (S = supernatant) or retention (*P* = pellet) from $TECs_{+58}$. (*B*) Transcription termination assays reveal the time-dependent release of nascent transcripts upon addition of Eta^{WT}. Percent transcript release is reported as mean ± SD of three replicates. (*C*) Transcription termination activities are significantly reduced when single amino acid variants Eta^{G644D} and Eta^{R804A} are compared to Eta^{WT}. Percent transcript release is reported as mean ± SD of three replicates. (*D*) Quantification and comparison of transcription termination activities of Eta and Eta variants. Error bars represent one SD from the mean ($n = 3$).

We tested whether several Eta variants that lacked robust termination functions could influence TEC activities beyond termination. The capacity to alter the dynamics of a population of TECs is suggestive of retained engagements with TECs when DNA-bound, whereas minimal changes to the dynamics of a population of stalled TECs is suggestive of impaired TEC interactions when Eta variants are known to retain ATPase, helicase, and dsDNA translocation activities similar to Eta^{WT}. When $TECs_{+58}$ are first generated, captured to a solid support, and washed to remove NTPs at reduced temperatures, nearly all (≥98%) transcripts are retained within TECs and no obvious reduction of nascent RNA length occurs (Fig. 6, lanes 1 to 4). When $TECs_{+58}$ are shifted to and incubated at 85 °C in the absence of any NTPs, most backtrack, and ∼75% show evidence of transcript cleavage (Fig. 6, lanes 7 and 8); the addition of even low concentrations of ATP, UTP, and CTP allows cycles of backtracking and resynthesis that maintain a much more uniform TEC_{+58} population (Fig. 6, lanes 5 and 6). The addition of Eta^{WT} alone (Fig. 6, lanes 11 and 12) does not influence population dynamics, whereas the addition of Eta^{WT} with an energy source exclusive to Eta (e.g., dATP) results in release of most transcripts to solution, and importantly, the number of TECs that backtrack is reduced from ∼70% to just ∼30% (Fig. 6, lanes 13 to 16). Loss of $TECs_{\sim+50–57}$ and coincident gains in $TECs_{+58}$ implies that proper engagement of Eta^{WT} with TECs TEC activities beyond termination, Eta^{R804A}, Eta^{F651A}, and Eta^{G644D}, each of which retains strong ATPase, helicase, and dsDNA translocase activities (Table 1), all fail not only to elicit efficient termination of TECs but also fail to significantly rescue backtracked complexes or influence TEC population dynamics (Fig. 6, lanes 17 to 28). While Eta^{WT} rescues ∼40% of backtracked complexes, Eta^{R804A}, Eta^{F651A}, and Eta^{G644D} rescue just 8%, 18%, and 15% of backtracked TECs. The most parsimonious explanation for the combined reductions in termination activity and loss of forward translocation of TECs is that although these Eta variants are likely fully capable of engaging DNA, these variants cannot correctly engage TECs, either by blocking direct Eta–RNAP interactions or altering Eta dynamics such that engagement of TECs is impaired.

Discussion

TECs are necessarily stable; many functional transcriptional units are thousands to millions of base pairs in length and must be processively transcribed. The requirement for regulated transcription termination ensures that TECs retain structural and energetic vulnerabilities that can be exploited by intrinsic and factor-dependent termination mechanisms. Eta-mediated termination provides an ideal model in Archaea for understanding TEC susceptibility and the activities of SF2 helicases. Determi-

both stimulated catalytically competent conformations of many TECs that permit extension of previously shortened transcripts and an overall reduction in backtracking that limits the potential to shorten transcripts. While Eta^{WT} clearly influences

nation of the X-ray crystal structure of $Eta^{\Delta N1D}$, combined with construction and analyses of more than a dozen Eta variants, ranging from large deletions to individual amino acid substitutions, defines the essentiality of the CTD for Eta-mediated

