nature microbiology

LETTERS

https://doi.org/10.1038/s41564-020-0667-3

Check for updates

FttA is a CPSF73 homologue that terminates transcription in Archaea

Travis J. Sanders[1], Breanna R. Wenck[1], Jocelyn N. Selan[1], Mathew P. Barker[1], Stavros A. Trimmer[1], Julie E. Walker[1,2] and Thomas J. Santangelo[1]✉

Regulated gene expression is largely achieved by controlling the activities of essential, multisubunit RNA polymerase transcription elongation complexes (TECs). The extreme stability required of TECs to processively transcribe large genomic regions necessitates robust mechanisms to terminate transcription. Efficient transcription termination is particularly critical for gene-dense bacterial and archaeal genomes[1–3] in which continued transcription would necessarily transcribe immediately adjacent genes and result in conflicts between the transcription and replication apparatuses[4–6]; the coupling of transcription and translation[7,8] would permit the loading of ribosomes onto aberrant transcripts. Only select sequences or transcription termination factors can disrupt the otherwise extremely stable TEC and we demonstrate that one of the last universally conserved archaeal proteins with unknown biological function is the Factor that terminates transcription in Archaea (FttA). FttA resolves the dichotomy of a prokaryotic gene structure (operons and polarity) and eukaryotic molecular homology (general transcription apparatus) that is observed in Archaea. This missing link between prokaryotic and eukaryotic transcription regulation provides the most parsimonious link to the evolution of the processing activities involved in RNA 3′-end formation in Eukarya.

Transcription termination (Extended Data Fig. 1), driven either by DNA sequence and encoded RNA structures (for example, intrinsic termination) or by protein factors (for example, factor-dependent termination), ensures the rapid dissociation of RNA polymerase (RNAP) from the DNA template to recycle RNAP and generate RNA 3′ ends[1,9]. While often prevalent within prokaryotic genomes, intrinsic termination sequences are typically neither sufficiently abundant nor efficient to mediate all termination events. Transcription termination factors must then efficiently recognize transcription elongation complexes (TECs) that are not intrinsically terminated and compete with continued elongation to mediate the release of the nascent transcript. While the identification of Eta provided evidence of factor-dependent archaeal termination[10], no kinetically efficient mechanism of factor-dependent archaeal transcription termination has been described. The retention of operon-organized archaeal genomes and the sensitivity of the archaeal transcription apparatus to bacterial Rho-mediated termination in vitro[2] (combined with the normal coupling of transcription and translation[8], the resultant polar suppression of downstream expression in the absence of such coupling in archaeal cells[11] and the conservation of Spt5/NusG in all genomes) implied the existence of a kinetically relevant archaeal transcription termination activity that might function akin to the bacterial Rho protein. However, Rho homologues are restricted to Bacteria[12], suggesting instead that conserved archaeal–eukaryotic or unique archaeal factors may drive factor-dependent archaeal transcription termination.

Only a core set of ~200 gene families (more properly, archaeal clusters of orthologous genes; arCOGs) are conserved in most archaeal genomes[13]; just ~129 arCOGs are strictly ubiquitous[14] and one is an obvious orthologue of a subunit of the cleavage and polyadenylation specificity factor (CPSF) complex[13–18]. The homology of most archaeal transcription components to eukaryotic factors suggested that the archaeal homologue of eukaryotic CPSF73 might function as the Factor that terminates transcription in Archaea (FttA). We challenged promoter-initiated TECs[19,20] (generated with an RNAP variant with a His_6–HA-epitope-tagged RpoL subunit[19,21,22]—where His6 represents six histidine residues, HA represents haemagglutinin and RpoL represents the RNA polymerase subunit L—and containing a radiolabelled nascent transcript) with FttA (Extended Data Fig. 2; FttA is the ~73.5 kDa protein product of gene TK1428) and monitored transcription termination by quantifying the release of transcripts from TECs (Fig. 1). TECs stalled by nucleotide deprivation with +125-nucleotide (nt) nascent transcripts ($TECs_{+125}$) remain stably associated in the absence of FttA (Fig. 1a, lanes 1–4). The addition of FttA to stalled TECs results in the cleavage and release of ~100 nt of the nascent transcript (Fig. 1a, lanes 5–6). However, despite repeated and exhaustive efforts to monitor FttA-mediated transcript cleavage within seconds of FttA addition, we never observed an ~25-nt 3′ transcript fragment. We were thus initially hesitant to assume that FttA-mediated transcript cleavage was coupled to bona fide transcription termination, as an ~25-nt transcript is sufficient to stabilize an archaeal TEC[2,23,24].

To fully validate that the cleavage and termination activity of FttA is distinct from that of a general RNase, we challenged TECs with either FttA or RNase I_f in parallel. If TECs remain intact following FttA-mediated cleavage of the nascent transcript then: radiolabelled 3′ nascent transcripts should remain associated with TECs; intact TECs should survive washes designed to remove transcripts not associated with TECs; nucleotide triphosphate (NTP) addition should permit the continued elongation of active TECs, allowing extension of the nascent transcript; and RNAP should remain within TECs. By contrast, if FttA-mediated cleavage of the transcripts inactivates and terminates transcription, RNAP should be released to the supernatant and resumed elongation following NTP supplementation will not be possible. Treatment of $TECs_{+125}$ with RNase I_f (ref. [10]; Fig. 1c) fulfilled all of the expectations of transcript processing that are not linked to transcription termination: stable $TECs_{+25}$ were observed (Fig. 1b, lanes 10–11), the addition of unlabelled NTPs (Fig. 1b, lane 12) or radiolabelled NTPs (Fig. 1b, lane

[1]Department of Biochemistry and Molecular Biology, Colorado State University, Fort Collins, CO, USA. [2]Present address: Watchmaker Genomics, Boulder, CO, USA. ✉e-mail: thomas.santangelo@colostate.edu

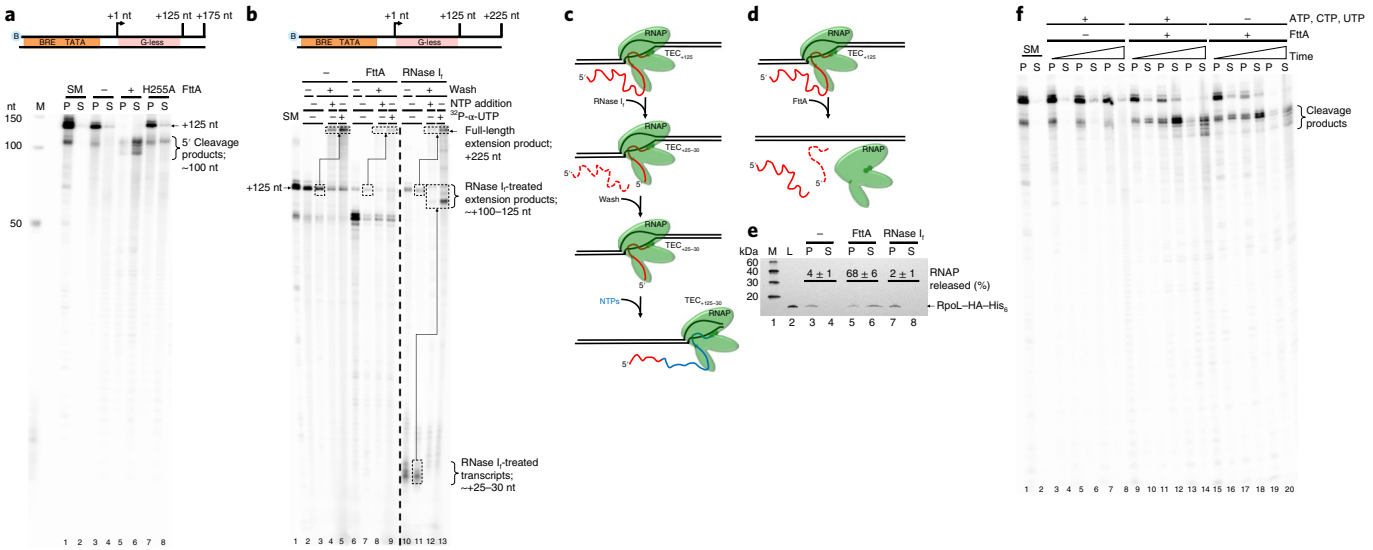


Fig. 1 | FttA is a bona fide archaeal transcription termination factor. a, Transcripts within intact TECs are retained in pellet (P) fractions; transcripts released from terminated TECs partition into the supernatant (S). Radiolabelled transcripts within starting material (SM) $TECs_{+125}$ and mock-treated $TECs_{+125}$ (–) are retained in P fractions (lanes 1–4), whereas $FttA^{WT}$ addition (+) results in the cleavage of nascent transcripts and the termination of most TECs (lanes 5–6). A catalytically deficient FttA variant ($FttA^{H255A}$) abrogates cleavage and RNA release (lanes 7–8). Lane M contains ^{32}P-labelled single-stranded DNA markers. **b**, FttA-mediated termination is distinct from RNase treatment of intact TECs. $TECs_{+125}$ (SM, lane 1) are resistant to repeated washes and readily resume elongation following NTP addition to generate +225-nt transcripts (lanes 2–5). Dashed boxes and arrows denote +125-nt transcripts that are elongated to +225-nt transcripts; the specific activity of +225-nt transcripts can be increased by the addition of more ^{32}P-α-UTP during resumed elongation. RNase I_f digestion of nascent transcripts associated with washed $TECs_{+125}$ results in the degradation of the nascent transcript to just ~20–30 nt, but TECs with shortened transcripts remain associated with the DNA and survive repeated washing (lanes 10–11). $TECs_{~+25–30}$ resulting from RNase I_f treatment of $TECs_{+125}$ readily resume elongation following NTP addition to generate ~+125-nt full-length transcripts (lanes 12–13). Dashed black line separates the FttA and RNase I_f samples. Dashed boxes and arrows denote ~+25-nt transcripts that are elongated to ~+125-nt transcripts; the specific activity of ~+125-nt transcripts can be increased by the addition of ^{32}P-α-UTP during resumed elongation. FttA addition to $TECs_{+125}$ disrupts most TECs with nascent transcript cleavage (lanes 6–9), resulting in the release of most TECs from the template; cleaved transcripts cannot be extended by NTP addition (lanes 8–9). B, biotin; BRE, TFB recognition element. **c**,**d**, Diagrams of the fate of $TECs_{+125}$ following RNase I_f (**c**) and FttA treatment (**d**). Solid red lines represent nascent RNA and the brown segments represent nascent RNA encapsulated within the TEC. Dashed red lines represent RNAs degraded by FttA or RNase I_f. Solid blue lines represent the newly synthesized nascent RNA from the NTP pool (blue). **e**, FttA releases RNAP from the DNA template into solution confirming the dissociation of the TEC and bona fide FttA-mediated transcription termination. RNAP was tracked and quantified by western blots ($n = 3$ independent replicates) with anti-HA antibodies that recognize the modified RpoL subunit. M, markers; L, load. **f**, FttA is not reliant on NTP hydrolysis to inactivate TECs, cleave nascent transcripts and terminate transcription. For **a**,**b**,**f**, similar results were observed in four independent experiments.

13) permitted all $TECs_{~+25}$ to resume elongation and RNAP partitioning confirmed that essentially all TECs remained intact (Fig. 1e, lanes 7–8). The treatment of identically prepared $TECs_{+125}$ with FttA (Fig. 1b) is, by contrast, fully supportive of FttA-mediated termination: the bulk of FttA-treated $TECs_{+125}$ do not survive washes and FttA activity releases ~70% of RNAP to solution (Fig. 1e, lanes 5–6). FttA is thus the second archaeal-encoded factor that can mediate transcription termination.

Addition of an FttA variant[15,17,25,26] ($FttA^{H255A}$) reduced but did not eliminate FttA-mediated termination (Fig. 1a, lanes 7–8). Termination activity is thus linked to FttA-mediated RNA cleavage, rather than FttA-mediated stimulation of the intrinsic cleavage activity of RNAP (ref. [27]). FttA-mediated cleavage of the nascent RNA to yield an ~100-nt 5′ transcript is consistent with FttA stimulating RNA cleavage at the first solvent-accessible phosphodiester linkage and the ~25 nt of nascent transcript protection is consistent with the results of previous digestions of intact archaeal[10] and eukaryotic TECs[28] with RNA exonucleases. In contrast to other prokaryotic transcription termination factors, FttA-mediated termination is not energy dependent (Fig. 1f).

