Direct binding of TFEα opens DNA binding cleft of RNA polymerase

Sung-Hoon Jun1,2,6,7✉, Jaekyung Hyun1,4,6, Jeong Seok Cha2, Hoyoung Kim2, Michael S. Bartlett5, Hyun-Soo Cho2,7✉ & Katsuhiko S. Murakami3,7✉

Opening of the DNA binding cleft of cellular RNA polymerase (RNAP) is necessary for transcription initiation but the underlying molecular mechanism is not known. Here, we report on the cryo-electron microscopy structures of the RNAP, RNAP-TFEα binary, and RNAP-TFEα-promoter DNA ternary complexes from archaea, *Thermococcus kodakarensis* (Tko). The structures reveal that TFEα bridges the RNAP clamp and stalk domains to open the DNA binding cleft. Positioning of promoter DNA into the cleft closes it while maintaining the TFEα interactions with the RNAP mobile modules. The structures and photo-crosslinking results also suggest that the conserved aromatic residue in the extended winged-helix domain of TFEα interacts with promoter DNA to stabilize the transcription bubble. This study provides a structural basis for the functions of TFEα and elucidates the mechanism by which the DNA binding cleft is opened during transcription initiation in the stalk-containing RNAPs, including archaeal and eukaryotic RNAPs.
transcription is the first step of gene expression, performed by cellular RNA polymerases (RNAPs) in the three life domains: bacteria, archaea, and eukaryotes. The structures of RNAP are highly conserved in archaea and eukaryotes, and consist of a "crab-claw" shape core with a protruding stalk domain.

For promoter-specific transcription, RNAP and general transcription factors (GTFs) are recruited to the promoter DNA to form the pre-initiation complex (PIC). The melting of DNA around the transcription start site (TSS) converts PIC to the open complex (OC), which is competent to initiate RNA synthesis. Archaeal RNAP requires three general transcription factors (GTFs) — TBP, TFB, and TFEα — whereas Pol II requires more complex GTFs including TFIIA, TFIIB, TFIIH/TBP, TFIIE, TFIIF, TFIH, and the Mediator complex. The orthologues of archaeal TBP, TFB, and TFEα are TBP, TFIIB, and TFIIEα-subunit, respectively, which indicates that the archaeal transcription system requires a subset of Pol II GTFs. Due to its simplicity, the archaeal transcription machinery is considered to be the evolutionary precursor and simplified form of the Pol II transcription apparatus that retains the conserved functions between two systems.

The DNA binding cleft of cellular RNAPs can adopt open or closed state by the conformational changes of the clamp domain of RNAP. Opening the DNA binding cleft is required for promoter DNA loading into the cleft. For Pol II, determination of PIC and OC structures was achieved in yeast and human. Yeast PIC structures show the closed cleft, whereas it shows the open cleft in human. Recent cryo-electron microscopy (cryo-EM) structure of yeast PIC reveals the closed clamp with a distorted promoter DNA in the closed cleft, which suggests that the transition of the clamp from the open conformation to the closed conformation induces DNA distortions to prime DNA melting. In the OC structures of both human and yeast, the cleft is closed with the closed clamp conformation. In contrast, the biophysical study of archaeal transcription complexes using a single-molecule fluorescence energy transfer (smFRET) assay suggested a closed cleft PIC and open cleft OC structure. Together, the status of the cleft during the transcription initiation process for archaeal RNAP suggests that the transition of the clamp from the open conformation to the closed conformation is induced by DNA distortions to prime DNA melting.

In general terms, archaeal transcription is initiated by the TFEα complex and the open cleft in human is converted to the closed cleft during the transcription initiation process. In contrast, the biophysical study of archaeal transcription complexes using a single-molecule fluorescence energy transfer (smFRET) assay suggested a closed cleft PIC and open cleft OC structure. Together, the status of the cleft during the transcription initiation process for archaeal RNAP suggests that the transition of the clamp from the open conformation to the closed conformation is induced by DNA distortions to prime DNA melting.