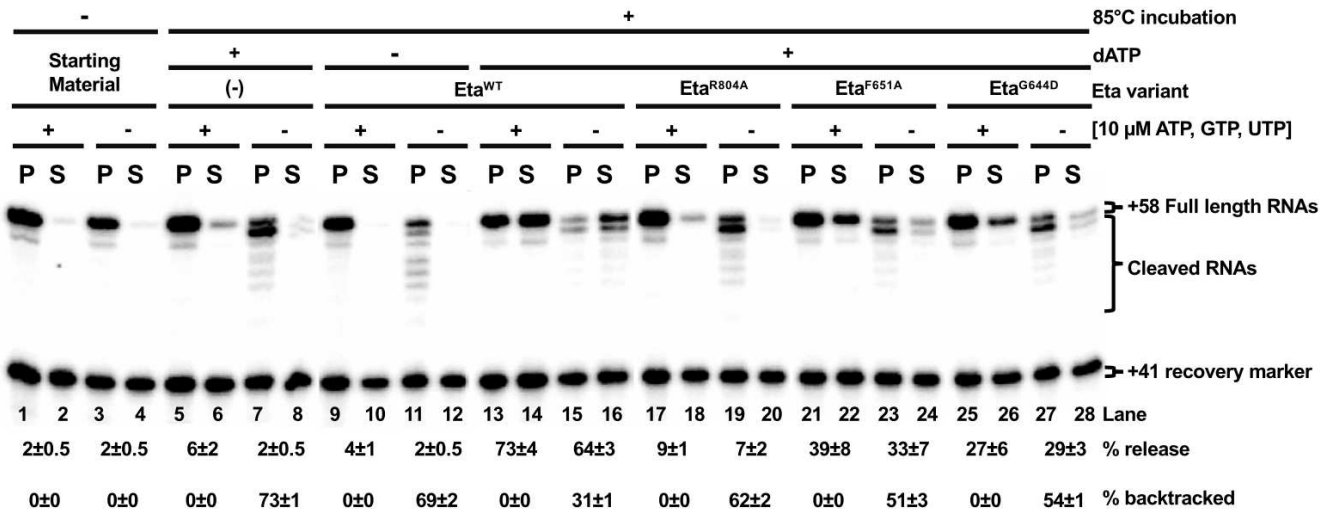

Fig. 6. Eta CTD variants cannot properly engage TECs. Stable $TECs_{+58}$ were assembled and incubated in the presence (lanes 1 and 2) or absence (lanes 3 and 4) of 10 μM ATP, CTP, and UTP which permit resynthesis to +58 following backtracking and transcript cleavage. Following extended incubation at 85 °C, $TECs_{+58}$ incubated in the presence of NTPs backtrack and likely cleave nascent transcripts but can immediately resynthesize RNA, producing RNA products that resolve at +58 (lanes 5 and 6). In contrast, following extended incubation at 85 °C, most (~73%) $TECs_{+58}$ lacking NTPs backtrack but cannot resynthesized shortened RNA products, resulting in multiple RNA species less than 58 nucleotides in length (lanes 7 and 8). Addition of Eta^{WT} without an energy source has no effect on the stability or backtracking of $TECs_{+58}$ (lanes 9 to 12), whereas addition of Eta^{WT} and dATP rescues and promotes nucleotide addition by backtracked TECs, allowing many TECs to resynthesize to +58 before being terminated (lanes 13 to 16). Rescue of backtracked TECs and TEC dissociation is severely abrogated when single-amino-acid CTD Eta variants (G644D, F651A, and R804A) are employed which retain ATPase, translocase, and helicase activities (lanes 17 to 28) adumbrating impaired TEC–Eta engagement.

transcription termination and highlights specific regions of Eta that play roles in ATPase, helicase, translocase, and termination activities (Table 1). The impacts of single-amino-acid substitutions on multiple activities of Eta imply cross-talk between the discrete domains resolved in the crystal structure that are critical to coordinate translocase activity into displacement of the transcription complex and transcription termination.

Eta has been implicated in protein networks likely involved in DNA repair (16, 52) and translation and RNA metabolism (54, 55), and thus it was anticipated that not all conserved residues in the CTD of Eta would be directly related to Eta-mediated transcription termination (Table 1). The NTD of Eta is dispensable for Eta-mediated termination (16) and may also have roles in stimulating additional activities. Modeling of the NTD alone, or as part of full-length Eta, supports an extended structure that may facilitate additional protein contacts that stimulate interactions other than those necessary and sufficient for transcription termination activity (56, 57). Substitutions to the near-universally conserved residues Q588 and E826, located at the extremities of the Eta CTD, resulted in no significant changes to ATPase, helicase, translocase, or transcription termination activities and likely contribute to Eta structure and function outside of termination activity. In contrast, Eta variants H774A and ΔWHD were severely compromised in all tested activities; given the thermostability of these variants, it is likely that they retained a folded structure but that substitutions to the CTD impacted the necessary coordination between the TDs and the CTD to convert (d)ATP hydrolysis into functional work. H774 is close to the twin-translocase domains of Eta and given the significant (~80%) loss of ATPase activity is likely important for correct domain packing to align the Walker motifs. The WHD of Hel308, another well-studied archaeal SF2 helicase, is required for proper DNA interactions (41). Given that DNA binding is essential for Eta ATPase activity (*SI Appendix*, Fig. S4), DNA in the absence of the WHD. Several highly conserved residues in the CTD of Eta are critical for Eta-mediated transcription termination, with little to no impact on ATPase, dsDNA translocase, and helicase activity: substitution of T641, F651, and C741 significantly reduced termination activity, while substitution of G644, F651 and R804 nearly or completely abolished transcription termination activity. For most of these variants, notable decreases in ssDNA translocase activities correlate with loses in termination efficiency, highlighting the likely requirement for ssDNA translocase activity for proper transcription termination. Select variants, including Eta^{G644D} and Eta^{R804A}, behaved, in termination assays, similarly to a complete deletion of the CTD, once again highlighting the necessity of the CTD for Eta-mediated termination and adumbrating direct contact of the TEC and Eta CTD.

The conservation of the TDs permits structural modeling based on other SF2 helicases with respect to the TEC. Modeling of $Eta^{\Delta NTD}$–TEC interactions using the cryogenic electron microscopy (cryo-EM) structures of *T. kodakarensis* RNAP (58) and bacterial Mfd–TEC complexes (15) provides a plausible arrangement of Eta and the archaeal TEC (Fig. 7). When the twin TDs of Eta and Mfd are superimposed, the C terminus of Eta conflicts with the path of the upstream DNA resolved in the Mfd–TEC structure, implying either that the essential C terminus of Eta is likely to rearrange upon TEC engagement or that Eta–DNA interactions generate a new path for the DNA that does not require CTD rearrangements. When bound to upstream DNA in the same manner as Mfd, Eta is positioned such that the highly conserved, functionally critical CTD could contact the β-protrusion domain of RNAP, in many ways analogous to the primary point of contact between Mfd and the bacterial TEC (59). The weak interactions supporting Eta–TEC engagements have beguiled imaging a cocomplex, but several Eta variants that retain most tested activities except termination

the ~30% reduction in ATPase activity and ~80% reduction in translocase activities in the $Eta^{\Delta WHD}$ variant suggests the WHD contributes to coordinating ATP-dependent translocase activities and that energy release is poorly coupled to motor activity on activity are also limited in altering backtracking (Fig. 6), implying a defect in Eta–TEC interactions.