FttA recognizes TECs through binding to nascent transcripts (Fig. 2). TECs stalled on G-less cassettes, and thus with A-, U- and C-rich RNAs, revealed a near-linear relationship between transcript length and FttA-mediated termination (Fig. 2b–d). Although FttA-mediated termination is possible with only short segments of solvent-accessible nascent transcript sequences (a notable discriminating feature between Rho- and FttA-mediated termination), the efficiency and rate of FttA-mediated termination are modest in such instances. By contrast, TECs stalled on C-less cassettes, and thus with A-, U- and G-rich RNAs, effectively abolish FttA activity (Fig. 2 and Extended Data Fig. 3a,b). FttA-mediated termination is thus stimulated by C-rich RNAs (as is the case for bacterial Rho-mediated termination) or inhibited by transcripts that are particularly G-rich. Rho activity can be stimulated at suboptimal *rut* sites by NusG; the archaeal–eukaryotic homologue of NusG, Spt5, together with its common binding partner Spt4, can likewise stimulate FttA when transcript sequences limit FttA recognition or FttA activity (Fig. 2c,d). As such, Spt4–Spt5 temper the nucleotide requirements of FttA. FttA is a known endo- and 5′–3′ exonuclease[15,29] and cleavage of nascent transcripts is stimulated by interactions with the archaeal TEC, but not RNAP alone (Extended Data Fig. 3d). FttA-mediated cleavage of TEC-associated nascent transcripts was completed within ~1–2 min, while incubations of FttA with purified RNA under identical conditions required ~30-times longer to generate even mild cleavage patterns (Fig. 3c), consistent with previous results[15,30]. FttA-mediated endonucleolytic cleavage of free RNA at CA and CC dinucleotide sequences is consistent with FttA activity on C-rich transcripts[15]. The consistently observed cleavage pattern

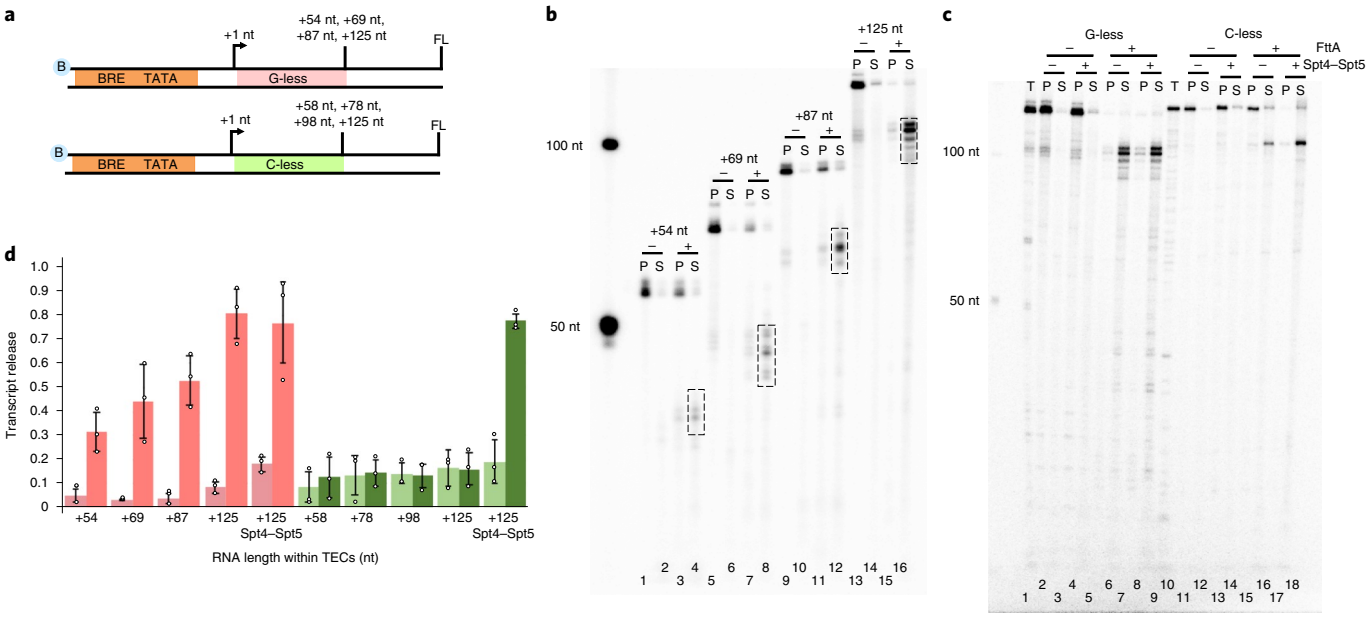


Fig. 2 | FttA-mediated termination shares mechanistic requirements of Rho-mediated bacterial transcription termination. a, Promoter-directed transcription of biotinylated templates encoding G-less or C-less cassettes permits the formation of TECs with increasingly long A-, C- and U-rich or A-, G- and U-rich nascent transcripts, respectively. FL, full-length; all templates permit elongation for 100 nt beyond the G- or C-less cassette. **b**, TECs remain stably associated and transcripts are primarily recovered in the P fraction in the absence (−) of FttA. When FttA is present (+), transcripts are cleaved and primarily recovered in the S fraction. Cleavage releases ~20–30-nt shorter transcripts (boxed). The left-most lane contains ^{32}P-labelled ssDNA markers. **c**, The addition of Spt4–Spt5 largely abrogates the RNA sequence requirements of FttA-mediated transcription termination. T, total reaction = P + S. The left-most lane contains ^{32}P-labelled ssDNA markers. **d**, Transcript release was quantified with (salmon/green) and without (pink/mint) FttA addition for TECs with transcripts of increasing length on G-less (pink/salmon) and C-less cassettes (mint/green), with and without Spt4–Spt5 addition for $TECs_{+125}$ formed on G- and C-less cassettes. Error bars were calculated as standard deviation from the mean ($n = 3$ independent experiments). For **b**,**c**, similar results were observed in three independent experiments.

on various substrates (reduced transcript length by ~20–30 nt) supports FttA-mediated cleavage and termination being dictated and positioned by RNAP–RNA interactions near the stalk domain and RNA exit channel and is further enhanced by Spt4–Spt5.

Archaeal transcription units are typically separated by only short (<100 base pair) intergenic regions[1,11,31,32]; thus, for FttA-mediated termination to be an effective mechanism of gene regulation in vivo, FttA must quickly recognize and disrupt TECs before transcription continues substantially into downstream genes or operons. To establish whether FttA-mediated termination was competitive with transcription elongation, stalled $TECs_{+125}$ were permitted to resume elongation with different [NTPs]. The differential elongation rates that resulted from varying [NTPs] provided a relative measure of the efficiency of FttA-mediated transcription termination in competition with transcription elongation (Fig. 3). At low [NTPs], TECs elongated slowly and many TECs were still transcribing after several minutes of incubation as evidenced by a mixture of nascent transcripts between 125–225 nt (Fig. 3b, lanes 3–4). At increasingly higher [NTPs], elongation rates increased until TECs were elongating at rates comparable to normal elongation rates in vivo (Fig. 3b, lanes 5–8). The addition of FttA to stalled TECs (Fig. 3b, lanes 17–18) resulted in near-complete termination, but as the rate of elongation increased with increasing [NTPs], FttA-mediated termination decreased. Although not an obligate subcomplex of archaeal RNAP, in vivo Spt4–Spt5 engages RNAP early during the elongation process and remains associated with TECs throughout long genes[33]. The ability of Spt4–Spt5 to temper the transcript requirements for FttA-mediated termination (Fig. 2) suggested that the addition of Spt4–Spt5 may accelerate FttA recognition of or action towards TECs. In support of this hypothesis, addition of Spt4–Spt5 greatly increased the termination efficiency of FttA, as demonstrated by the release of transcripts >+125 nt but <+225 nt (Fig. 3b, lanes 27–32). The results demonstrate that FttA is kinetically coupled to RNAP via the elongation factors Spt4–Spt5, a striking analogy to the stimulation of the unrelated bacterial Rho protein by NusG (ref. [34]) and to the observed stimulation of Pol II termination by CPSF73/Xrn2 (ref. [35]). To ensure that FttA mediates termination when combined with Spt4–Spt5 (and that termination observed in the presence of all three factors was not a new activity of Spt4–Spt5) we employed a variant of FttA ($FttA^{H255A}$) that retains only partial activity (Extended Data Fig. 4a,b).

Interactions between Rho and the C-terminal KOW domain of NusG stimulate Rho-mediated termination[36–38]. The NusG KOW domain is normally engaged with the ribosome and becomes available only when transcription becomes uncoupled from translation[39]. Archaeal transcription and translation are coupled[8]; we asked whether the isolated KOW domain of Spt5 would suffice to stimulate FttA-mediated termination. Addition of the Spt5 KOW domain ($Spt5^{\Delta NGN}$, which remains thermostable) alone does not influence the activities of FttA or RNAP in vitro (Extended Data Fig. 4c). Spt5 is often in a heterodimeric partnership with Spt4 and this partnership is critical to kinetically couple FttA activity to RNAPs (Extended Data Fig. 4d). As with the nuclear eukaryotic RNAPs, the archaeal RNAP contains a stalk domain[40]. The stalk provides binding surfaces for conserved initiation and elongation factors and the nascent transcript[1,32,41–43]. Purified stalkless RNAP ($RNAP^{\Delta E/\Delta F}$), when combined with the TATA-binding protein (TBP) and transcription factor B (TFB), is competent for transcription initiation, elongation and intrinsic termination[21], but fails to respond correctly to FttA-mediated termination (Extended Data Fig. 4e). Even when continued elongation was prohibited, the termination activities of FttA were stunted by the loss of the

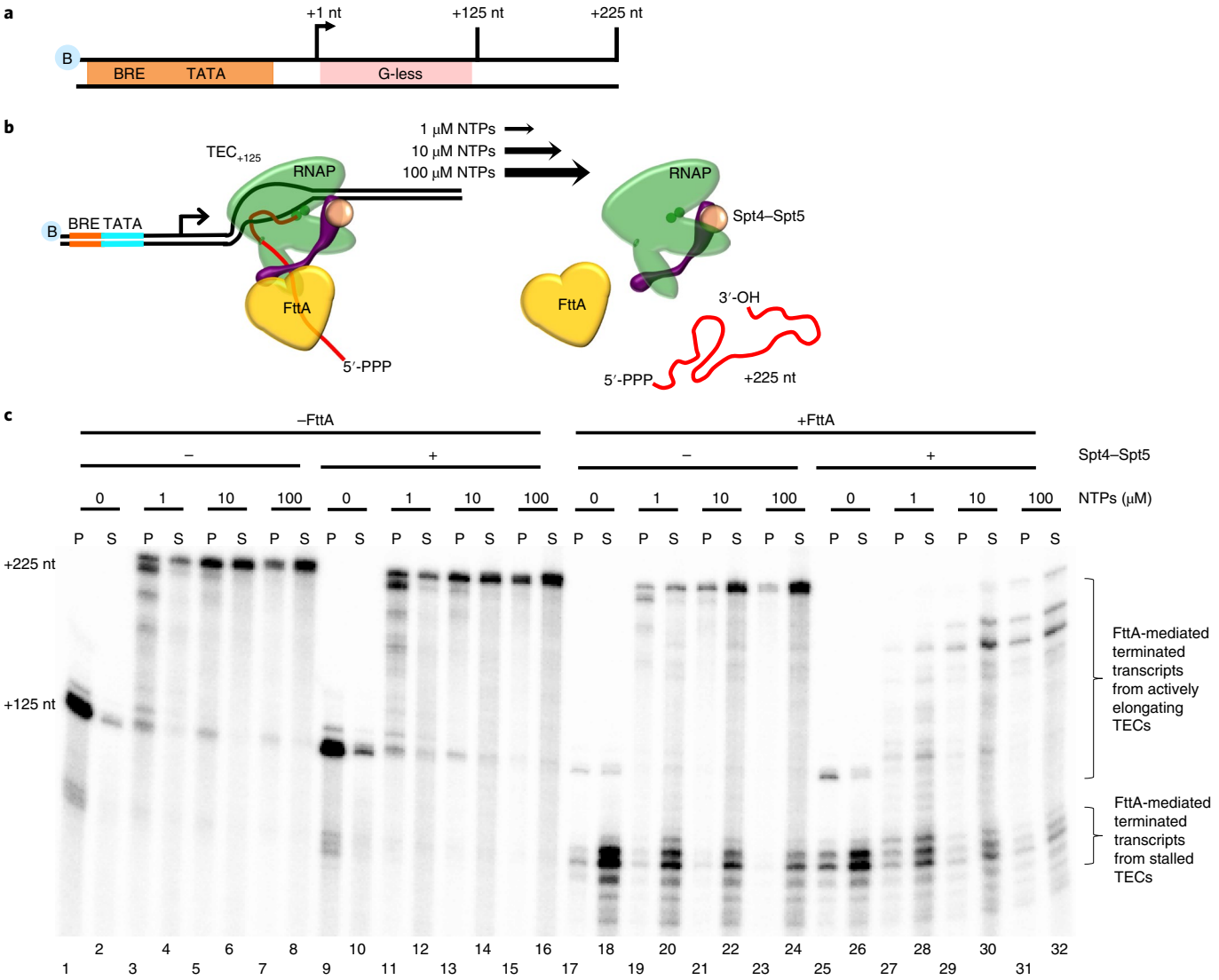


Fig. 3 | FttA-mediated transcription termination is competitive with transcription elongation. a, Washed, NTP-deprived $TECs_{+125}$ were assembled on biotinylated templates with a +125-nt G-less cassette. **b**, Resumed elongation following differential [NTPs] addition permits transcription to generate +225-nt transcripts, albeit at different rates. Arrow sizes reflect the increasing [NTP]. Solid red lines represent nascent RNA and the brown segment represents nascent RNA encapsulated within the TEC. **c**, FttA readily terminates stalled or slowly elongating TECs (lanes 17–24) and FttA-mediated termination becomes competitive with transcription elongation even at high [NTPs] in the presence of Spt4–Spt5 (lanes 25–32). Similar results were observed in three independent experiments.

RNAP stalk domain (Extended Data Fig. 4e, lanes 17–18). Addition of Spt4–Spt5 to TECs assembled with $RNAP^{\Delta E/\Delta F}$ does not stimulate FttA-mediated termination to rates that are competitive with continued elongation even at low [NTPs] (Extended Data Fig. 4e, lanes 27–32). The results suggest roles for both the stalk domain and Spt4–Spt5 in accelerating FttA-mediated termination to permit kinetically competitive termination in vitro.