TFEα is a conserved GTF in archaeal-eukaryal transcription systems and is an essential factor in vivo. It consists of an N-terminal extended winged-helix domain (eWHD) and a C-terminal zinc ribbon domain (ZRD), and the X-ray crystal structure of TFEα in yeast was determined from Sulfolobus solfataricus (Sso). TFEα enhances transcription efficiency of RNAP in vitro, particularly for weak promoters under sub-optimal conditions. It facilitates the formation of PIC and promoter DNA melting to form OC, and also stabilizes OC. In relation to the clamp conformation, the smFRET study suggested that TFEα allosterically aids in opening the clamp during OC formation. However, the molecular mechanisms of the functions of TFEα remain elusive due to lack of information about its structural conformation in the transcription complexes.

TFIIEα is a heterodimeric complex of TFIIEα/β in human and Tfs1/2 in yeast. The N-terminal region of TFIIEα or Tfs1 is essential and corresponds to archaeal TFEα, while TFIIEβ or Tfs2 is conserved only in eukaryotes and is functionally redundant. The interactions between Pol II and TFIIE in the PIC have been studied from the cryo-EM structures in human and yeast, but the dynamic and transient nature of the PIC limits the resolution of the structures. Yeast PIC structures reveal the closed clamp, and there is no direct interaction between TFIIEα and the clamp, whereas the human PIC structure is in the open clamp, enabling TFIIEα interaction with the clamp. This human PIC structure suggests that TFIIEβ interacts with the clamp. Additionally, biochemical probes and crosslinking suggest that Tfs1 interacts with the clamp in PIC. Thus, the molecular details of interactions between TFIIE and Pol II in the PIC remain unresolved and the molecular mechanism of how TFIIE is involved in the clamp conformation change is unknown.

In this study, we determined the X-ray structure of TFEα and the cryo-EM structures of three forms of archaeal RNAP, including the apo-form, RNAP–TFEα binary complex, and RNAP–TFEα–DNA ternary complex. The structures reveal that direct binding of TFEα can shift the clamp conformation from a closed conformation to an open conformation. Our results show dynamic structural interactions among TFEα, RNAP, and promoter DNA, which provides structural information for the TFEα functions and transcription mechanism.

Results
X-ray crystal structure of TFEα. To determine the structure of TFEα, we first tried, unsuccessfully, to crystallize full-length TFEα from euarchaeal Thermococcus kodakarensis (Tk) and Pyrococcus furiosus (Pfu). After screening derivatives of TFEα, we found that the 13 kDa fragment Pfu TFEα (TFEαAZRD, 1–110 amino acid residues) was crystallized. The structure was solved by a single wavelength anomalous dispersion (SAD) at 3.2 Å and the resolution of the structure was improved to 2.6 Å with a native crystal (Supplementary Table 1). The TFEαAZRD structure consists of the N-terminal eWHD (residues 5–86) and a long α-helix (a5) that connects the eWHD and the ZRD. The eWHD structure is similar with the X-ray crystal structures of the eWHD of Sso TFEα and the eWHD of human TFIIEα.

We find a previously unrecognized feature of the TFEα structure from the TFEαAZRD structure. There were eight molecules in the asymmetric unit of the SAD structure and two molecules in the asymmetric unit of the native structure. Superposition of the globular part of the eWHD from the 10 crystallographically independent structures (average root-mean-square deviation in α-carbon positions is 0.369–0.589 Å over residues 8–62) reveal flexibilities in the wing and the a5 linker (Fig. 1c). The conformation of the wing can change about 24° (~13 Å), and the a5 can be bent at ~21° (~12 Å) for the eWHD (Fig. 1d). These results show a potential structural adaptability of TFEα during the transcription initiation process.

Cryo-EM structures of apo-RNAP and the binary complex of RNAP–TFEα. To investigate molecular details of the interactions among RNAP, TFEα, and promoter DNA during transcription initiation, we determined the cryo-EM structures of Tk RNAP in three distinct states in the transcription cycle (Fig. 2a, b). The three forms were the apo-form, the TFEα bound form (RNAP–TFEα binary complex), and the TFEα and DNA bound form (RNAP–TFEα–DNA ternary complex) at nominal resolutions of 3.9 Å, 4.0 Å, and 3.8 Å, respectively (Supplementary Table 2 and Supplementary Figs. 1–3).