While termination-deficient variants (T641A, G644D, F651A, F705A, C741A, and R804A) are dispersed in linear

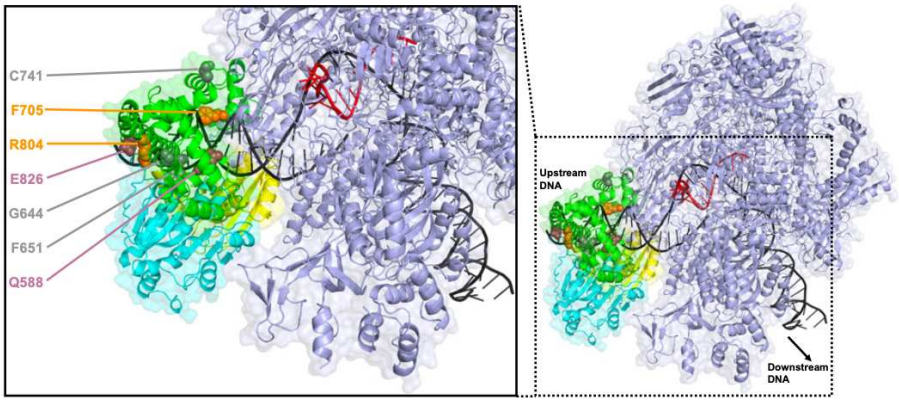

Fig. 7. Model of Eta–TEC engagement. The high conservation of the core of bacterial and archaeal RNAPs and the conserved structures of the SF2 translocase domains of Mfd and Eta allow modeling of the Eta–TEC complex using the previously reported Mfd–TEC structure. The TEC is oriented such that the direction of transcription is to the right (black arrow) with upstream and downstream DNA sequences identified. Conserved, solvent-exposed Eta residues substituted and investigated here are highlighted (raspberry, Q588 and E826, result in minimal impact on Eta function; gray, G644, F651, and C741, result in reduced ssDNA translocase and termination efficiencies; orange, F705 and R804, result in termination specific deficiencies) cluster near the β-protrusion of RNAP. Eta TD1 (cyan), TD2 (yellow), and CTD (green); RNAP (slate green); DNA (light and dark gray); RNA (red).

sequence, folding of the CTD positions clusters these residues such that several could plausibly interact with the TEC (Fig. 7). The mode of TEC engagement with Mfd, and that modeled for Eta, suggests that both termination factors drive RNAP forward using their motor activities to initiate transcription termination. Coordinating Eta-mediated ATP hydrolysis to movement and ultimately dissociation of the TEC likely involves not only direct contacts between Eta and RNAP but also intramolecular movements within Eta that are potentially hindered by substitutions that reduce, but do not eliminate, Eta-mediated transcription termination.

In Bacteria, Mfd acts as the transcription repair coupling factor and terminates TECs stalled at bulky DNA lesions while concurrently recruiting bacterial nucleotide excision repair (NER) enzymes to the site of damage. NER or the transcription-coupled repair (TCR) subpathway have not been explicitly demonstrated in Archaea, although tantalizing evidence implies NER may be active in diverse archaeal clades (16, 52, 60–62). Eta-mediated transcription termination resembles Mfd-mediated termination, and deletion of both factors from their respective organisms results in sensitivity of cells to ultraviolet-induced DNA damage (16, 63). However, continued efforts will be necessary to detail whether Eta-mediated transcription termination is linked to DNA repair and to identify the factors that may direct archaeal NER and TCR.

Materials and Methods

X-Ray Crystal Structure Determination of Eta$^{\Delta NTD}$. Eta$^{\Delta NTD}$ crystals were prepared by hanging-drop vapor diffusion at 22 °C by mixing equal volumes of Eta$^{\Delta NTD}$ (~8 mg/mL) and crystallization solution (0.1 M Tris·HCl [pH 8.0], 4.5 M NaCl, and ~3% ethylene glycol). Crystals were plunge-frozen in liquid nitrogen. To obtain experimental phase information, native crystals were soaked in crystallization solution containing 0.1 mM hexatantalum bromide cluster (Ta_6Br_{12}) for 4 h before flash freezing in liquid nitrogen. Diffraction data were collected at the F1 beamline of the Cornell University High Energy Synchrotron Source (CHESS, Ithaca, NY) at 100 K and data were processed by HKL2000 (64). Both native and Ta_6Br_{12}-containing crystals belonged to space group $P3_2$ and contained four molecules per asymmetric unit. With both isomorphous and anomalous signals from native and Ta_6Br_{12} datasets, 11 Ta_6Br_{12} sites in the asymmetric unit were located and the experimental phase (figure of used to predict a model of Eta that was used as a reference for the model building and refinement. The final model contains residues from 194 to the C terminus. Final coordinates and structure factors have been deposited in the Protein Data Bank (PDB) with ID code 7UJB.