FttA is probably sufficiently abundant (~2,100 ± 500 molecules per cell; Extended Data Fig. 5) to monitor global transcription (RNAP is estimated at ~3,000 molecules per cell)[19]. FttA is a metallo-beta lactamase-fold protein containing a beta-CASP domain[15] and we predicted that general inhibitors of metallo-beta-lactamase proteins[44] may impair FttA activity. 2,6-pyridine dicarboxylic acid (dipicolinic acid; DPA) nearly completely inhibited FttA-mediated termination in vitro (Fig. 4a,b). *Thermococcus kodakarensis* cultures challenged with DPA demonstrated a reduced, then complete inhibition of growth (Fig. 4c). While DPA may impact several factors in vivo, we rationalized that monitoring RNA 3′ ends following DPA addition may reveal changes due to the inactivation of FttA. Quantitative PCR with reverse transcription (qRT–PCR) analyses revealed substantial changes to the 3′ ends at several loci in vivo (Fig. 4d–f) following DPA addition. The fold changes in the extended 3′ ends of each transcript generally increased in magnitude compared to untreated cultures both with respect to the distance from the translation stop codon and with increasing [DPA]. Altered 3′ termini stemming from exposure to a general metallo-beta-lactamase inhibitor cannot be definitively attributed to direct inhibition of FttA activity in vivo. Given our desire to directly demonstrate that reduced FttA activity impacts termination in vivo, coupled with our inability to generate *T. kodakarensis* strains encoding enzymatically impaired FttA variants, we next reduced FttA activity by limiting FttA expression and altering steady-state wild-type ($FttA^{WT}$) protein levels. To ensure that the introduced and regulated expression of TK1428 did not impact TK1429 expression[45], we separated TK1428

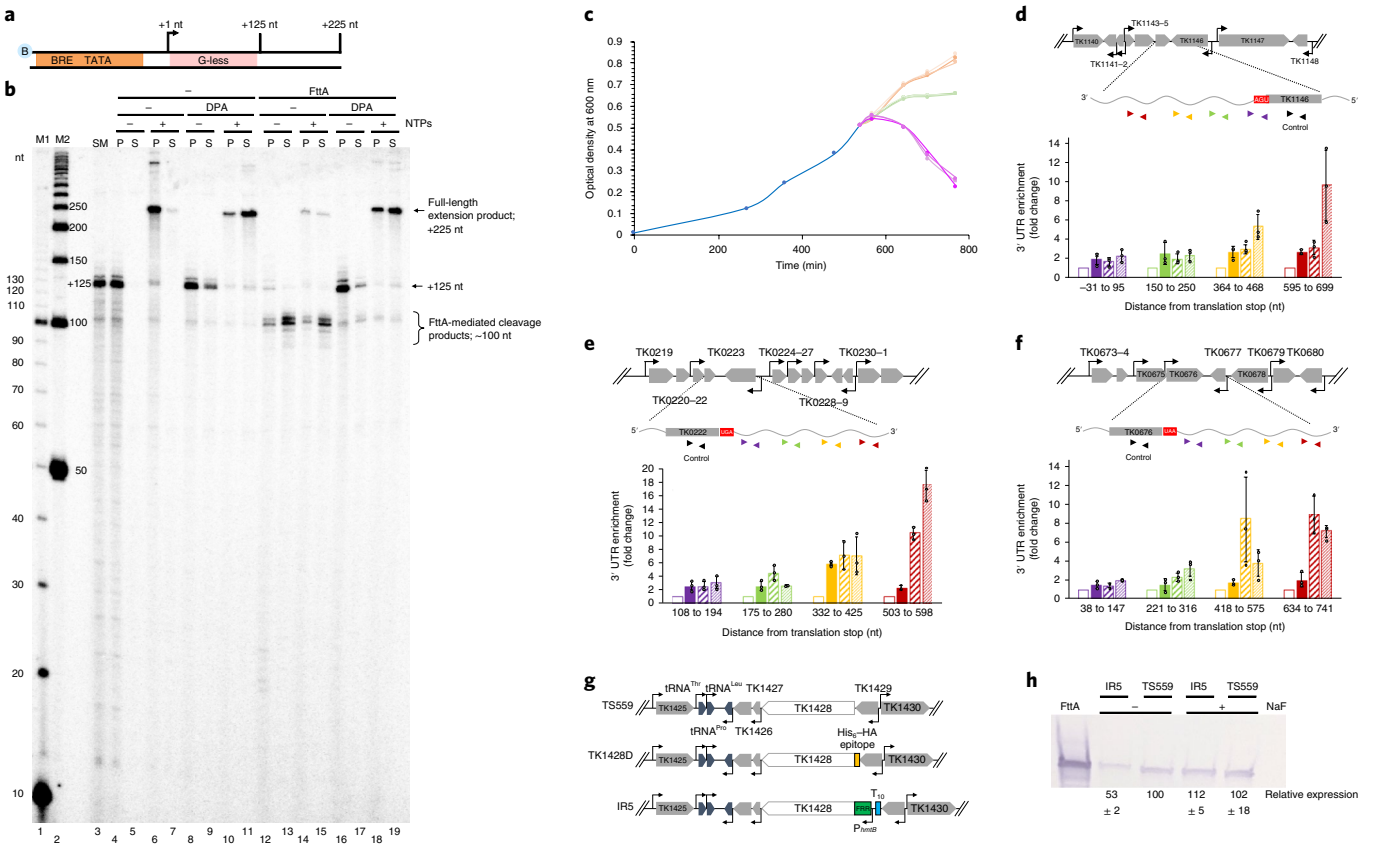


Fig. 4 | Inhibition of FttA activity abolishes transcription termination in vitro and reduced FttA expression or activity alters steady-state RNA 3′ termini in vivo. a, DNA templates supporting transcription. **b**, $TECs_{+125}$ (lane 3) resume elongation following NTP addition to generate +225-nt transcripts with or without 25 mM DPA (lanes 4–11). FttA addition results in transcript cleavage and the release of most TECs to the supernatant without DPA (lanes 12–13), inhibiting resumed elongation following NTP addition (lanes 14–15). Pre-incubation of FttA with 25 mM DPA inhibits FttA-mediated termination (lanes 16–17), permitting $TECs_{+125}$ to resume elongation (lanes 18–19). Lanes M1 and M2 contain ^{32}P-labelled 10- and 50-nt ssDNA markers, respectively. Similar results were observed in five independent experiments. **c**, The inhibition of metallo-beta-lactamase/beta-CASP protein activity impairs the growth of *T. kodakarensis*. A mid-log culture of *T. kodakarensis* strain TS559 was split into nine cultures, with three biological replicates exposed to 0 mM (peach), three to 12.5 mM (green) and three to 25 mM DPA (purple). **d**–**f**, RNAs recovered 1 h post-DPA addition to cultures of TS559, or from cultures of IR5 grown in the absence of NaF, displayed altered 3′ termini. TRIzol-extracted RNAs were reverse transcribed with primers complementary to the nascent transcript sequences of TK1146 (**d**), TK0222 (**e**) and TK0676 (**f**) to generate complementary DNAs that were quantified and normalized to internal controls. Inhibiting FttA activity with DPA or lowering steady-state FttA levels by riboswitch-mediated controlled expression impacted the abundance of RNAs with extended 3′ termini in vivo. RNA abundance in untreated TS559 cultures (open bars) was set to 1.0, and fold changes in the abundance of amplicons reflecting RNA transcripts with extended 3′ sequences at increasing distances from the translation stop site (purple, green, orange and red) are shown for strain IR5 (solid bars), TS559 + 12.5 mM DPA (wide stripes) and TS559 + 25 mM DPA (narrow stripes). Errors were calculated at a 95% confidence interval with the centre value as the mean of three biological replicates. **g**, Maps of the TK1428 locus in parental (TS559), N-terminally tagged (TK1428D) and riboswitch-regulated expression (IR5) strains. **h**, A western blot demonstrating the reduction in FttA protein levels in strain IR5 following the removal of NaF from the medium; $n = 3$ independent replicates. Size standards are identified by molecular weight (left).

expression from TK1429 by the introduction of a new promoter and intrinsic termination sequence and then placed the TK1428 coding sequences downstream of sequences encoding an archaeal fluoride-responsive riboswitch[46], thereby generating strain IR5 (Fig. 4g). Construction of IR5 was possible only when cultures were continuously provided with fluoride even though fluoride impairs the growth of *T. kodakarensis*[46], supporting the idea that very limited expression of TK1428 was not compatible with life. Steady-state FttA levels in IR5 strains grown in the absence and presence of fluoride revealed a modest ~twofold change in FttA levels in vivo when fluoride was removed from the cultures (Fig. 4h and Supplementary Fig. 2), yet even this modest alteration substantially and reproducibly impacted transcription termination in vivo (Fig. 4d–f). The increased abundance of RNA with extended 3′ untranslated regions (UTRs) in strains with reduced FttA protein abundance is supportive of FttA normally directing transcription termination in vivo.

FttA is conserved in all archaeal genomes[13,15,47,48], including the severely reduced genomes of symbiotic Nanoarchaeota, and it was perhaps not surprising that exhaustive attempts to delete or generate variants that radically impair the activity of FttA[26,49,50] in *T. kodakarensis* were unsuccessful; our failures were supported by the essentiality of FttA in other Archaea[47,48]. We were able to generate a strain (termed TK1428D) encoding a His_6-affinity and HA-epitope-tagged FttA (Fig. 4g). Strain TK1428D growth was indistinguishable from the parental strain[51,52] and N-terminally tagged FttA was easily recovered directly from TK1428D cell lysates in abundance (Extended Data Fig. 6). Proteins copurifying with FttA from TK1428D were identified by MuDPIT[53,54], returning only a small number of proteins (Extended Data Fig. 7) that have minimal inferred activity related to transcription and gene expression. No obvious stoichiometric FttA interaction partners were recovered, supportive of our in vitro demonstration that FttA alone can disrupt

archaeal TECs. Affinity purification of FttA does not return RNAP subunits nor Spt4–Spt5, suggesting that FttA transiently encounters and disrupts TECs rather than forming stable interactions with TEC components or Spt4–Spt5.

The essentiality of FttA in *T. kodakarensis* and other Archaea[47,48], the complete conservation of FttA in Archaea[14,15,18], the demonstrated in vitro ability of FttA-mediated termination to compete with productive elongation (Fig. 3) and the demonstrated changes to RNA 3′ ends in strains in which FttA activity is reduced by two independent mechanisms (Fig. 4) suggests that FttA is probably responsible for the 3′-end formation of transcripts that are not directed by intrinsic termination, and further that FttA-mediated termination is probably responsible for the polarity in archaeal cells[11]. By establishing the requirements for FttA-mediated transcription termination (Figs. 1–4, Extended Data Figs. 3 and 4) we complete the archaeal transcription cycle and describe an additional mechanism of 3′-end formation (Extended Data Fig. 8). The described activities of FttA suggest that the steady-state 3′ termini of in vivo transcripts terminated by FttA do not reflect the actual position of termination of the archaeal RNAP. Thus, consensus termination sequences derived from next-generation sequencing and Term-seq data[1,55] should be re-evaluated given that FttA-terminated transcripts probably lack at least ~20–30 nt from the 3′ terminus; additional RNA processing events are likely to further complicate attempts to map the 3′ termini of transcripts that reflect the true position of TEC dissociation. Even transcripts derived from loci encoding putative intrinsic termination sequences should be re-evaluated, as FttA activity may influence the efficiency of intrinsic termination or serve as a back-up mechanism of transcription termination for genes/operons with less efficient intrinsic termination signals.

The requirements for FttA-mediated termination suggest that the long 5′ UTRs observed for some archaeal transcripts may serve as points of regulation for premature termination upstream of the coding sequences[45,56]. It will be interesting to determine how the transcription of stable RNAs, including ribosomal RNAs, is protected from FttA-mediated termination. The exclusion of Spt4–Spt5 from TECs transcribing stable RNAs, or structures within the nascent transcript, may suffice to hinder FttA loading or FttA-mediated termination of archaeal TECs; a delayed mechanism of Spt5 recruitment to rRNA and CRISPR loci has been identified[33]. Full-length FttA homologues are found in the genomes of several bacterial species, suggesting that FttA may function as a termination factor in multiple domains (Supplementary Fig. 1). It will be of immediate interest to determine whether the bacterial FttA proteins can direct transcription termination and, if they can, whether they cooperate with or can substitute for Rho. It will be similarly interesting to determine whether FttA activity can disrupt eukaryotic TECs formed with Pol I, II and III, given that the structure of FttA is nearly identical to the CPSF73 subunit of the eukaryotic CPSF complex (Extended Data Fig. 2)[15,17,25,57,58]. The combined activities of CPSF and Xrn2 are necessary for normal termination patterns in Eukarya[59,60]. FttA retains all the necessary activities within a single protein: FttA can bind TECs, mediate the cleavage and release of the nascent transcript and use 5′–3′ exonuclease activities to degrade the 3′ transcript. We propose that the eukaryotic CPSF complex[61], which contains at least four homologous but non-identical subunits, arose from archaeal FttA. The ability of the CPSF complex to directly terminate transcription was probably lost during specialization and the partnership with factors that direct RNA 3′ maturation in Eukarya.

Methods

***T. kodakarensis* culturing conditions.** *T. kodakarensis* strain TS559 and derivatives of such were grown at 85 °C under anaerobic conditions as previously described[52]. DPA (Sigma) was added at neutral pH to either 12.5 or 25 mM as shown in Fig. 4. NaF was added to 4 mM when necessary.

Protein purifications. Archaeal RNAPs (WT and ΔE/F variant) containing His_6–HA-epitope-tagged RpoL subunits, TBP and TFB were purified as previously described[19,21]. *T. kodakarensis* Spt5 and His_6–Spt4 were purified as previously described[20]. $Spt5^{\Delta NGN}$ was purified as was full-length Spt5. WT and a H255A variant of FttA were purified from Rosetta2 *Escherichia coli* cells carrying pQE-80L (Qiagen) expression vectors with the WT or variant TK1428 coding sequence (Extended Data Fig. 2). Cells were grown in LB medium at 37 °C with shaking (~220 r.p.m.) with 30 μg ml^{-1} chloramphenicol and 100 μg ml^{-1} ampicillin to an optical density at 600 nm of 0.5 before expression was induced with 0.5 mM isopropyl β-D-1-thiogalactopyranoside. Cultures were grown for an additional 3 h at 37 °C with shaking before biomass was harvested via centrifugation (~8,000*g*, 20 min, 4 °C), resuspended and lysed via sonication (3 ml g^{-1} of biomass) in 20 mM Tris-HCl pH 8.0, 5 mM 2-mercaptoethanol, 10 mM $MgCl_2$, 100 mM NaCl. Cellular lysates were clarified by centrifugation (~20,000*g*, 20 min, 4 °C), heated to 85 °C for 30 min to denature most host proteins and clarified again by centrifugation (~20,000*g*, 20 min, 4 °C). Heat-treated clarified cell lysates were resolved through a 5-ml HiTrap-heparin column (GE Healthcare) with a linear gradient from 0.1–1.0 M NaCl dissolved in 20 mM Tris-HCl pH 8.0, 5 mM 2-mercaptoethanol, 10 mM $MgCl_2$. Fractions containing >95% pure FttA were identified by SDS–PAGE, pooled and dialysed into 25 mM Tris-HCl pH 8.0, 100 mM KCl, 10 mM 2-mercaptoethanol, 50% glycerol before storage at −80 °C. All protein concentrations were quantified using a Bradford Assay[62].