We previously solved the X-ray crystal structure of Tk RNAP. To resolve accurate structural changes of Tk RNAP in solution through interactions with TFEα and promoter DNA, we determined the cryo-EM structure of Tk apo-RNAP. Three-dimensional classification of the apo-RNAP dataset showed that 82% particles formed complete apo-RNAP structures whereas 18% particles formed a stalkless RNAP structure (Supplementary Fig. 1c). The class of apo-RNAP comprising 36% particles formed the complete apo-RNAP structure and the C-terminus is resolved (Supplementary Table 2 and Supplementary Figs. 1–3).
We discovered that the cryo-EM structure of apo-RNAP adopts a closed clamp conformation (Supplementary Fig. 4). Another 3D class covering 22%, referred as a minor class, was also refined and the resulting structure was similar with the major class structure with 3.3° shift in clamp conformation to the open direction compared to the major class structure, indicating that there is a slight mobility in the clamp of apo-RNAP. Consistent with this,
the results of the smFRET experiment suggested that free apo-RNAP in solution is predominantly in a closed clamp conformational state. To visualize the effect of TFEα binding on RNAP and molecular details of the RNAP and TFEα interaction, we determined the cryo-EM structure of the RNAP–TFEα binary complex. The amino acid sequences of TFEα from Pfu, and Tko are 71% identical (Fig. 3a), which allowed us to use the X-ray structure of TFEαΔZRD determined in this study (Fig. 1b) as a guide to build the complete model of TFEα with the eWHD, ZRD, and two α helix bundles linking these domains (Figs. 2b and 3b). The overall shape of Tko TFEα structure resembles the

![Diagram](image-url)
TFIIEα within the human PIC structure. The distinction between the two is that TFEα in the RNAP–TFEα binary complex interacts directly with the tip of the clamp whereas there is no direct interaction between TFIIEα and the tip of clamp in the human PIC (Supplementary Fig. 5). For coordinating Zn atoms in the ZRD, Tko TFEα uses three Cys residues (C117, C137, and C140) together with a His residue (H122) instead of using four Cys residues in case of the eukaryotic TFIIEα (Supplementary Fig. 5a). Archaeal TFEα contains a positively-charged tail (16 residues) at its C-terminus but its density is not traceable (Fig. 3a, b).

To reveal the precise effect of TFEα binding on RNAP, we compared the cryo-EM structures of the apo-RNAP and RNAP–TFEα binary complex. In the latter, the clamp swings away from the DNA binding cleft at 11.8° relative to the apo-form RNAP, and this movement is coupled with the stalk swinging (4.7°) away from the clamp (Fig. 3c and Supplementary Movie 1). The structure shows that TFEα connects two mobile modules of RNAP, the clamp and stalk. It demonstrates that direct binding of TFEα on RNAP shifts the clamp and the stalk conformation from a closed state to an open state.

The RNAP–TFEα binary complex structure shows the molecular details of the interplays between RNAP and TFEα to open the DNA binding cleft of RNAP. Two α helices between the eWHD and ZRD form an α-helix bundle linker, placing the ZRD over 50 Å away from the eWHD and allowing the eWHD and ZRD of TFEα to locate adjacent to the clamp and stalk domains of RNAP, respectively (Fig. 3d). The hydrophobic patch of the eWHD (residues Y55, F65, W75, and Y78) faces the tip of the coiled-coil (CC) of the clamp domain and the wing fits into the cavity between the CC and clamp head domain. Several hydrophilic residues of TFEα (R51, R52, D59, K66, R67, and R81) are also involved in the interaction with the RNAP clamp domain. The ZRD establishes its interaction with RpoE in the stalk domain of RNAP via hydrophobic surfaces (residues Y114, Y115, F127, M131, F135, and Y147). The D128 amino acid of TFEα also forms a salt bridge with K152 of RpoE. Consistent with the structure, deletion of the coiled-coil tip, amino acid substitutions at hydrophobic patch of RpoE (Y95E/F98E/L107E), or deletion of a short loop of RpoE (AS150-R157) disrupts TFEα binding to RNAP. The contact surface areas between the eWHD and clamp and between the ZRD and stalk are 632 and 441 Å², respectively. Amino acid residues involved in the TFEα–RNAP interaction are conserved in archaeal TFEα and only a few are conserved in eukaryotic TFIIEα (Fig. 3a). This explains the much lower interaction affinity in eukaryotic Pol II counterparts.