Modeling the Tko RNAP Elongation Complex with Eta$^{\Delta NTD}$. We modeled the *T. kodakarensis* RNAP elongation complex with Eta$^{\Delta NTD}$ using the cryo-EM structure of *T. kodakarensis* RNAP (PDB ID code: 6KF3) (58), the cryo-EM structure of the *E. coli* RNAP elongation complex with Mfd (PDB ID code: 6X 2N) (15), and the crystal structure of Eta$^{\Delta NTD}$ determined in this study. *T. kodakarensis* and *E. coli* RNAPs were superimposed using their catalytic domains including double-psi beta barrel domains in their largest (RpoA1/β′) and second-largest subunits (RpoB/β). Eta$^{\Delta NTD}$ was aligned with Mfd using the translocase domains (TD1 and TD2).

Multiple Sequence Alignment and Identification of Conserved Amino Acids within the CTD of Eta. The amino acid sequence for TK0566, encoding Eta in *T. kodakarensis*, was queried using the blastp suite and aligned using COBALT (67) versus the 100 top matches from the blastp query (https://www.ncbi.nlm.nih.gov/). Individual residues in the CTD (aa 577 to 832) were scored for conservation based on retention frequency. Residues conserved in >90% of sequences were deemed conserved. Surface modeling of the Eta CTD in the crystal structure identified conserved residues likely to be solvent-exposed for mutagenesis.

Site-Directed Mutagenesis. Site-directed mutagenesis was performed on plasmid pTS481 (16) with the QuikChange II XL kit (Agilent Technologies). Conserved and like solvent-exposed residues were replaced with an amino acid predicted to disrupt the normal function of the residue (Q588A, T641A, G644D, F651A, F705A, C741A, H774A, R804A, and E826A). A small deletion of the WHD (aa 630 to 645) was generated in the same manner.

Protein Purifications. RNAP (RpoL-HA-6xHis) was purified from *T. kodakarensis* strain TS413 as previously described (68). Eta^{WT}, Eta$^{\Delta NTD}$, Eta$^{D344A/E345A}$, Eta$^{\Delta CTD}$, Eta$^{\Delta WHD}$, Eta^{Q588A}, Eta^{T641A}, Eta^{G644D}, Eta^{F651A}, Eta^{F705A}, Eta^{C741A}, Eta^{H774A}, Eta^{R804A}, and Eta^{E826A} were expressed and purified from Rosetta 2 (DE3) cells cultured in Luria-Bertani broth supplemented with 3% D-sorbitol, 100 μg/mL ampicillin, and 34 μg/mL chloramphenicol. Expression was induced by addition of 0.2 mM isopropyl β-D-1-thiogalactopyranoside, and cultures were grown overnight at 22 °C. Biomass was pelleted and initially lysed via sonication in 15 mM Tris·HCl, pH 8.0, 10 mM $MgCl_2$, 100 mM NaCl, and 2 mM β-mercaptoethanol. Material was centrifuged at 22,000 × *g* for 15 min at 4 °C to produce a clarified lysate and debris pellet. The debris pellet was lysed again in 15 mM Tris·HCl, pH 8.0, 10 mM $MgCl_2$, 500 mM NaCl, and 2 mM β-mercaptoethanol and reclarified as described above. The supernatant from this reclarification in higher salt contained

merit: 0.480) was calculated using Automated structure solution (AutoSol) in PHENIX (65). Density modification by Automated model building (AutoBuild) in PHENIX yielded a map which could be used for manual model building using Coot (66), and the structure was refined using PHENIX. AlphaFold (56) was

Eta and was heat-treated at 85 °C for 30 min. The high-salt heat-treated lysate was again clarified by centrifugation at 22,000 × *g* for 15 min at 4 °C, and the resulting supernatant was dialyzed into 15 mM Tris·HCl, pH 8.0, 10 mM $MgCl_2$, 300 mM NaCl, and 2 mM β-mercaptoethanol and loaded onto a 5-mL

HiTrap Heparin column (Cytiva) using an AKTA Pure FPLC system (GE Healthcare). Proteins were eluted over a 60-mL gradient to 15 mM Tris·HCl, pH 8.0, 10 mM $MgCl_2$, 1 M NaCl, and 2 mM β-mercaptoethanol. Fractions containing Eta were identified by sodium dodecyl sulfate (SDS)-PAGE, pooled, and dialyzed into storage buffer (15 mM Tris·HCl, pH 8.0, 10 mM $MgCl_2$, 250 mM KCl, 2 mM β-mercaptoethanol, and 50% glycerol) and concentrated by column centrifugation (Vivaspin 50-kDa molecular weight cutoff). Resulting enzymes was quantified using a Qubit Protein Assay Kit (Invitrogen).