DNA templates. The double-stranded DNA templates used in all transcription reactions were PCR amplified from plasmids and gel purified as previously described[19,20,24]. All transcription templates contained a non-template 5′-strand biotin–TEG moiety to provide attachment to streptavidin-coated paramagnetic beads (Promega).

In vitro transcription assays. The assembly of preinitiation complexes and elongation via NTP deprivation was carried out as described previously[19,20,24]. To obtain stalled TECs on G-less cassette templates, preinitiation complexes were assembled using 10 nM template, 20 nM RNAP, 40 nM TBP, 40 nM TFB in a 20 μl total volume of transcription buffer (20 mM Tris-HCl pH 8.0, 250 mM KCl, 5 mM $MgCl_2$, 1 mM DTT) with 75 μM ApC for 3 min at 85 °C before the addition of 200 μM ATP, 200 μM CTP, 10 μM uridine triphosphate (UTP) and 10 μCi [^{32}P-α]-UTP for 3 additional min at 85 °C, then chilled to 4 °C. To obtain stalled TECs on C-less cassette templates, reactions were identical to those above, with the substitution of 200 μM GTP for 200 μM CTP. RNAP-bound templates were captured with HisPur Ni-NTA magnetic particles (Thermo Fisher Scientific) and washed three times with 100 μl 20 mM Tris-HCl pH 8.0, 1 mM EDTA, 500 mM KCl.

For Fig. 1a, washed TECs were resuspended in 10 mM Tris-HCl pH 8.0, 125 mM KCl, 5 mM $MgCl_2$, 1 mM DTT, with 10 μM each of ATP, CTP and UTP before the addition of 1 μM FttA or $FttA^{H255A}$ for 5 min at 85 °C. Reactions were chilled to 4 °C followed by the separation of pellet and supernatant fractions by the addition of streptavidin-coated paramagnetic particles (Promega). Similar results were observed in four independent experiments. For Fig. 1b, washed TECs (lane 1) were resuspended in 10 mM Tris-HCl pH 8.0, 125 mM KCl, 5 mM $MgCl_2$, 1 mM DTT with 10 μM each of ATP, CTP and UTP and were incubated at 85 °C for 7 min (lane 2) before being chilled on ice, bound to streptavidin-coated paramagnetic beads and washed with 100 μl 20 mM Tris-HCl pH 8.0, 1 mM EDTA, 500 mM KCl (lane 3). Washed TECs were incubated at 85 °C for 1 min before the addition of 100 μM NTPs (lane 4) or 100 μM ATP, CTP and GTP, 10 μM UTP containing 1 μC ^{32}P-α-UTP (lane 5) and continued incubation at 85 °C for 3 min. Washed TECs were exposed to 1 μM FttA (lane 6) at 85 °C for 7 min before being chilled on ice, bound to streptavidin-coated paramagnetic beads and washed with 100 μl 20 mM Tris-HCl pH 8.0, 1 mM EDTA, 500 mM KCl (lane 7). FttA-treated, washed TECs were incubated at 85 °C for 1 min before the addition of 100 μM NTPs (lane 8) or 100 μM ATP, CTP and GTP, 10 μM UTP containing 1 μC ^{32}P-α-UTP (lane 9) and continued incubation at 85 °C for 3 min. Washed TECs were exposed to 50 U RNase I_f (lane 10) at 37 °C for 7 min before being chilled on ice, bound to streptavidin-coated paramagnetic beads and washed with 100 μl 20 mM Tris-HCl pH 8.0, 1 mM EDTA, 500 mM KCl (lane 11). RNase I_f-treated, washed TECs were incubated at 85 °C for 1 min before the addition of 100 μM NTPs (lane 12) or 100 μM ATP, CTP and GTP, 10 μM UTP containing 1 μC ^{32}P-α-UTP (lane 13) and continued incubation at 85 °C for 3 min. Similar results were observed in four independent experiments. For Fig. 1f, washed TECs were resuspended in 10 mM Tris-HCl pH 8.0, 125 mM KCl, 5 mM $MgCl_2$, 1 mM DTT, with or without 10 μM each of ATP, CTP and UTP before the addition of 1 μM FttA at 85 °C. Reaction aliquots were removed after 1, 2 or 5 min, chilled to 4 °C and then pellet and supernatant fractions were separated by the addition of streptavidin-coated paramagnetic particles (Promega). Similar results were observed in four independent experiments.

For Fig. 2b,c and Extended Data Fig. 3, washed TECs were assembled as above on G-less or C-less cassettes of various lengths. Washed TECs were resuspended in 10 mM Tris-HCl pH 8.0, 125 mM KCl, 5 mM $MgCl_2$, 1 mM DTT, with or without 1 μM FttA at 85 °C for 3 min. Reactions were chilled to 4 °C, then pellet and supernatant fractions were separated by the addition of streptavidin-coated paramagnetic particles (Promega). For Fig. 2b,c and Extended Data Fig. 3, similar results were observed in three independent experiments.

For Fig. 3 and Extended Data Fig. 4, washed TECs were assembled with WT or ΔE/F RNAP as above on a +125-nt G-less cassette template. Continued elongation was permitted by the addition of 0, 1, 10 or 100 μM NTPs in the presence or absence of combinations of 6 μM Spt4, 6 μM Spt5, 6 μM $Spt5^{\Delta NGN}$, 1 μM FttA or $FttA^{H255A}$. After 5 min at 85 °C, reactions were chilled to 4 °C followed by the separation of pellet and supernatant fractions by the addition of streptavidin-coated paramagnetic particles (Promega). For Fig. 3c, similar results were observed in three independent experiments. Extended Data Fig. 4 was performed once.

For Fig. 4b, stalled TECs on a G-less cassette template were assembled using 10 nM template, 20 nM RNAP, 40 nM TBP, 40 nM TFB in a 20-μl total volume of transcription buffer (20 mM Tris-HCl pH 8.0, 250 mM KCl, 5 mM $MgCl_2$, 1 mM DTT) with 75 μM ApC for 3 min at 85 °C before the addition of 200 μM ATP, 200 μM CTP, 10 μM UTP and 10 μCi [α-^{32}P]-UTP and incubation for 3 additional min at 85 °C followed by chilling to 4 °C. RNAP-bound templates were captured with HisPur Ni-NTA magnetic particles (Thermo Fisher Scientific) and washed three times with 100 μl 20 mM Tris-HCl pH 8.0, 1 mM EDTA, 500 mM KCl. Washed TECs were resuspended in 10 mM Tris-HCl pH 8.0, 125 mM KCl, 5 mM $MgCl_2$, 1 mM DTT, with 10 μM each of ATP, CTP and UTP before the addition of reaction buffer (10 mM Tris-HCl pH 8.0, 120 mM KCl, 8 mM DTT and 1.25 mM $MgCl_2$) ± 25 mM DPA and/or ± 1 μM FttA at 85 °C. Reaction aliquots were incubated for 3 min and then chased with 250 μM ATP, CTP, UTP and GTP for 2 min to allow for elongation to +225 nt. Reactions were chilled to 4 °C, then the pellet and supernatant fractions were separated by the addition of streptavidin-coated paramagnetic particles (Promega). Similar results were observed in five independent experiments.

Radiolabelled transcripts from Figs. 1–4, Extended Data Fig. 3 and Extended Data Fig. 4 were recovered by the addition of five volumes of STOP buffer (600 mM Tris-HCl pH 8.0, 30 mM EDTA) containing 7 μg of transfer RNA (total) and equal volume phenol/chloroform/isoamyl alcohol (25:24:1, v/v/v) extractions. RNA was recovered by precipitations of the aqueous phase with 2.6 volumes of 100% ethanol. Precipitated transcripts were resuspended in 95% formamide, 1× TBE, heated to 99 °C for 2 min, rapidly chilled on ice, loaded and resolved in 10–20% polyacrylamide, 8 M urea, 1× TBE denaturing gels. Radiolabelled RNA was detected using phosphorimaging (GE Healthcare). Gel images were analysed using GE Imagequant 5.2.

To generate the +125-nt [α-^{32}P]-UTP-labelled transcripts used in Extended Data Fig. 3, transcription reactions were assembled and terminated as above, with the substitution of 10 μg of glycogen for 7 μg of tRNA during reaction clean-up. Radiolabelled transcripts were incubated at 85 °C with or without 1 μM FttA or at 37 °C with 5 U RNaseA (Thermo Fisher Scientific) in *T. kodakarensis* transcription buffer (20 mM Tris-HCl pH 8.0, 250 mM KCl, 5 mM $MgCl_2$, 1 mM DTT) to monitor cleavage. When present, *T. kodakarensis* RNAP was added to a final concentration of 40 nM.

Western blot analysis of RNAP release to solution . Anti-HA antibodies (BioLegend 901513) were employed as previously described[10,19,52] to quantify RpoL levels in the pellet and supernatant fractions (Fig. 1e).

Western blot analysis of FttA protein levels. Purified, recombinant full-length FttA was used as an antigen to prepare polyclonal antibodies in mice (Cocalico Biologicals) (Fig. 4h and Extended Data Fig. 5). Known quantities of purified FttA were resolved as comparative quantification standards in adjacent lanes to clarified cell lysates derived from known quantities of cells. Proteins were separated via SDS–PAGE, transferred to PVDF membranes and probed with primary anti-FttA antibodies. Blots were developed by the addition of an IgG-AP conjugated antimouse secondary antibody allowing for detection by NBT/BCIP (Roche). A linear regression FttA signal intensity to FttA amount (in ng) was generated.

Construction of strains TK1428D and IR5. The *T. kodakarensis* strains used here were constructed from the parental strain TS559 (ref. [51]) as previously described[52].

Purification of FttA directly from lysates of strain TK1428D and MuDPIT analysis. Purification procedures and MuDPIT analysis were performed largely as previously described[10,53,54]. Briefly, 2 l of early exponential phase cultures (optical density at 600 nm = ~0.3) of *T. kodakarensis* strains TK1428D and TS559 grown in ASW-YT-S° medium at 85 °C were rapidly chilled to 4 °C and harvested by centrifugation (20,000*g*). All subsequent procedures were carried out at 4 °C and completed as rapidly as possible to retain native in vivo protein–protein interactions. Cells from each strain were individually resuspended in 3 ml of 25 mM Tris-HCl pH 8.0, 500 mM NaCl, 10% glycerol per gram of wet biomass and lysed by repetitive sonication. The resulting lysates were clarified via centrifugation (20,000*g*) before passage through 5-ml HiTRAP chelating columns (GE Healthcare) precharged with $NiSO_4$ and equilibrated in 25 mM Tris-HCl pH 8.0, 500 mM NaCl, 10% glycerol. The lysate components that did not bind the columns (for example, the flowthrough) were discarded and the columns were washed with ~20 column volumes 25 mM Tris-HCl pH 8.0, 500 mM NaCl, 10% glycerol until no additional proteins were eluted. Bound proteins were eluted using a linear gradient from (an initial) 25 mM Tris-HCl pH 8.0, 500 mM NaCl, 10% glycerol to (a final) 25 mM Tris-HCl pH 8.0, 100 mM NaCl, 150 mM imidazole, 10% glycerol. Fractions that contained the tagged TK1428 protein were identified by western blotting (Extended Data Fig. 6), pooled and dialysed twice in 3-kDa molecular weight cut-off dialysis tubing against 1 l 25 mM Tris-HCl pH 8.0, 500 mM NaCl, 0.5 mM EDTA, 2 mM dithiothreitol. FttA was not identified by western blots within fractions that resulted from processing TS559 biomass, but identical fractions were collected and processed to identify native *T. kodakarensis* present in such fractions that spontaneously bind the chelating resin. Protein concentrations were determined by Bradford assays[62] and 30-μg aliquots of the proteins present in solution were precipitated by adding trichloroacetic acid (15% final concentration). The trichloroacetic acid-precipitated proteins were identified by multidimensional protein identification technology at the Ohio State University mass spectrometry facility (https://www.ccic.osu.edu/MSP) using the MASCOT search engine. We required a minimum of two unique peptide fragments to positively identify a protein. TS559 control samples identified several proteins that bound and eluted from the Ni^{2+}-charged matrix in the absence of a His_6-tagged protein. All of the proteins identified in the experimental samples that had MASCOT scores of >100 and were not also present in the control samples are shown in Extended Data Fig. 7.

TRIzol-based RNA purifications from *T. kodakarensis* cultures and qRT–PCR. RNA extractions were performed essentially as previously described[45] from strain TS559 before or 1 h after DPA addition. RNA extractions from IR5 and TS559, in the absence and presence of 4 mM NaF, were performed as previously described[45]. qRT–PCR reactions were performed as previously described[63], except that 500 ng of total RNA was used during cDNA synthesis.