The structure of RNAP–TFEα–DNA ternary complex: DNA loading to the RNAP cleft closes the clamp. To examine the conformation of archaeal RNAP with TFEα in the DNA bound form, and how TFEα stabilizes the RNAP and DNA complex, we determined the cryo-EM structure of RNAP–TFEα–DNA ternary complex (Supplementary Table 2). We used a synthetic scaffold DNA in the form of a downstream fork junction promoter DNA containing the downstream double-stranded DNA and the partial transcription bubble (Fig. 2a), which functionally mimics open promoter DNA. Strong densities are observed in the ternary complex for double-stranded DNA within the DNA binding cleft and 5 bases of single-stranded template DNA (from −3 to +2) positioned near the active site of RNAP. However, density for the non-template DNA in the bubble is not traceable (Fig. 4a). The DNA position within the RNAP DNA binding cleft is the same as that of the cryo-EM structures of the yeast and human OC.

To identify the structural changes of RNAP and TFEα upon the binding of the transcription bubble-mimicking promoter DNA, we compared the cryo-EM structures of the RNAP–TFEα binary complex and RNAP–TFEα–DNA ternary complex and found that DNA binding to the DNA binding cleft of RNAP changes the conformations of the clamp and stalk back to the closed state (Fig. 4b). Although the clamp and stalk swing 17° (14 Å) and 7.3° (3.7 Å), respectively, TFEα still maintains the same interactions with these domains, as observed in the RNAP–TFEα binary complex. Relative orientations of the eWHD and ZRD in TFEα change in the binary and ternary complexes (Fig. 4c). The structural flexibility of TFEα, including the α5 linker (Fig. 1c, d), allows it to bind the domains and function dynamically in both open and closed states.

Closing of the clamp establishes a new interaction between the TFEα wing and the lobe domain of RpoB, including the interaction of R69 side chain (TFEα) with Q285 side chain and G283 main chain at the lobe loop (Fig. 4d). This interaction encloses the DNA binding main cleft and provides a structural basis for TFEα stabilization of the OC, which is reminiscent of the mechanism for Spt/5 stabilizing the elongation complex (Supplementary Fig. 6).

Models of the PIC and OC. We modeled the archaeal PIC and OC (Fig. 5 and Supplementary Movie 2) using the cryo-EM structures of the RNAP–TFEα and RNAP–TFEα–DNA complexes determined in this study. For modeling DNA, TBP, and TFB, we used the cryo-EM structures of the yeast PIC containing a straight B-form DNA (PDB: 6GYK) and human OC (PDB: 5IYB) containing both the template and non-template DNA strands in the transcription bubble as templates. Cryo-EM structures of the Tko RNAP–TFEα (open clamp form) and RNAP–TFEα–DNA complexes (closed clamp form) were aligned with Pol II in the PIC and OC, respectively, using the catalytic core of RNAP containing the double-phi-β-barrel (DPBB) domains of the largest (RpoA1/Rpb1) and the second largest (RpoB/Rpb2) subunits.
In the PIC model, the open clamp state of RNAP can accommodate double-stranded DNA between the TFEα wing and RpoB lobe loop (Fig. 5a), whereas clamp closure results in a steric clash between the TFEα wing and DNA. The lobe loop positions right below a major groove of DNA around −3T/−2T position, whereas the TFEα wing does not establish any contact with DNA. Thus, the opening state facilitates the subsequent positioning of the double-stranded upstream DNA into the DNA binding cleft, which explains the transcription enhancing activity of TFEα.

In the OC model, the RNAP clamp can be closed without steric clash with single-stranded DNA strands in the transcription bubble (Fig. 5b). The TFEα wing is located adjacent to the non-template strand and the upstream edge of the transcription bubble, which is similar to the eukaryotic OC structures. We observe that β2 of the wing is on the side of the non-template DNA and the aromatic side chain of W75 is positioned to establish the stacking interaction with the base of the non-template DNA at −9 position where the DNA bubble starts. This is reminiscent of the interaction between the bacterial RNAP σ factor region 2.3 and the non-template DNA in the open promoter complex, RPo. This suggests a molecular mechanism of how TFEα stabilizes the transcription bubble during the transcription initiation by the interaction with the upstream DNA of OC.