Translocase (Streptavidin Displacement) Assays. CM0082-IB (5′-CTGGCT GTGGCGTGTTTCTGGTGG[bio-dT]ACCTAGGTCTTAGCCGTCTACGCCTCACT-3′) was radiolabeled with T4 Polynucleotide Kinase (New England Biolabs) in the presence of [γP32]ATP (PerkinElmer). For dsDNA assays, radiolabeled CM0082-IB was incubated in fivefold molar excess CM0082-RC-IB (5′-AGTGAGGCGTAGACGGCTAAGA CCTAGG[bio-dT]ACCACCAGAAACACGCCACAGCCAG-3′, heated to 95 °C for 2 min, and slowly cooled to room temperature to anneal the two strands. Single-stranded radiolabeled CM0082-IB or radiolabeled double-stranded CM0082-IB/CM0082-RC-IB was bound to 40-fold molar excess streptavidin. Translocase reactions were assembled with final concentrations of 5 nM ssDNA or dsDNA substrate and 15 nM Eta in 20 mM Tris·HCl, pH 8.0, 10 mM $MgCl_2$, 10 μg/mL bovine serum albumin (BSA), and 1 mM dithiothreitol (DTT). Reactions were heated to 50 °C and started by addition of dATP and excess biotin (400 μM) to trap streptavidin displacement events. At the appropriate time points, 10-μL aliquots were removed and stopped on ice with the addition of 3 μL 1.5% SDS and 50% glycerol. Terminated reactions were separated on 15% native polyacrylamide gels, and radiolabeled products were detected by exposure of the gel to a phosphorimaging screen (GE Healthcare) and analyzed using GE Imagequant 5.2.

Helicase Assays. The short DNA strand of the 3′-overhung dsDNA substrate (CM0080; 5′-GACCTAGGAACCACCAGAAACACGCCACAGCCAG-3′) was radiolabeled with T4 Polynucleotide Kinase (New England Biolabs) in the presence of [γ^{P32}]ATP (PerkinElmer). Equal molar amounts of CM0080 and the long DNA strand of the 3′-overhung dsDNA substrate (CM0082; 5′-CTGGCTGTGGCGTGTTT CTGGTGGTTCCTAGGTCTTAGCCGTCTACGCCTCACT-3′) were mixed, heated to 95 °C for 2 min, and slowly cooled to room temperature to anneal the two strands, forming the substrate. Helicase reactions were assembled with final concentrations of 5 nM substrate and 15 nM Eta in 20 mM Tris·HCl, pH 8.0, 10 mM $MgCl_2$, 10 μg/mL BSA, and 1 mM DTT. Reactions were heated to 50 °C and started by addition of dATP and trap DNA (CM0079; 5′-CTGGCTGTGGCGT GTTTCTGGTGGTTCCTAGGTC-3′) to 3.5 mM and 25 nM, respectively. At the appropriate time points, 10-μL aliquots were removed and stopped on ice with the addition of 3 μL 1.5% SDS and 50% glycerol. Terminated reactions were separated on 15% native polyacrylamide gels, and radiolabeled products were detected by exposure of the gel to a phosphorimaging screen (GE Healthcare) and analyzed using GE Imagequant 5.2.

Malachite Green ATPase Assays. One-hundred-microliter reactions were assembled with 500 nM substrate (prepared from CM0080 and CM0082 DNA oligonucleotides, as above) and 250 nM Eta in 20 mM Tris·HCl, pH 8.0, 10 mM $MgCl_2$, 10 μg/mL BSA, and 1 mM DTT with 200 μM ATP. Reactions were heated to 50 °C for 90 s, stopped by cooling on ice and the addition of 20 μL Malachite Green reagent (MAK-307, Sigma-Aldrich), followed by 15-min incubation at room temperature in a 96-well plate to develop. Absorbance at 620 nm was determined for each reaction in the 96-well plate using a BioTek Synergy 2. Two-hundred-fifty-nanomolar Eta^{WT}, in the presence of 500 nM DNA substrate, converts ∼11% (correcting for nonenzymatic breakdown of ATP) of the available 200 μM ATP to free phosphate at 50 °C under the conditions tested. Assuming Eta is functioning at maximal velocity (V_{max}), with one active site, the determined catalytic constant is 56 ATP converted per minute per enzyme (2.3 nmol phosphate per 90 s per 27 pmol Eta^{WT}).

In Vitro Transcription Termination Assays. For in vitro transcription termination assays presented in Fig. 5, assembly of preinitiation complexes and elongation to generate $TECs_{+58}$ were performed as described (68). $TECs_{+58}$ were captured using nickel-coated magnetic beads (Thermo Fisher), washed three times in 20 mM Tris·HCl, pH 8.0, 0.1 mM ethylenediaminetetraacetic acid (EDTA), 250 mM KCl, 4 mM $MgCl_2$, and 20 μg/mL BSA, and resuspended 250 mM KCl, 20 mM Tris·HCl, pH 8.0, 10 mM $MgCl_2$, 2 mM DTT, 10 μM ATP, 10 μM GTP, and 10 μM UTP. $TECs_{+58}$ were incubated in the absence or presence of 500 nM Eta (or Eta variant) and 5 mM dATP at 85 °C for 2, 5, or 10 min. Reaction aliquots (20 μL) were removed to ice-cold tubes containing streptavidin paramagnetic particles (Promega) that permit separation of intact and disrupted TECs through binding to biotin conjugated to the 5′ end of the nontemplate strand of DNA substrates. Supernatants were added to 100 μL 0.6 M Tris·HCl, pH 8.0, and 12 mM EDTA and streptavidin bead pellets were resuspended in 120 μL 0.5 M Tris·HCl, pH 8.0, and 10 mM EDTA. Pellets and supernatants of each reaction were then subjected to an equal volume P/C/I (25:24:1 [vol/vol/vol] phenol/chloroform/isoamyl alcohol) extraction. Nucleic acids were precipitated with ethanol and resuspended in 4 μL formamide loading dye before separation by electrophoresis through a 15% polyacrylamide/8 M urea denaturing gel. Radiolabeled RNAs were detected by exposure to a phosphorimaging screen (GE Healthcare) and analyzed using GE ImageQuant 5.2. Release of RNAs was calculated by quantifying transcripts in the supernatant samples divided by transcripts quantified in both the pellet and supernatant samples.