Sequences of DNA templates employed in transcription reactions. For all DNA templates, the BRE and TATA sequences are shown in capital text and italicized, transcription start sites are in capital bold text, transcribed regions are underlined and positions used to stall TECs by NTP deprivation are shown in bold. The critical parameters of each template (template name, G-less or C-less cassette length and the length of full-length run-off transcripts) are highlighted above the text of each DNA template:

TJS0 125-nt G-less stalled transcript, 175-nt run-off transcript. cgcgcgtaatacgactcactatagggGCGAtataTTTATATAgggatatagtaatagataaatatc**A**catctctcttctcaaatattctacttattactcaacaccatctaccttccactccatcctacctttccacatctctcttctcaccctccctctccaccaactacttccacttcaccatcaaccat**ggg**tcctgttccgttggaccgtggtttccaggcggtgcgtctgctgtcgg

TJS1 125-nt G-less stalled transcript, 225-nt run-off transcript. cgcgcgtaatacgactcactatagggGCGAtataTTTATATAgggatatagtaatagataaatatc**A**catctctcttctcaaatattctacttattactcaacaccatctaccttccactccatcctacctttccacatctctcttctcaccctccctctccaccaactacttccacttccaccatcaaccat**ggg**tcctgttccgttggaccgtggtttccaggcggtgcgtctgctgtcgggtccctctggtgactggctagatggacccttttagtgagggttaattgcgg

TJS2 87-nt G-less stalled transcript, 187-nt run-off transcript. cgcgcgtaatacgactcactatagggGCGAtataTTTATATAgggatatagtaatagataatatc**A**catctctcttctcaaatattctacttattactcaacaccatctaccttccactccatcctaccttccacatctctcttctcaccct**ggg**tcctgttccgttggaccgtggtttccaggcggtgcgtctgctgtcgggtccctctggtgactggctagatggacccttttagtgagggttaattgcgg

TJS3 69-nt G-less stalled transcript, 169-nt run-off transcript. cgcgcgtaatacgactcactatagggGCGAtataTTTATATAgggatatagtaatagataaatatc**A**catctctcttctcaaatattctacttattactcaacaccatctaccttccactccatcctaccttcca**ggg**tcctgttccgttggaccgtggtttccaggcggtgcgtctgctgtcgggtccctctggtgactggctagatggacccttttagtgagggttaattgcgg

TJS4 54-nt G-less stalled transcript, 154-nt run-off transcript. cgcgcgtaatacgactcactatagggGCGAtataTTTATATAgggatatagtaatagataaatatc**A**catctctcttctcaaatattctacttattactcaacaccatctaccttccact**ggg**tcctgttccgttggaccgtggtttccaggcggtgcgtctgctgtcgggtccctctggtgactggctagatggacccttttagtgagggttaattgcgg

TJS5 125-nt C-less stalled transcript, 180-nt run-off transcript. cgcgcgtaatacgactcactatagggGCGAtataTTTATATAgggatatagtaatagataaatatc**A**catgtgtgttgtgaaatattgtagttggatgtaggttggagtggagtgtgtggtgtgtgtatggtgatagattaagtgttaatgtggttgtagataaagttaatgtatattggtaatatgagtt**ccc**tcctgttcgtcgccgttggaccgtggttccctttagtgagggttaattgcgg

TJS6 98-nt C-less stalled transcript, 153-nt run-off transcript. cgcgcgtaatacgactcactatagggGCGAtataTTTATATAgggatatagtaatagataaatatc**A**catgtgtgttgtgaaatattgtagttggatgtaggttggagtggagtgtgtggtgtgtgtatggtgatagattaagtgttaatgtggttgtagataa**ccc**tcctgttcgtcgccgttggaccgtggttccctttagtgagggttaattgcgg

TJS7 78-nt C-less stalled transcript, 133-nt run-off transcript. cgcgcgtaatacgactcactatagggGCGAtataTTTATATAgggatatagtaatagataaatatc**A**catgtgtgttgtgaaatattgtagttggatgtaggttggagtggagtgtgtggtgtgtgtatggtgatagattaagt**ccc**tcctgttcgtcgccgttggaccgtggttccctttagtgagggttaattgcgg

TJS8 58-nt C-less stalled transcript, 113-nt run-off transcript. cgcgcgtaatacgactcactatagggGCGAtataTTTATATAgggatatagtaatagataaatatc**A**catgtgtgttgtgaaatattgtagttggatgtaggttggagtggagtgtgtggtgtg**ccc**tcctgttcgtcgccgttggaccgtggttccctttagtgagggttaattgcgg

Reporting Summary. Further information on research design is available in the Nature Research Reporting Summary linked to this article.

Data availability

The data that support the findings of this study are available from the authors on reasonable request. Source data for Figs. 1e and 4h, as well as Extended Data Figs. 2a, 3d and 5 are included in this article and its Supplementary Information files.

Received: 30 August 2019; Accepted: 6 January 2020;
Published online: 24 February 2020

References

1. Maier, L.-K. & Marchfelder, A. It's all about the T: transcription termination in archaea. *Biochem. Soc. Trans.* **47**, 461–468 (2019).
2. Santangelo, T. J. & Reeve, J. N. Archaeal RNA polymerase is sensitive to intrinsic termination directed by transcribed and remote sequences. *J. Mol. Biol.* **355**, 196–210 (2006).
3. Sela, I., Wolf, Y. I. & Koonin, E. V. Theory of prokaryotic genome evolution. *Proc. Natl Acad. Sci. USA* **113**, 11399–11407 (2016).
4. Helmrich, A., Ballarino, M., Nudler, E. & Tora, L. Transcription-replication encounters, consequences and genomic instability. *Nat. Struct. Mol. Biol.* **20**, 412–418 (2013).
5. Washburn, R. S. & Gottesman, M. E. Transcription termination maintains chromosome integrity. *Proc. Natl Acad. Sci. USA* **108**, 792–797 (2011).
6. Shin, J.-H., Santangelo, T. J., Xie, Y., Reeve, J. N. & Kelman, Z. Archaeal minichromosome maintenance (MCM) helicase can unwind DNA bound by archaeal histones and transcription factors. *J. Biol. Chem.* **282**, 4908–4915 (2007).
7. Miller, O. L., Hamkalo, B. A. & Thomas, C. A. Visualization of bacterial genes in action. *Science* **169**, 392–395 (1970).
8. French, S. L., Santangelo, T. J., Beyer, A. L. & Reeve, J. N. Transcription and translation are coupled in Archaea. *Mol. Biol. Evol.* **24**, 893–895 (2007).
9. Ray-Soni, A., Bellecourt, M. J. & Landick, R. Mechanisms of bacterial transcription termination: all good things must end. *Annu. Rev. Biochem.* **85**, 319–347 (2016).
10. Walker, J. E., Luyties, O. & Santangelo, T. J. Factor-dependent archaeal transcription termination. *Proc. Natl Acad. Sci. USA* **114**, E6767–E6773 (2017).
11. Santangelo, T. J. et al. Polarity in archaeal operon transcription in *Thermococcus kodakaraensis*. *J. Bacteriol.* **190**, 2244–2248 (2008).
12. D'Heygere, F., Rabhi, M. & Boudvillain, M. Phyletic distribution and conservation of the bacterial transcription termination factor Rho. *Microbiology* **159**, 1423–1436 (2013).
13. Makarova, K., Wolf, Y. & Koonin, E. Archaeal clusters of orthologous genes (arCOGs): an update and application for analysis of shared features between Thermococcales, Methanococcales, and Methanobacteriales. *Life* **5**, 818–840 (2015).
14. Wolf, Y. I., Makarova, K. S., Yutin, N. & Koonin, E. V. Updated clusters of orthologous genes for Archaea: a complex ancestor of the Archaea and the byways of horizontal gene transfer. *Biol. Direct* **7**, 46 (2012).
15. Phung, D. K. et al. Archaeal β-CASP ribonucleases of the aCPSF1 family are orthologs of the eukaryal CPSF-73 factor. *Nucleic Acids Res.* **41**, 1091–1103 (2013).
16. Mandel, C. R. et al. Polyadenylation factor CPSF-73 is the pre-mRNA 3′-end-processing endonuclease. *Nature* **444**, 953–956 (2006).
17. Nishida, Y. et al. Crystal structure of an archaeal cleavage and polyadenylation specificity factor subunit from *Pyrococcus horikoshii*. *Proteins* **78**, 2395–2398 (2010).
18. Galperin, M. Y., Kristensen, D. M., Makarova, K. S., Wolf, Y. I. & Koonin, E. V. Microbial genome analysis: the COG approach. *Brief. Bioinform.* **20**, 1063–1070 (2017).
19. Santangelo, T. J., Cubonová, L., James, C. L. & Reeve, J. N. TFB1 or TFB2 is sufficient for *Thermococcus kodakaraensis* viability and for basal transcription in vitro. *J. Mol. Biol.* **367**, 344–357 (2007).
20. Sanders, T. J. et al. TFS and Spt4/5 accelerate transcription through archaeal histone-based chromatin. *Mol. Microbiol.* **111**, 784–797 (2019).
21. Hirata, A. et al. Archaeal RNA polymerase subunits E and F are not required for transcription in vitro, but a *Thermococcus kodakarensis* mutant lacking subunit F is temperature-sensitive. *Mol. Microbiol.* **70**, 623–633 (2008).
22. Santangelo, T. J. & Reeve, J. N. Deletion of switch 3 results in an archaeal RNA polymerase that is defective in transcript elongation. *J. Biol. Chem.* **285**, 23908–23915 (2010).
23. Gehring, A. M. & Santangelo, T. J. Archaeal RNA polymerase arrests transcription at DNA lesions. *Transcription* **8**, 288–296 (2017).
24. Gehring, A. M. & Santangelo, T. J. in *Bacterial Transcriptional Control. Methods in Molecular Biology* Vol. 1276 (eds Artsimovitch, I. & Santangelo, T.) 263–279 (Humana Press, 2015).
25. Mir-Montazeri, B., Ammelburg, M., Forouzan, D., Lupas, A. N. & Hartmann, M. D. Crystal structure of a dimeric archaeal cleavage and polyadenylation specificity factor. *J. Struct. Biol.* **173**, 191–195 (2011).
26. Kolev, N. G., Yario, T. A., Benson, E. & Steitz, J. A. Conserved motifs in both CPSF73 and CPSF100 are required to assemble the active endonuclease for histone mRNA 3′-end maturation. *EMBO Rep.* **9**, 1013–1018 (2008).
27. Orlova, M., Newlands, J., Das, A., Goldfarb, A. & Borukhov, S. Intrinsic transcript cleavage activity of RNA polymerase. *Proc. Natl Acad. Sci. USA* **92**, 4596–4600 (1995).
28. Dengl, S. & Cramer, P. *Torpedo* nuclease Rat1 is insufficient to terminate RNA polymerase II in vitro. *J. Biol. Chem.* **284**, 21270–21279 (2009).
29. Phung, D. K. & Clouet-d'Orval, B. in *RNA Remodeling Proteins. Methods in Molecular Biology* Vol. 1259 (ed. Boudvillain, M.) 453–466 (Humana Press, 2015).
30. Silva, A. P. G. et al. Structure and activity of a novel archaeal β-CASP protein with N-terminal KH domains. *Structure* **19**, 622–632 (2011).
31. Ray, W. C. & Daniels, C. J. PACRAT: a database and analysis system for archaeal and bacterial intergenic sequence features. *Nucleic Acids Res.* **31**, 109–113 (2003).
32. Gehring, A. M., Walker, J. E. & Santangelo, T. J. Transcription regulation in Archaea. *J. Bacteriol.* **198**, 1906–1917 (2016).
33. Smollett, K., Blombach, F., Reichelt, R., Thomm, M. & Werner, F. A global analysis of transcription reveals two modes of Spt4/5 recruitment to archaeal RNA polymerase. *Nat. Microbiol.* **2**, 17021 (2017).
34. Cardinale, C. J. et al. Termination factor Rho and its cofactors NusA and NusG silence foreign DNA in *E. coli*. *Science* **320**, 935–938 (2008).
35. Cortazar, M. A. et al. Control of RNA Pol II speed by PNUTS-PP1 and Spt5 dephosphorylation facilitates termination by a "sitting duck torpedo" mechanism. *Mol. Cell* **76**, 896–908 (2019).
36. Lawson, M. R. & Berger, J. M. Tuning the sequence specificity of a transcription terminator. *Curr. Genet.* **65**, 729–733 (2019).
37. Lawson, M. R. et al. Mechanism for the regulated control of bacterial transcription termination by a universal adaptor protein. *Mol. Cell* **71**, 911–922 (2018).
38. Mitra, P., Ghosh, G., Hafeezunnisa, M. & Sen, R. Rho protein: roles and mechanisms. *Annu. Rev. Microbiol.* **71**, 687–709 (2017).
39. Burmann, B. M. et al. A NusE:NusG complex links transcription and translation. *Science* **328**, 501–504 (2010).
40. Werner, F. & Grohmann, D. Evolution of multisubunit RNA polymerases in the three domains of life. *Nat. Rev. Microbiol.* **9**, 85–98 (2011).
41. Nagy, J. et al. Complete architecture of the archaeal RNA polymerase open complex from single-molecule FRET and NPS. *Nat. Commun.* **6**, 6161 (2015).
42. Plaschka, C. et al. Architecture of the RNA polymerase II–Mediator core initiation complex. *Nature* **518**, 376–380 (2015).
43. Walker, J. E. & Santangelo, T. J. Analyses of in vivo interactions between transcription factors and the archaeal RNA polymerase. *Methods* **86**, 73–79 (2015).
44. Horsfall, L. E. et al. Competitive inhibitors of the CphA metallo-β-lactamase from *Aeromonas hydrophila*. *Antimicrob. Agents Chemother.* **51**, 2136–2142 (2007).
45. Jäger, D., Förstner, K. U., Sharma, C. M., Santangelo, T. J. & Reeve, J. N. Primary transcriptome map of the hyperthermophilic archaeon *Thermococcus kodakarensis*. *BMC Genomics* **15**, 684 (2014).
46. Speed, M. C., Burkhart, B. W., Picking, J. W. & Santangelo, T. J. An archaeal fluoride-responsive riboswitch provides an inducible expression system for hyperthermophiles. *Appl. Environ. Microbiol.* **84**, e02306–e02317 (2018).
47. Sarmiento, F., Mrázek, J. & Whitman, W. B. Genome-scale analysis of gene function in the hydrogenotrophic methanogenic archaeon *Methanococcus maripaludis*. *Proc. Natl Acad. Sci. USA* **110**, 4726–4731 (2013).
48. Zhang, C., Phillips, A. P. R., Wipfler, R. L., Olsen, G. J. & Whitaker, R. J. The essential genome of the crenarchaeal model *Sulfolobus islandicus*. *Nat. Commun.* **9**, 4908 (2018).
49. Garas, M., Dichtl, B. & Keller, W. The role of the putative 3′ end processing endonuclease Ysh1p in mRNA and snoRNA synthesis. *RNA* **14**, 2671–2684 (2008).
50. Ryan, K., Calvo, O. & Manley, J. L. Evidence that polyadenylation factor CPSF-73 is the mRNA 3′ processing endonuclease. *RNA* **10**, 565–573 (2004).
51. Santangelo, T. J., Cubonová, L. & Reeve, J. N. *Thermococcus kodakarensis* genetics: TK1827-encoded β-glycosidase, new positive-selection protocol, and targeted and repetitive deletion technology. *Appl. Environ. Microbiol.* **76**, 1044–1052 (2010).
52. Gehring, A., Sanders, T. & Santangelo, T. J. Markerless gene editing in the hyperthermophilic archaeon *Thermococcus kodakarensis*. *Bio. Protoc.* **7**, e2604 (2017).
53. Li, Z., Santangelo, T. J., Cuboňová, L., Reeve, J. N. & Kelman, Z. Affinity purification of an archaeal DNA replication protein network. *mBio* **1**, e00221-10 (2010).