Crosslinking experiments in the archaenal PIC and OC. To investigate how TFEα interacts with promoter DNA and the non-template DNA strand observed in the OC model, we used a highly specific UV-inducible crosslinking system based on the site-specific incorporation of unnatural amino acid para-benzoyl-phenylalanine (Bpa) (Fig. 6). We prepared a series of Bpa substituted TFEα variants, including Bpa substitutions at I17 and N49 in the globular part of eWHD and Y56, F66, R70, W76, Y78, and Y79 in the wing (Fig. 6a). The Bpa reacts preferentially with unreactive C–H bonds when exposed to long wave UV light, with a reactive spherical radius of 3.1 Å, and these covalent cross-links can provide very specific proximity information for interacting macromolecular surfaces.

First, to establish whether the TFEα Bpa substitutions affected transcription activation by TFEα, we examined each TFEα variant for gdh promoter-dependent transcription in transcription complexes containing TBP and TFB2 (Fig. 6c and Supplementary Fig. 7). TFB2 was used since TFEα activation is greater with TFB2 than with TFB1. A modest activation was observed for wild type TFEα and all Bpa variants except for Y79Bpa, which lacked any activation, possibly because of interference by the bulkier Bpa R-group. This indicated that the transcription activation by TFEα was not affected by Bpa substitutions in most variants.

Next, to monitor the site-specific contact between TFEα and DNA, we prepared three sets of gdh promoter DNA containing radiolabel at −9/−8/−7 positions of non-template DNA (−9NT), at −3 position of non-template DNA (−3NT), and at −4 position of template DNA (−4T) (Fig. 6b). The −9NT and −3NT/−4T promoter positions are located at the upstream edge and at the middle of the DNA bubble in the open complex of the eukaryotic Pol II transcription system. Previous DNA to protein cross-links

---

**Fig. 4 DNA interaction with RNAP in the presence of TFEα.** (a) Close-up view of the DNA binding channel of the RNAP-TFEα-DNA complex is shown with the cryo-EM density map (yellow and transparent) of DNA. DNA is shown as a stick model (template DNA, dark green; non-template DNA, light green). Template DNA strand (tDNA), downstream double-stranded DNA (dDNA) and transcription start site (+1) are indicated. Mg$^{2+}$ bound at the active site of RNAP and bridge helix (BH) are depicted. (b) The structures of the RNAP-TFEα binary (red) and the RNAP-TFEα-DNA ternary (green) complexes are superimposed. The conformational changes of the clamp and stalk by DNA positioning in the ternary complex are indicated. (c) Conformational change of TFEα upon RNAP binding DNA. The ZRD of TFEα was used as a reference to superimpose the RNAP and TFEα in the binary (color) and ternary complexes (light gray), revealing movement of the eWHD of TFEα together with the RNAP clamp indicated by an arrow. (d) Close-up view of the RNAP lobe and TFEα wing interaction. Amino acid residues involved in the RNAP lobe and TFEα wing interaction are depicted as stick models and indicated. Surface exposed aromatic residues (W75 and Y77) of TFEα involved in the DNA interaction are also indicated.
indicated that TFEα in preinitiation complexes cross-links strongly to −11 and −9 of the NT strand, but not to the T strand. We observed crosslinking signals for the DNA radiolabelled at −9NT position with TFEα variants with Bpa at R70, W76, Y78, and Y79 (Fig. 6d and Supplementary Fig. 8). The strongest signal was observed in W76Bpa, located at the tip of wing, indicating that W76 is positioned proximally to the upstream edge of transcription bubble. This result is consistent with the observation in the OC model (Fig. 5b) that the corresponding residue W75 is located to interact with the non-template DNA at the upstream edge of transcription bubble. In the OC model, when the clamp is open, the distance between −9NT base and W75 of TFEα is 21 Å, which is too far for crosslinking, whereas the distance becomes 10 Å when the clamp is closed. Thus, the crosslinking result also indicates that the RNAP clamp is closed in the OC.