In vitro transcription termination experiments presented in Fig. 6 required assembly of preinitiation complexes and elongation to generate $TECs_{+58}$ as previous described (68). After nickel capture and washing of TECs as described above, TECs were either resuspended in Buffer A (250 mM KCl, 20 mM Tris·HCl, pH 8.0, 10 mM $MgCl_2$, 2 mM DTT, 10 μM ATP, 10 μM GTP, and 10 μM UTP) to prevent backtracking of TECs or Buffer B (250 mM KCl, 20 mM Tris·HCl, pH 8.0, 10 mM $MgCl_2$, and 2 mM DTT) to allow backtracking of TECs. Buffer A- and buffer B-resuspended TECs were incubated in the absence or presence of 500 nM Eta (or Eta variant) and 5 mM dATP at 85 °C for 5 min. Reactions (20 μL) were removed to ice-cold tubes containing streptavidin paramagnetic particles (Promega) that permit separation of intact and disrupted TECs through binding to biotin conjugated to the 5′ end of the nontemplate strand of DNA substrates. Supernatants were added to 100 μL 0.6 M Tris·HCl, pH 8.0, and 12 mM EDTA and streptavidin bead pellets were resuspended in 120 μL 0.5 M Tris·HCl, pH 8.0, and 10 mM EDTA. Pellets and supernatants of each reaction were then subjected to an equal volume P/C/I (25:24:1 vol/vol/vol) extraction. Nucleic acids were precipitated with ethanol and resuspended in 4 μL formamide loading dye before separation by electrophoresis through a 15% polyacrylamide/8 M urea denaturing gel. Radiolabeled RNAs were detected by exposure to a phosphorimaging screen (GE Healthcare) and analyzed using GE ImageQuant 5.2. Release of RNAs was calculated by quantifying transcripts in the supernatant samples divided by transcripts quantified in both the pellet and supernatant samples.

Data Availability. Coordinates and structure factors have been deposited in the PDB (ID code 7UJB). All other study data are included in the article and/or *SI Appendix*.

ACKNOWLEDGMENTS. We thank the Macromolecular X-ray science staff at the Cornell High Energy Synchrotron Source (MacCHESS) for supporting the crystallographic data collection. Experimental and personnel costs at Colorado State University were supported by National Institute of General Medical Sciences funding (R01 GM100329) and NSF funding (MCB-2016857) to T.J.S. Efforts at Pennsylvania State University were supported by NIH funding (R01 GM087350 and R35 GM131860) to K.S.M.

Author affiliations: [a]Department of Biochemistry and Molecular Biology, Colorado State University, Fort Collins, CO 80523; and [b]Department of Biochemistry and Molecular Biology, The Pennsylvania State University, University Park, PA 16802

Author contributions: C.J.M., K.S.M., and T.J.S. designed research; C.J.M., M.Z.Q., J.E.W., and T.J.S. performed research; C.J.M., M.Z.Q., K.S.M., and T.J.S. analyzed data; and C.J.M., M.Z.Q., J.E.W., K.S.M., and T.J.S. wrote the paper.

The authors declare no competing interest.

1. W. Lin *et al.*, Structural basis of ECF-σ-factor-dependent transcription initiation. *Nat. Commun.* **10**, 710 (2019).
2. D. Grohmann *et al.*, The initiation factor TFE and the elongation factor Spt4/5 compete for the RNAP clamp during transcription initiation and elongation. *Mol. Cell* **43**, 263–274 (2011).
3. A. Agapov, D. Esyunina, D. Pupov, A. Kulbachinskiy, Regulation of transcription initiation by Gfh factors from *Deinococcus radiodurans*. *Biochem. J.* **473**, 4493–4505 (2016).
4. M. Lilic, S. A. Darst, E. A. Campbell, Structural basis of transcriptional activation by the *Mycobacterium tuberculosis* intrinsic antibiotic-resistance transcription factor WhiB7. *Mol. Cell* **81**, 2875–2886.e5 (2021).