54. Burkhart, B. W., Febvre, H. P. & Santangelo, T. J. Distinct physiological roles of the three ferredoxins encoded in the hyperthermophilic archaeon *Thermococcus kodakarensis*. *mbio* **10**, e02807-18 (2019).
55. Dar, D., Prasse, D., Schmitz, R. A. & Sorek, R. Widespread formation of alternative 3′ UTR isoforms via transcription termination in archaea. *Nat. Microbiol.* **1**, 16143 (2016).
56. Cohen, O. et al. Comparative transcriptomics across the prokaryotic tree of life. *Nucleic Acids Res.* **44**, W46–W53 (2016).
57. Sun, Y. et al. Molecular basis for the recognition of the human AAUAAA polyadenylation signal. *Proc. Natl Acad. Sci. USA* **115**, E1419–E1428 (2018).
58. Clerici, M., Faini, M., Muckenfuss, L. M., Aebersold, R. & Jinek, M. Structural basis of AAUAAA polyadenylation signal recognition by the human CPSF complex. *Nat. Struct. Mol. Biol.* **25**, 135–138 (2018).
59. Eaton, J. D. et al. Xrn2 accelerates termination by RNA polymerase II, which is underpinned by CPSF73 activity. *Genes Dev.* **32**, 127–139 (2018).
60. Proudfoot, N. J. Transcriptional termination in mammals: stopping the RNA polymerase II juggernaut. *Science* **352**, aad9926 (2016).
61. Hill, C. H. et al. Activation of the endonuclease that defines mRNA 3′ ends requires incorporation into an 8-subunit core cleavage and polyadenylation factor complex. *Mol. Cell* **73**, 1217–1231 (2019).
62. Bradford, M. M. A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. *Anal. Biochem.* **72**, 248–254 (1976).
63. Mattiroli, F. et al. Structure of histone-based chromatin in Archaea. *Science* **357**, 609–612 (2017).
64. Roberts, J. W. Mechanisms of bacterial transcription termination. *J. Mol. Biol.* **43**, 4030–4039 (2019).
65. Mishra, S. & Maraia, R. J. RNA polymerase III subunits C37/53 modulate rU:dA hybrid 3′ end dynamics during transcription termination. *Nucleic Acids Res.* **47**, 310–327 (2019).
66. Sanders, T. J., Marshall, C. J. & Santangelo, T. J. The role of archaeal chromatin in transcription. *J. Mol. Biol.* **431**, 4103–4115 (2019).
67. Dar, D. et al. Term-seq reveals abundant ribo-regulation of antibiotics resistance in bacteria. *Science* **352**, aad9822 (2016).
68. Spitalny, P. & Thomm, M. A polymerase III-like reinitiation mechanism is operating in regulation of histone expression in archaea. *Mol. Microbiol.* **67**, 958–970 (2008).
69. Blombach, F., Matelska, D., Fouqueau, T., Cackett, G. & Werner, F. Key concepts and challenges in archaeal transcription. *J. Mol. Biol.* **431**, 4184–4201 (2019).
70. Shashni, R., Qayyum, M. Z., Vishalini, V., Dey, D. & Sen, R. Redundancy of primary RNA-binding functions of the bacterial transcription terminator Rho. *Nucleic Acids Res.* **42**, 9677–9690 (2014).
71. Valabhoju, V., Agrawal, S. & Sen, R. Molecular basis of NusG-mediated regulation of Rho-dependent transcription termination in bacteria. *J. Biol. Chem.* **291**, 22386–22403 (2016).
72. Peters, J. M. et al. Rho and NusG suppress pervasive antisense transcription in *Escherichia coli*. *Genes Dev.* **26**, 2621–2633 (2012).
73. Fong, N. et al. Effects of transcription elongation rate and Xrn2 exonuclease activity on RNA polymerase II termination suggest widespread kinetic competition. *Mol. Cell* **60**, 256–267 (2015).

Acknowledgements

We thank members of the Santangelo laboratory for critical review of the manuscript. This work was supported by the National Institutes of Health grant no. GM100329 (to T.J. Santangelo).

Author contributions

T.J. Sanders performed in vitro transcription and RNase digestions, designed and purified FttA variants, designed and assisted in the construction of *T. kodakarensis* strain IR5, purified proteins and templates, generated qRT–PCR data, analysed data, built structural models, performed western blots, prepared figures and helped write the manuscript. B.R.W. performed in vitro transcription and RNase digestions, designed templates and FttA variants, manipulated TK1428 genomic sequences and prepared figures. J.N.S. generated FttA variants, purified transcription proteins, performed bioinformatic analyses, manipulated TK1428 genomic sequences to generate strain TK1428D, built structural models and prepared figures. M.P.B. and S.A.T. manipulated TK1428 genomic sequences, generated and analysed qRT–PCR data and assisted with western blots and protein purifications. J.E.W. purified FttA variants and performed in vitro transcription. T.J. Santangelo conceived and directed the project, wrote the manuscript, analysed data and prepared figures.

Competing interests

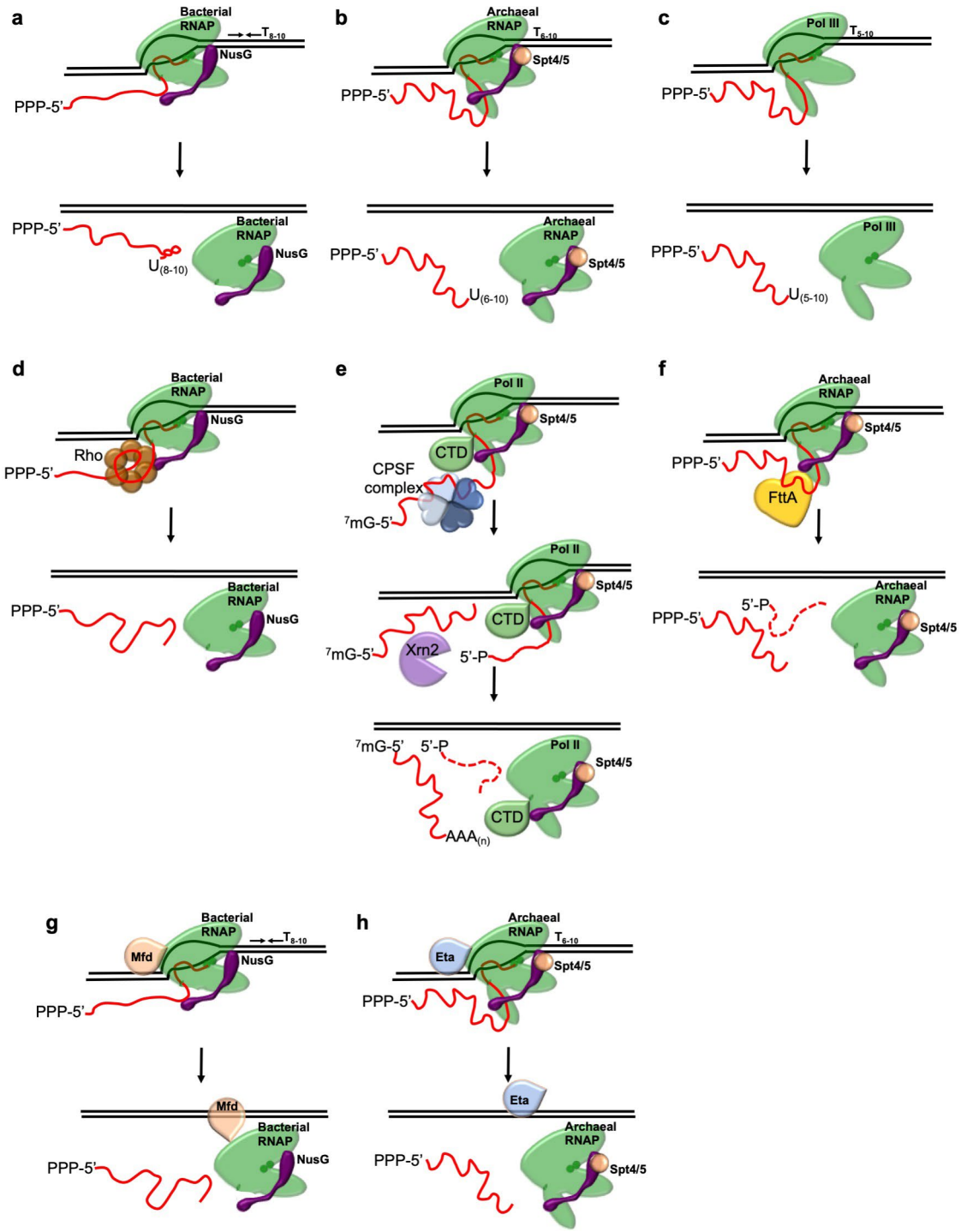
The authors declare no competing interests.

Additional information

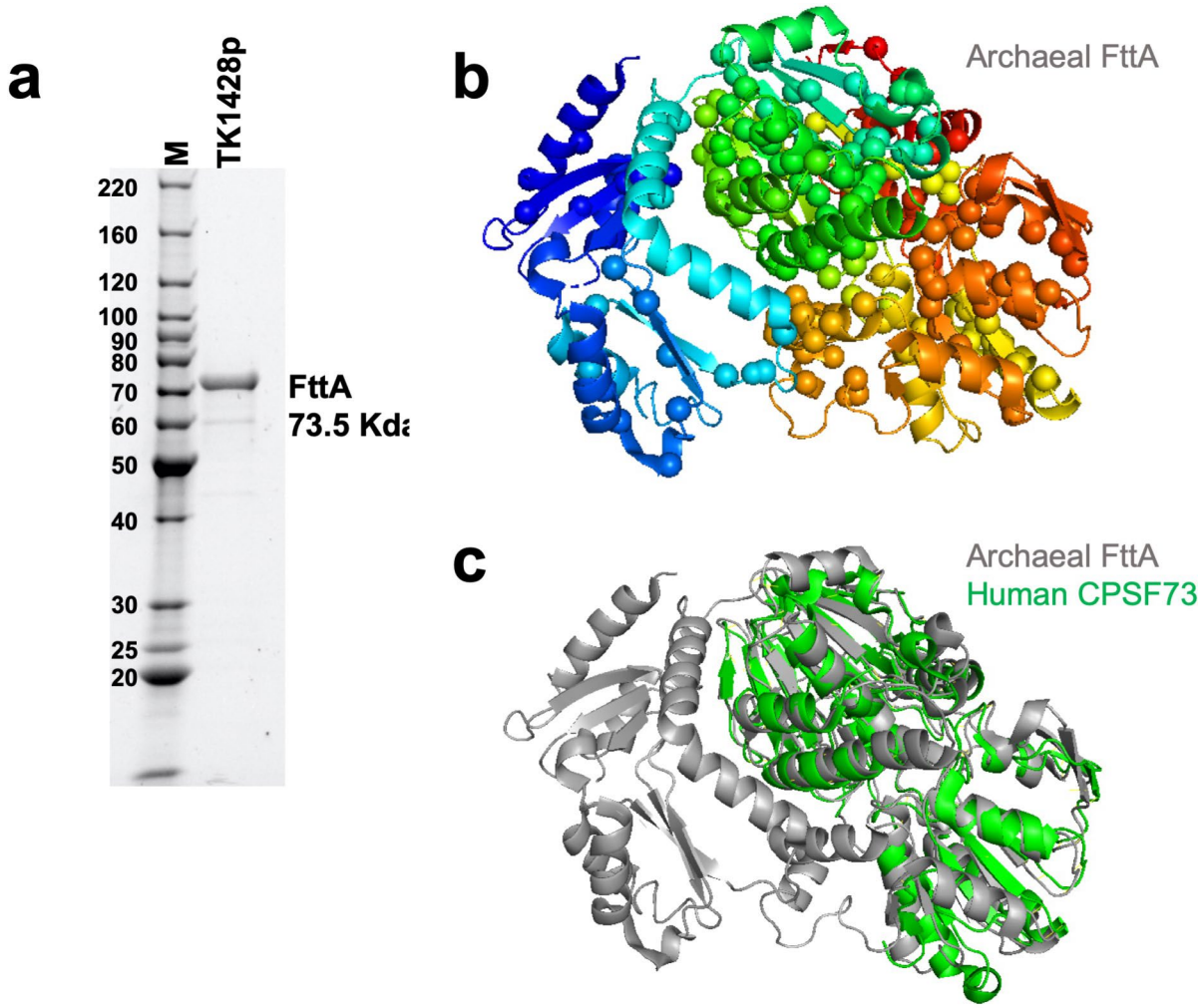
Extended data is available for this paper at https://doi.org/10.1038/s41564-020-0667-3.

Supplementary information is available for this paper at https://doi.org/10.1038/s41564-020-0667-3.