**Discussion**

Here, we report the cryo-EM structures of apo-RNAP, RNAP–TFEα binary complex, and RNAP–TFEα-promoter DNA ternary complex that visualize dynamic structural changes of TFEα and archaeal RNAP, especially in the clamp and stalk, during transcription initiation (Supplementary Movies 1). Stable binding of TFEα to Tko RNAP without other GTFs and promoter DNA allowed us to demonstrate that the direct binding of TFEα opens the DNA binding cleft of RNAP. We suggest a simplified transcription initiation model (Supplementary Movie 2) in archaea based on the model structures of PIC and OC in this study (Fig. 5). In this model, the clamp of apo-RNAP is closed in a conformation with a slight mobility. In the PIC, the binding of TFEα on the clamp and stalk shifts the conformations of the mobile domains of RNAP to an open state (Fig. 5a) and then, the double-stranded downstream promoter DNA is placed into the DNA binding cleft. The positioning of the downstream DNA into the DNA binding cleft induces the clamp closing by strong ionic interactions between positively-charged residues covering the inside of the DNA binding cleft and a negative-charged DNA backbone. During the clamp closure, DNA distortion may be induced as observed from the yeast Pol II system, and we suggest that the wing of TFEα may be inserted into the DNA minor groove around −6 position where the transcription bubble is formed (Supplementary Movie 2). This also explains how TFEα facilitates the unwinding of the DNA. In the OC, the clamp is in a closed state and the wing of TFEα is located to stabilize OC structure by its interactions with the opposite region of the cleft (Supplementary Fig. 6) and the non-template strand DNA at the upstream edge of the transcription bubble (Fig. 5b).

This model shows how the DNA binding cleft of stalk-containing RNAPs can be opened in PIC and provides the structural basis of TFEα functions to facilitate transcription initiation. A previous study using smFRET assay suggested TFEα helps clamp opening allosterically during OC formation. However, our results show that the direct binding of TFEα opens the clamp conformation. In the cryo-EM structures of eukaryotic Pol II PIC, the interaction between Pol II and TFIIE was not resolved in relation with the conformation of the clamp and stalk because of the dynamic and transient nature of PIC structure as well as the complex composition (Supplementary Fig. 9). Given that the essential N-terminal part of TFIIεα in humans (Tf1l in yeast) corresponds to TFEα and their structures are
conserved (Supplementary Fig. 5), our results suggest that the clamp opening activity of TFEα in the PIC by the direct interaction with the tip of the clamp is conserved in TFIIEα (Tfa1 in yeast) as well. This is consistent with other studies8,33 where the intermediate PIC structure having an open clamp conformation has been observed in humans, and has been reported as a requirement for the priming of DNA melting10. Additionally, our results suggest that the clamp opening activity of TFEα is located close to the clamp in the PIC22.

Our results also suggest how the binding of TFEα to RNAP can shift the conformation of the clamp and stalk from a closed state to an open state. The cryo-EM reconstruction of the RNAP–TFEα binary complex in the high threshold of the map shows that the density between the ZRD and the stalk is rigid, whereas the density of the eWHD and clamp is weak (Supplementary Fig. 10), which suggests a modest structural heterogeneity in the eWHD and clamp. The asymmetric interaction of TFEα on the stalk and clamp shows the mode of TFEα binding on RNAP and suggests a model of how TFEα binding on the clamp and stalk domains shifts their conformations from a closed state to an open state. In this model, the interaction between the ZRD and the stalk tethers TFEα on the RNAP. Subsequently, the eWHD of TFEα interacts with the coiled-coil tip of the clamp, which physically connects the center region of the stalk and the tip of the clamp. The bridging of the two domains of RNAP by TFEα mechanically allows the conformational change of the clamp from a closed state to an open state via a leverage effect. There is a slight structural heterogeneity in the clamp and eWHD due to the flexibility of the domains. The conformational change in the clamp produces a coupled conformation change in the stalk, as observed previously in other RNAP structures2,34.

In eukaryotic Pol II, a clash of synthesized RNA with TFIIB causes promoter escape35 and the subsequent initiation–elongation transition alters the path of upstream DNA without significant conformational change in Pol II7,35,36.
The conformation of RNAP in the RNAP–TfEx–DNA ternary complex is in the closed clamp state (Figs. 4b) as archaeal RNAP in the RNAP–Spt4/5 complex structure26 as well as Pol II in the OC structure27–29 and the transcription elongation complex (EC) structure30. TfEx can bind to RNAP stably in the RNAP–TfEx–DNA ternary complex structure (Figs. 2 and 4), which supports the notion that TfEx is a part of early EC19 and Spt4/5 displaces TfEx in the EC to enhance transcription processivity of the EC16.

smFRET study in solution characterized the conformational positioning of the archaeal RNA polymerase clamp during the transcription cycle17. The structural information for the clamp conformation in this report is consistent with the smFRET work (Figs. 3c and 5a) but there is a difference after DNA melting and the OC formation (Figs. 4b and 5b). Recent reports showed that the OC formation is a multi-step process and clamp conformation changes are crucial event during DNA melting and the OC formation32,38. The structures of complete transcription complexes including GTFs such as TBP and TFB along with further biophysical studies would give refined models of transcription initiation mechanism in archaea.