5. A. Wu *et al.*, DOT1L complex regulates transcriptional initiation in human erythroleukemic cells. *Proc. Natl. Acad. Sci. U.S.A.* **118**, e2106148118 (2021).
6. D. F. Browning, S. J. W. Busby, Local and global regulation of transcription initiation in bacteria. *Nat. Rev. Microbiol.* **14**, 638–650 (2016).
7. J. Y. Kang *et al.*, Structural basis for transcript elongation control by NusG family universal regulators. *Cell* **173**, 1650–1662.e14 (2018).
8. J. Ma *et al.*, Transcription factor regulation of RNA polymerase's torque generation capacity. *Proc. Natl. Acad. Sci. U.S.A.* **116**, 2583–2588 (2019).
9. Z. Hao *et al.*, Pre-termination transcription complex: Structure and function. *Mol. Cell* **81**, 281–292.e8 (2021).
10. V. Svetlov, E. Nudler, Reading of the non-template DNA by transcription elongation factors. *Mol. Microbiol.* **109**, 417–421 (2018).
11. T. J. Sanders *et al.*, TFS and Spt4/5 accelerate transcription through archaeal histone-based chromatin. *Mol. Microbiol.* **111**, 784–797 (2019).
12. W. Kobayashi, H. Kurumizaka, Structural transition of the nucleosome during chromatin remodeling and transcription. *Curr. Opin. Struct. Biol.* **59**, 107–114 (2019).
13. S. M. Fuchs, R. N. Laribee, B. D. Strahl, Protein modifications in transcription elongation. *Biochim. Biophys. Acta* **1789**, 26–36 (2009).
14. J. P. Hsin, J. L. Manley, The RNA polymerase II CTD coordinates transcription and RNA processing. *Genes Dev.* **26**, 2119–2137 (2012).
15. J. Y. Kang *et al.*, Structural basis for transcription complex disruption by the Mfd translocase. *eLife* **10**, e62117 (2021).
16. J. E. Walker, O. Luyties, T. J. Santangelo, Factor-dependent archaeal transcription termination. *Proc. Natl. Acad. Sci. U.S.A.* **114**, E6767–E6773 (2017).
17. N. Said, *et al.*, Steps toward translocation-independent RNA polymerase inactivation by terminator ATPase ρ. *Science.* **371**, eabd1673 (2021).
18. V. Epshtein, D. Dutta, J. Wade, E. Nudler, An allosteric mechanism of Rho-dependent transcription termination. *Nature* **463**, 245–249 (2010).
19. J. D. Eaton *et al.*, Xrn2 accelerates termination by RNA polymerase II, which is underpinned by CPSF73 activity. *Genes Dev.* **32**, 127–139 (2018).
20. H. E. Mischo *et al.*, Cell-Cycle Modulation of Transcription Termination Factor Sen1. *Mol. Cell* **70**, 312–326.e7 (2018).
21. Y. Jiang, M. Liu, C. A. Spencer, D. H. Price, Involvement of transcription termination factor 2 in mitotic repression of transcription elongation. *Mol. Cell* **14**, 375–385 (2004).
22. C. J. Marshall, "Factors and mechanisms of archaeal transcription termination and DNA repair," PhD thesis, Colorado State University, Ft. Collins, CO (2022).
23. A. G. Arimbasseri, K. Rijal, R. J. Maraia, Transcription termination by the eukaryotic RNA polymerase III. *Biochim. Biophys. Acta* **1829**, 318–330 (2013).
24. T. J. Santangelo, L. Cubonová, K. M. Skinner, J. N. Reeve, Archaeal intrinsic transcription termination in vivo. *J. Bacteriol.* **191**, 7102–7108 (2009).
25. J. W. Roberts, Mechanisms of bacterial transcription termination. *J. Mol. Biol.* **431**, 4030–4039 (2019).
26. K. M. Kazmierczak, E. K. Davydova, A. A. Mustaev, L. B. Rothman-Denes, The phage N4 virion RNA polymerase catalytic domain is related to single-subunit RNA polymerases. *EMBO J.* **21**, 5815–5823 (2002).
27. C. L. Turnbough Jr., Regulation of bacterial gene expression by transcription attenuation. *Microbiol. Mol. Biol. Rev.* **83**, e00019-19 (2019).
28. T. J. Sanders, C. J. Marshall, B. R. Wenck, J. N. Selan, T. J. Santangelo, "Transcription termination" in *Reference Module in Life Sciences* (Elsevier, ed. 3, 2021).
29. T. J. Sanders *et al.*, FttA is a CPSF73 homologue that terminates transcription in Archaea. *Nat. Microbiol.* **5**, 545–553 (2020).
30. P. Mitra, G. Ghosh, M. Hafeezunnisa, R. Sen, Rho protein: Roles and mechanisms. *Annu. Rev. Microbiol.* **71**, 687–709 (2017).
31. M. S. Ciampi, Rho-dependent terminators and transcription termination. *Microbiology (Reading)* **152**, 2515–2528 (2006).
32. A. Weixlbaumer, F. Grünberger, F. Werner, D. Grohmann, Coupling of transcription and translation in Archaea: Cues from the bacterial world. *Front. Microbiol.* **12**, 661827 (2021).
33. S. L. French, T. J. Santangelo, A. L. Beyer, J. N. Reeve, Transcription and translation are coupled in Archaea. *Mol. Biol. Evol.* **24**, 893–895 (2007).
34. N. J. Proudfoot, Transcriptional termination in mammals: Stopping the RNA polymerase II juggernaut. *Science.* **352**, aad9926 (2016).
35. J. Shi *et al.*, Structural basis of Mfd-dependent transcription termination. *Nucleic Acids Res.* **48**, 11762–11772 (2020).
36. Z. Xie, D. Price, *Drosophila* factor 2, an RNA polymerase II transcript release factor, has DNA-dependent ATPase activity. *J. Biol. Chem.* **272**, 31902–31907 (1997).