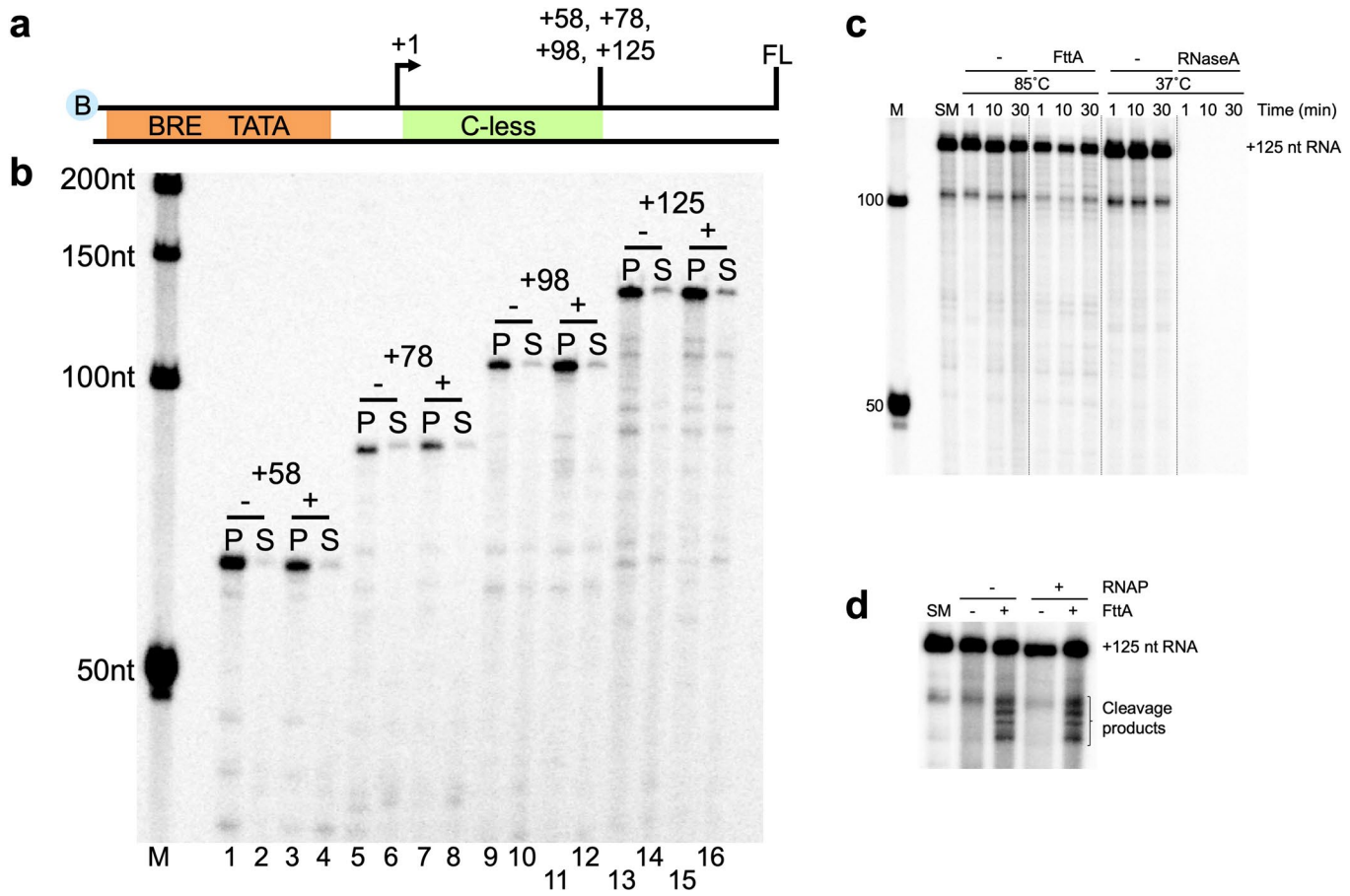
Correspondence and requests for materials should be addressed to T.J.S.



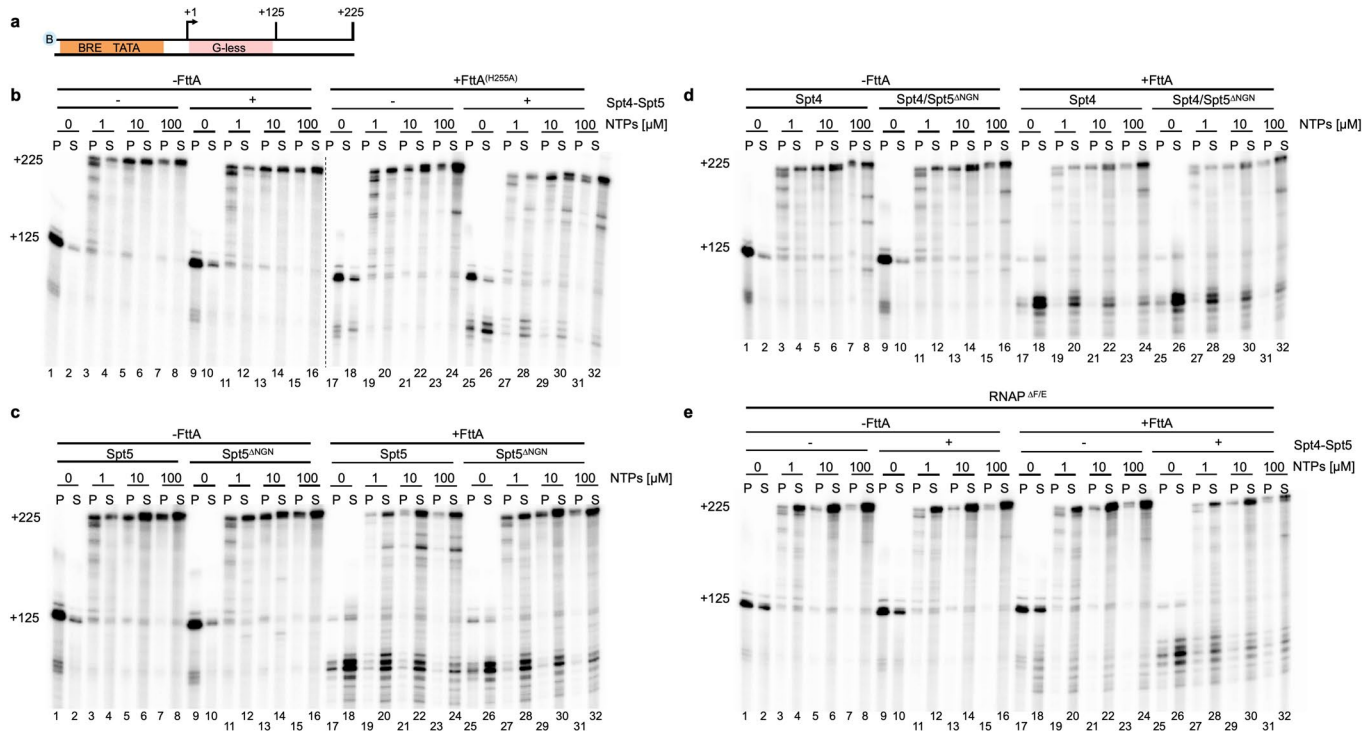
Extended Data Fig. 1 | Transcription termination mechanisms commonly employed in Bacteria, Eukarya and Archaea. Intrinsic transcription termination in Bacteria (**a**), Archaea (**b**), and for eukaryotic Pol III (**c**) results in release of the entire 5′-triphosphate-containing RNA transcript following transcription through a region of dyad-symmetry encoding an RNA hairpin immediately proceeded a T-rich non-template strand sequence (Bacteria)[64] or T-rich non-template strand sequences (Archaea and eukaryotic Pol III)[20,32,65–69]. **d**, **g**, Factor-mediated bacterial transcription termination[64], driven by rho or Mfd, also directs release of the entire nascent transcript and results in collapse of the TEC and recycling of RNAP. Rho-mediated termination is aided by NusG (Spt5 in Eukarya and Archaea)[34,36,70–72]. **e**, Release of the majority of the nascent transcript cannot be considered a *bona fide* termination event in-of-itself. RNA processing events, such as the endonucleolytic cleavage of the nascent RNA within eukaryotic Pol II TECs by the cleavage and polyadenylation factor complex (CPSF)[16,25,50,57,61] yield a 5′-fragment that is often further processed – typically by the addition of a 3′-polyA tail for many Pol II transcripts – but also a 3′-fragment that is encapsulated within a still-stable TEC;[28,60] the combined activities of CPSF and Xrn2 are necessary for normal termination patterns in Eukarya[59,60,73]. **f**, FttA can cleave the nascent transcript and terminate the archaeal transcription apparatus. **g**, **h**, Both bacterial Mfd and archaeal Eta can disrupt stalled TECs and release full-length transcripts by rewinding the transcription bubble.



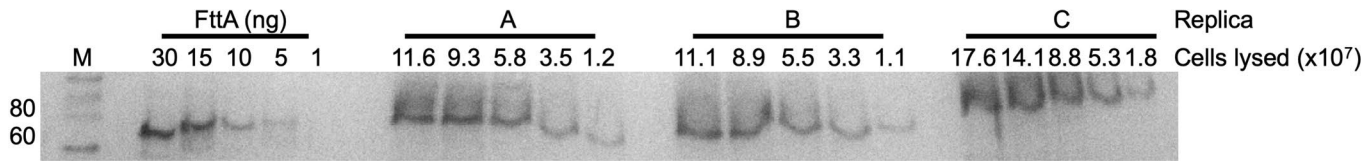
Extended Data Fig. 2 | FttA is highly conserved and shares structural and sequence homology with eukaryotic CPSF73. Recombinant FttA (the protein product of *T. kodakarensis* TK1428; TK1428p) is 73.5 Kda, 85˚C thermotolerant and free of contaminating proteins. Lane M contains molecular weight standards in Kda. Data shown are from a single experiment. **b**, The crystal structure of FttA from *Pyrococcus horikoshii* (PDB: 3AF5)[17] shown in chainbow coloring (N-terminus in purple to C-terminus in red) reveals two N-terminal KH-domains attached by a linker to a C-terminal MBL fold. Alpha-carbons of highly conserved residues in archaeal FttA homologues are shown in colored spheres. **c**, The MBL-fold of FttA is nearly structural identical to the MBL-fold of the human CPSF73 protein (PDB: 2I7T)[16].



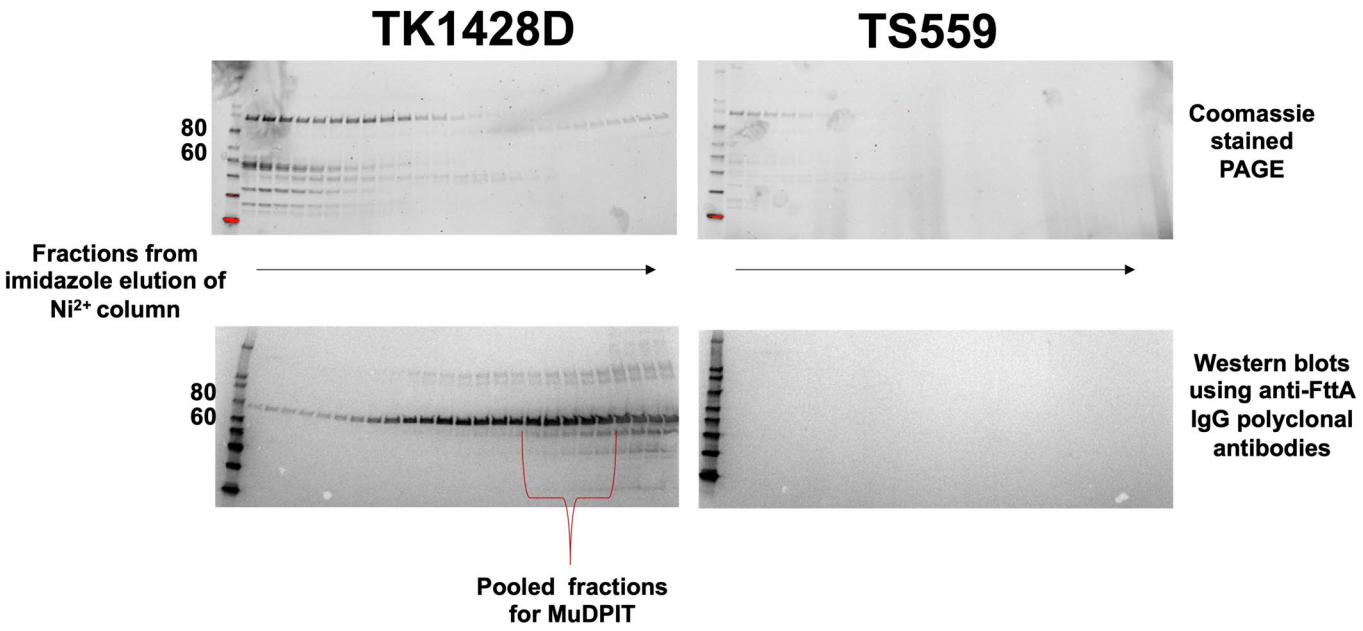
Extended Data Fig. 3 | The RNA cleavage activity of FttA is stimulated by interactions with the archaeal TEC and FttA-mediated termination prefers C-rich transcripts. Promoter-directed transcription of biotinylated templates encoding a C-less cassette permits formation of TECs with increasing length A-, G-, and U-rich nascent transcripts, respectively. FL = full-length; all templates permit elongation for 100 nts beyond the C-less cassette. **b**, TECs remain stably associated and transcripts are primarily recovered in the pellet (P) fraction in the absence (-) of FttA. When FttA is present (+), but nascent transcripts are devoid of CMP, minimal FttA-mediated transcript cleavage or termination occurs, and transcripts are not released to the supernatant (S). Lane M contains ^{32}P-labeled ssDNA markers. Similar results were observed in 3 independent experiments and quantified (n = 3) in Fig. 2d. **c**, FttA demonstrates minimal RNase activity on an isolated +125 nt transcript. Control reactions with RNaseA demonstrate that the purified transcript is not resistant to the activity of RNases. Similar results were observed in 2 independent experiments. **d**, Addition of *T. kodakarensis* RNAP to reactions containing purified +125 nt transcripts does not stimulate FttA activity over 30 min. Similar results were obtained in two independent experiments.