The simpler bacterial system suggested detailed models for promoter DNA melting and subsequent arrangement of the transcription bubble37–39. Biophysical study assigned bacterial RNAP clamp conformation at each steps in transcription initiation and elongation39 and real-time fluorescent study monitored that a transient clamp closure is required to nucleate DNA melting37. Universal structural features of RNAP within the DNA binding cleft play key roles in the late steps of promoter DNA melting38 and clamp opening is required for the template DNA strand to enter into the active site of RNAP37,38. Our results support the universally conserved mechanisms of transcription initiation in all three domains of life by showing that direct binding of TfEx, a conserved GTF in archaeal-eukaryal RNAP systems, can open the clamp of the stalk-containing RNAPs. In Thermus and Mycobacterium, interestingly, an additional GTF CarD stabilizes the OC at the upstream edge of the transcription bubble40,41 as TfEx does in archaea, which also suggests some domain- or lineage-specific transcription factors incorporated during evolution to carry out common functions.

Methods

TfEx cloning, expression, and purification. The DNA fragment encoding the N-terminal 110 amino acids of the Pfu TfEx (TfExAZRD) was cloned into a pSMUO vector (LifeSensors) with Bsal/Xhol restriction sites (Supplementary Table 3). The plasmid was transformed into BL21(DE3) CodonPlus cells (Stratagene), and the vector (LifeSensors) with BsmBI/BamHI restriction sites (Supplementary Table 3) and TdP18000 (Stratagene) was transformed into BL21(DE3) CodonPlus cells (Stratagene), and the vector (LifeSensors) with BsaI/XhoI restriction sites (Supplementary Table 3). The mutant model was built in the crystallographically free flap region of RNAP-EC31. Manual model building was done using Coor45 nm, and the structure was refined using PHENIX. The structure of native Pfu TfExAZRD was determined by the molecular replacement using Automated molecular replacement (Phaser-MR) in PHENIX followed by the structure refinement. Final coordinates and structure factor data were listed in the Supplementary Movie 1.

Preparations of RNAP, RNAP–TfEx and RNAP–TfEx–DNA complexes. Tko RNAP was purified from Δpol4 strain KUWLVB16 and then the 11-subunit apo-Tko RNAP was reconstituted in vitro by supplementing recombinant Rpo4 and Rpo72. About 50 g of cells were suspended in 200 ml lyso buffer (10 mM Tris-Cl (pH 8.0), 500 mM KCl, 10% glycerol, 10 mM midaclate, 10 μM ZnCl2, 5 mM 2-mercaptoethanol, 0.3 μM leupeptin, 0.3 μM benzamidine, 0.1 mM ethylenediaminetetraacetic acid (EDTA), and 0.5 mM phenylmethylsulfonyl fluoride) and lysed by sonication. The supernatant was loaded to Ni-NTA affinity column (Qiagen) equilibrated with the lysis buffer and washed with the same buffer containing 20 mM imidazole. Proteins were eluted with the lysis buffer containing 200 mM imidazole and precipitated with ammonium sulfate (final 80% saturation). The pellet was suspended in TGED buffer (20 mM Tris-Cl (pH 8.0), 10% glycerol, 0.5 μM EDTA and 5 mM DTT) and RNAP was further purified by binding and elution from HiTrap Q HP (GE Healthcare). To reconstitute RNAP containing all subunits, RNAP pools from HiTrap Q were mixed with recombinant Rpo4/Rpo7 and Rpo4 at a ratio of 1:4 for Rpo4/Rpo7/Rpo4 and were further purified by size exclusion and passage and elution from HiTrap Heparin and Superdex200 columns (GE Healthcare). The gene encoding Tko TfEx was cloned into a pSMUO vector (LifeSensors) with BsmB/BamH restriction sites (Supplementary Table 3) and TfEx was expressed in BL21(DE3) CodonPlus cells (Stratagene) using a pSMUO vector (LifeSensors) in 2×YT media at 37 °C. The temperature was shifted to 18 °C when OD600 of culture reached 0.6 and protein expression was induced by 0.2 mM IPTG for 2 h. After harvesting cells, TfEx was purified as TfExAZRD.