37. O. Adebali, Y. Y. Chiou, J. Hu, A. Sancar, C. P. Selby, Genome-wide transcription-coupled repair in *Escherichia coli* is mediated by the Mfd translocase. *Proc. Natl. Acad. Sci. U.S.A.* **114**, E2116–E2125 (2017).
38. J. Griesenbeck, H. Tschochner, D. Grohmann, Structure and function of RNA polymerases and the transcription machineries. *Subcell. Biochem.* **83**, 225–270 (2017).
39. A. Hirata, B. J. Klein, K. S. Murakami, The X-ray crystal structure of RNA polymerase from Archaea. *Nature* **451**, 851–854 (2008).
40. T. J. Santangelo, J. N. Reeve, Archaeal RNA polymerase is sensitive to intrinsic termination directed by transcribed and remote sequences. *J. Mol. Biol.* **355**, 196–210 (2006).
41. S. J. Northall *et al.*, DNA binding and unwinding by Hel308 helicase requires dual functions of a winged helix domain. *DNA Repair (Amst.)* **57**, 125–132 (2017).
42. B. G. Wilson, C. W. M. Roberts, SWI/SNF nucleosome remodellers and cancer. *Nat. Rev. Cancer* **11**, 481–492 (2011).
43. L. He, M. St John James, M. Radovcic, I. Ivancic-Bace, E. L. Bolt, Cas3 protein–A review of a multi-tasking machine. *Genes (Basel)* **11**, 208 (2020).
44. A. K. Byrd, K. D. Raney, Superfamily 2 helicases. *Front. Biosci.* **17**, 2070–2088 (2012).
45. S. J. Black *et al.*, Molecular basis of microhomology-mediated end-joining by purified full-length Polθ. *Nat. Commun.* **10**, 4423 (2019).
46. F. He, K. DuPrez, E. Hilario, Z. Chen, L. Fan, Structural basis of the XPB helicase-Bax1 nuclease complex interacting with the repair bubble DNA. *Nucleic Acids Res.* **48**, 11695–11705 (2020).
47. T. van Eeuwen *et al.*, Cryo-EM structure of TFIIH/Rad4-Rad23-Rad33 in damaged DNA opening in nucleotide excision repair. *Nat. Commun.* **12**, 3338 (2021).
48. L. Holm, Using Dali for protein structure comparison. *Methods Mol. Biol.* **2112**, 29–42.
49. J. D. Richards *et al.*, Structure of the DNA repair helicase hel308 reveals DNA binding and autoinhibitory domains. *J. Biol. Chem.* **283**, 5118–5126 (2008).
50. E. Absmeier, K. F. Santos, M. C. Wahl, Molecular mechanism underlying inhibition of intrinsic ATPase activity in a Ski2-like RNA helicase. *Structure* **28**, 236–243.e3 (2020).
51. T. Kouba *et al.*, Mycobacterial HelD is a nucleic acids-clearing factor for RNA polymerase. *Nat. Commun.* **11**, 6419 (2020).
52. C. J. Marshall, T. J. Santangelo, Archaeal DNA repair mechanisms. *Biomolecules* **10**, 1472 (2020).
53. A. Fujiwara *et al.*, Application of a Euryarchaeota-specific helicase from *Thermococcus kodakarensis* for noise reduction in PCR. *Appl. Environ. Microbiol.* **82**, 3022–3031 (2016).
54. B. Clouet-d'Orval *et al.*, Insights into RNA-processing pathways and associated RNA-degrading enzymes in Archaea. *FEMS Microbiol. Rev.* **42**, 579–613 (2018).
55. D. K. Phung *et al.*, RNA processing machineries in Archaea: The 5′-3′ exoribonuclease aRNase J of the β-CASP family is engaged specifically with the helicase ASH-Ski2 and the 3′-5′ exoribonucleolytic RNA exosome machinery. *Nucleic Acids Res.* **48**, 3832–3847 (2020).
56. L. A. Kelley, S. Mezulis, C. M. Yates, M. N. Wass, M. J. E. Sternberg, The Phyre2 web portal for protein modeling, prediction and analysis. *Nat. Protoc.* **10**, 845–858 (2015).
57. J. Jumper *et al.*, Highly accurate protein structure prediction with AlphaFold. *Nature* **596**, 583–589 (2021).
58. S. H. Jun *et al.*, Direct binding of TFEα opens DNA binding cleft of RNA polymerase. *Nat. Commun.* **11**, 6123 (2020).
59. A. J. Smith, N. J. Savery, RNA polymerase mutants defective in the initiation of transcription-coupled DNA repair. *Nucleic Acids Res.* **33**, 755–764 (2005).
60. N. Stantial, J. Dumpe, K. Pietrosimone, F. Baltazar, D. J. Crowley, Transcription-coupled repair of UV damage in the halophilic archaea. *DNA Repair (Amst.)* **41**, 63–68 (2016).
61. A. M. Gehring, T. J. Santangelo, Archaeal RNA polymerase arrests transcription at DNA lesions. *Transcription* **8**, 288–296 (2017).
62. M. Ogrünç, D. F. Becker, S. W. Ragsdale, A. Sancar, Nucleotide excision repair in the third kingdom. *J. Bacteriol.* **180**, 5796–5798 (1998).
63. C. P. Selby, Mfd protein and transcription-repair coupling in *Escherichia coli*. *Photochem. Photobiol.* **93**, 280–295 (2017).
64. Z. Otwinowski, W. Minor, Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* **276**, 307–326 (1997).
65. P. D. Adams *et al.*, PHENIX: A comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr. D Biol. Crystallogr.* **66**, 213–221 (2010).
66. P. Emsley, K. Cowtan, Coot: Model-building tools for molecular graphics. *Acta Crystallogr. D Biol. Crystallogr.* **60**, 2126–2132 (2004).
67. J. S. Papadopoulos, R. Agarwala, COBALT: Constraint-based alignment tool for multiple protein sequences. *Bioinformatics* **23**, 1073–1079 (2007).
68. A. M. Gehring, T. J. Santangelo, Manipulating archaeal systems to permit analyses of transcription elongation-termination decisions in vitro. *Methods Mol. Biol.* **1276**, 263–279 (2015).