Extended Data Fig. 4 | The active center of FttA, an intact Spt4/5 complex and the stalk domain of the archaeal RNAP are necessary for efficient and kinetically-competitive FttA-mediated termination *in vitro*. TEC_{s+125} were assembled using promoter directed, biotinylated DNA templates. Intact TECs are bound to pellet fractions (P) while released transcripts are recovered from the supernatant (S). **b**, $FttA^{H255A}$ retains only minimal cleavage and termination activity alone, and inefficiently terminates stalled or slowly elongating TECs (lanes 17–24). $FttA^{H255A}$-mediated termination becomes more efficient upon addition of Spt4-Spt5 but remains non-competitive with transcription elongation at high [NTP] (lanes 25–32). Note that the left part of this figure (lanes 1–16) is a duplication of the left part of Fig. 3c. **c**, Spt4/5 complexes stimulate FttA-mediated termination (Fig. 3), however, addition of Spt5 alone, containing (lanes 17–24) or lacking (lanes 25–32) the N-terminal NGN domain, fails to stimulate FttA-mediated termination to be competitive with transcription elongation at high [NTPs]. **d**, Spt4 alone, or together with the KOW domain of Spt5, is insufficient to stimulate FttA-mediated termination to be competitive with transcription elongation at high [NTPs]. **e**, While TECs assembled with $RNAP^{WT}$ support kinetically-competitive FttA-mediated termination (Fig. 3), TECs generated with $RNAP^{\Delta F/\Delta E}$ only support FttA-mediated termination of stalled or slowly elongating complexes. The absence of the stalk domain impairs both FttA-mediated cleavage and release of the nascent transcript, and while FttA-activity can be stimulated by the addition of Spt4/5, the hinderance to FttA-mediated termination in the absence of the stalk domain impairs FttA-mediated termination under condition of high [NTP]. Each experiment (**b**–**e**) was performed once independently.



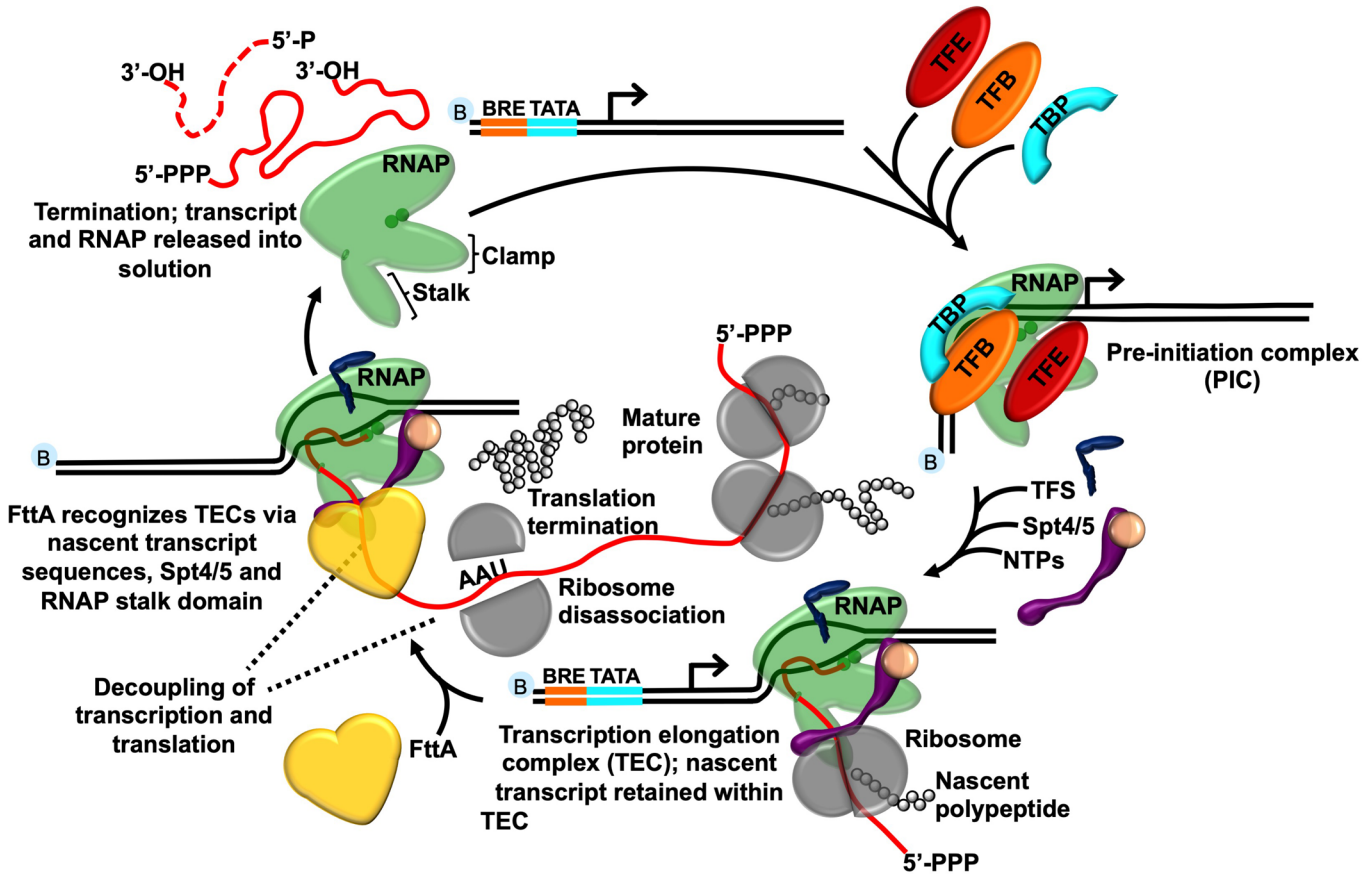
Extended Data Fig. 5 | FttA is an abundant protein likely responsible for 3′-end formation in archaeal cells. Quantitative Western blots employing anti-FttA antibodies, purified recombinant FttA, and total cellular lysates derived from known numbers of lysed *T. kodakarensis* cells reveal that FttA is present at ~2,100 -/+500 copies per cell. Cell counts and protein calculations were performed as described[20]. **A**, **B** and **C** represent independent biological samples.



Extended Data Fig. 6 | Gentle-purification of FttA directly from lysates of *T. kodakarensis* strain TK1428D. Top panels show SDS-PAGE gels of fractions recovered from imidazole elutions of total cell lysates from strains TK1428D (left) and TS559 (right) resolved over 5 ml Ni^{2+}-charged chelating columns (GE Healthcare). Bottom panels are Western blots of the same fractions from above probed with anti-HA antibodies to identify fractions within TK1428D lysates that contain FttA. The fractions pooled and analyzed by MuDPIT are identified. Magic Mark protein ladders are identified by molecular weight in Kda to the left of the gels. Data shown are from a single experiment.

Gene	Annotation	Mascot Score	Number of Unique Peptides
TK 1428	Cleavage and Polyadenylation specificity factor homologue	6388	47
TK 2215	tRNA splicing endonuclease	633	12
TK 0011	Uncharacterized protein	585	7
TK 1557	Predicted dehydrogenase	537	18
TK 1165	Predicted AP endonuclease	468	18
TK 2250	Serine/Threonine protein kinase	421	17
TK 0976	Putative snRNP Sm-like protein	351	8
TK 1509	Probable tRNA pseudouridine synthase	202	13
TK 0528	Serine hydroxyethyltransferase	168	13
TK 0211	Amidophosphoribosyltransferase	147	9
TK 1305	Probable translation initiation factor IF-2	133	11

Extended Data Fig. 7 | Proteins identified as co-eluting partners of FttA from lysates of strain TK1428D.



Extended Data Fig. 8 | FttA-mediated transcription termination completes the archaeal transcription cycle. Promoter-directed assembly of pre-initiation complexes requires RNAP, TFB and TBP and is often assisted by TFE. *De novo* RNA synthesis permits promoter escape and transcription initiation factors are replaced by transcription elongation factors TFS and Spt4-Spt5. The absence of a nuclear compartment permits translation initiation and the normal coupling of the archaeal transcription and translation apparatuses throughout transcription of the gene or operon, but this coupling is disrupted by translation termination. The exposed nascent transcript likely permits loading of FttA to TECs and FttA activity mediates cleavage of nascent transcripts and release of RNAP to solution.

nature research

Corresponding author(s): Santangelo, Thomas J.

Last updated by author(s): December 11, 2019

Reporting Summary

Nature Research wishes to improve the reproducibility of the work that we publish. This form provides structure for consistency and transparency in reporting. For further information on Nature Research policies, see Authors & Referees and the Editorial Policy Checklist.

Statistics

For all statistical analyses, confirm that the following items are present in the figure legend, table legend, main text, or Methods section.

n/a	Confirmed	
☐	☒	The exact sample size (*n*) for each experimental group/condition, given as a discrete number and unit of measurement
☐	☒	A statement on whether measurements were taken from distinct samples or whether the same sample was measured repeatedly
☐	☒	The statistical test(s) used AND whether they are one- or two-sided *Only common tests should be described solely by name; describe more complex techniques in the Methods section.*
☒	☐	A description of all covariates tested
☒	☐	A description of any assumptions or corrections, such as tests of normality and adjustment for multiple comparisons
☐	☒	A full description of the statistical parameters including central tendency (e.g. means) or other basic estimates (e.g. regression coefficient) AND variation (e.g. standard deviation) or associated estimates of uncertainty (e.g. confidence intervals)
☒	☐	For null hypothesis testing, the test statistic (e.g. *F*, *t*, *r*) with confidence intervals, effect sizes, degrees of freedom and *P* value noted *Give P values as exact values whenever suitable.*
☒	☐	For Bayesian analysis, information on the choice of priors and Markov chain Monte Carlo settings
☒	☐	For hierarchical and complex designs, identification of the appropriate level for tests and full reporting of outcomes
☒	☐	Estimates of effect sizes (e.g. Cohen's *d*, Pearson's *r*), indicating how they were calculated

Our web collection on statistics for biologists contains articles on many of the points above.

Software and code

Policy information about availability of computer code

Data collection	ImageQuant version 5.2 was used to analyze phoshorimages. The current study has not developed any novel software or code.
Data analysis	Isotopically-labeled RNA species were quantified using GE Imagequant software v5.2 using established techniques common to the field. Mass spectrometry analyses were completed at the Ohio State University mass spectrometry facility (https://www.ccic.osu.edu/MSP) using the MASCOT search engine from Matrix Science (http://www.matrixscience.com).

For manuscripts utilizing custom algorithms or software that are central to the research but not yet described in published literature, software must be made available to editors/reviewers. We strongly encourage code deposition in a community repository (e.g. GitHub). See the Nature Research guidelines for submitting code & software for further information.

Data

Policy information about availability of data

All manuscripts must include a data availability statement. This statement should provide the following information, where applicable:

- Accession codes, unique identifiers, or web links for publicly available datasets
- A list of figures that have associated raw data
- A description of any restrictions on data availability

All raw images used for data analysis is available upon request. Representative images are included in the current manuscript.

Field-specific reporting

Please select the one below that is the best fit for your research. If you are not sure, read the appropriate sections before making your selection.

☒ Life sciences ☐ Behavioural & social sciences ☐ Ecological, evolutionary & environmental sciences

For a reference copy of the document with all sections, see nature.com/documents/nr-reporting-summary-flat.pdf

Life sciences study design

All studies must disclose on these points even when the disclosure is negative.

Sample size	Transcription termination efficiencies were reported as the standard mean from at least three independent experiments. The exact number of independent replicates is reported for each figure. Experimentation was largely in vitro. Our in vivo work is reliant on strain constructions that we pioneered and all details are available in published literature.
Data exclusions	No data points were excluded from our analyses.
Replication	Figures and legends report means and standard deviations demonstrating reproducibility of our experimentation. Figures and legends report means and standard deviations demonstrating reproducibility of our experimentation.
Randomization	Randomization was not required for the current studies. In vitro biochemistry does not necessitate randomization.
Blinding	Blinding was not necessary for the current studies. In vitro biochemistry does not require blinding.

Reporting for specific materials, systems and methods

We require information from authors about some types of materials, experimental systems and methods used in many studies. Here, indicate whether each material, system or method listed is relevant to your study. If you are not sure if a list item applies to your research, read the appropriate section before selecting a response.

Materials & experimental systems

n/a	Involved in the study
☐	☒ Antibodies
☒	☐ Eukaryotic cell lines
☒	☐ Palaeontology
☐	☒ Animals and other organisms
☒	☐ Human research participants
☒	☐ Clinical data

Methods

n/a	Involved in the study
☒	☐ ChIP-seq
☒	☐ Flow cytometry
☒	☐ MRI-based neuroimaging

Antibodies

Antibodies used	Custom polyclonal anti-FttA antibodies were generated in two mice at Cocalico Biolabs using full-length recombinant FttA as an antigen. Total serum from each animal was provided and used in Western blots in the current work. Commercially available monoclonal anti-HA antibodies previously available from Covance Research, but now marketed by BioLegend (Covance #MMS-101R; BioLegend 901513; both stem from clone #16B12) were used in this study. Lot #GR3295537-1 was used at a 1:4,000 dilution. Custom antibodies against recombinant FttA were raised in commercial facilities by Coalico Biologicals (http://www.cocalicobiologicals.com; Stevens, PA). Pre- and post-immunization serum were tested on purified, mass-spec verified FttA as well as extracts from WT cells. The effective dilutions of each lot of FttA-antibodies are empirically determined in house, but typically range from 1:1,000 - 1:10,000.
Validation	Pre-immune serum and post-inoculated serum were used with purified FttA to confirm antibody specificity Purified FttA was used as a standard to confirm FttA specificity in total cell lysates. The efficacy of all lots and aliquots of our existing Covance stocks, and any new BioLegend provided materials have and will continue to be determined using Western blots with control extracts from a E. coli strain harboring a sequence-confirmed plasmid encoding a ~35 kDa HA-tagged MBP fusion protein. Anti-FttA antibodies were confirmed with Thermococcus kodakarensis lysates and Thermococcus kodakarensis derived and purified FttA.

Animals and other organisms

Policy information about studies involving animals; ARRIVE guidelines recommended for reporting animal research

Laboratory animals	Components were purified from Thermococcus kodakarensis. Strain TS559 was used for all genetic constructions.

Wild animals	No wild animals were used in this study.
Field-collected samples	No field-collected samples were used in this study.
Ethics oversight	No ethics oversight was necessary or applied to the current study.

Note that full information on the approval of the study protocol must also be provided in the manuscript.