The binary complex of RNAP and TfEx was formed by mixing 3 n mole RNAP with 5-fold molar excess of TfEx at room temperature for 1 h. The sample was applied to a Superose 6 10/300 GL column (GE Healthcare) eluted with 10 mM Tris-Cl (pH 8.0), 500 mM NaCl, 5 mM DTT, 0.1% Triton X-100 and collected in 300 μl fractions. The fraction corresponding to the peak of the complex, in about 0.3 mg/ml, was used directly in the subsequent EM analysis.

For the preparation of the RNAP–TfEx–promoter DNA ternary complex, DNA sample was soaked in TGED buffer (20 mM Tris-Cl (pH 8.0), 10% glycerol, 0.5 mM EDTA and 5 mM DTT) and RNAP was further purified by binding and elution from HiTrap Q HP (GE Healthcare). To reconstitute RNAP containing all subunits, RNAP pools from HiTrap Q were mixed with recombinant Rpo4/Rpo7 and Rpo4 at a ratio of 1:3:1 for Rpo4/Rpo7/Rpo4 and were further purified by size exclusion and passage and elution from HiTrap Heparin and Superdex200 columns (GE Healthcare). The gene encoding Tko TfEx was cloned into a pSMUO vector (LifeSensors) with BsmB/BamH restriction sites (Supplementary Table 3) and TfEx was expressed in BL21(DE3) CodonPlus cells (Stratagene) using a pSMUO vector (LifeSensors) in 2×YT media at 37 °C. The temperature was shifted to 18 °C when OD600 of culture reached 0.6 and protein expression was induced by 0.2 mM IPTG for 2 h. After harvesting cells, TfEx was purified as TfExAZRD.

The binary complex of RNAP and TfEx was formed by mixing 3 n mole RNAP with 5-fold molar excess of TfEx at room temperature for 1 h. The sample was applied to a Superose 6 10/300 GL column (GE Healthcare) eluted with 10 mM Tris-Cl (pH 8.0), 500 mM NaCl, 5 mM DTT, 1% Triton, 0.1 mM EDTA and collected in 300 μl fractions. The fraction corresponding to the peak of the complex, in about 0.3 mg/ml, was used directly in the subsequent EM analysis.

Cryo-electron microscopy data acquisition. Samples of 4 μl of purified apo-RNAP, RNAP–TfEx, and RNAP–TfEx–DNA of ~0.3–0.5 mg/ml were loaded onto freshly glow-discharged holey carbon grids (Quantifoil R1.2/1.3 Cu200, Quantifoil C Tecnor) wet in a drop of water for 5 s and blotted dry with a filter paper. The grids were fixed by a 100 kV electron microscope (Techno Electron, Germany) or a FEI Tecnai 300 kV electron microscope equipped with a Gatan Imaging Filter (GIF), multi-tilt stage, and a direct electron detector. Imaging was performed in nanoscale probe mode, spot size 3, 70 μm C2 aperture and 100 μm objective lens aperture. Cs corrector was tuned just within 2 days by hanging drop vapor diffusion method at 21 °C with the crystallization solution containing 10% PEG 8000, 0.2 M MgCl2, and 0.1 M Tris–HCl (pH 7.7). Crystals were cryoprotected by adding ethylene glycol to a final concentration of 20 % prior to plunge-freezing in liquid nitrogen. Diffraction data were collected at the A1 beamline of the Cornell University High Energy Synchrotron Source (CHESS, Ithaca, NY) at 100 K and data were processed by HKL200041. Native crystals belonged to P212121, and contained two molecules in an asymmetric unit, whereas SeMet labeled crystals belonged to P212121, and contained eight molecules in an asymmetric unit.

X-ray crystal structure determination and refinement. With the anomalou signal from SeMet, 32 selenium sites in the asymmetric unit (4 sites in each of the 8 molecules in an asymmetric unit) were located and the experimental phase (figure of merit: 0.390) was calculated using Automated structure solution (AutoSol) in PHENIX46. Density modification by Automated model building (AutoBuild) in PHENIX yielded an excellent map and most of the model was built automatically. Model manual building was done using Coor45 nm, and the structure was refined using PHENIX. The structure of native Pfu TfExAZRD was determined by the molecular replacement using Automated molecular replacement (Phaser-MR) in PHENIX followed by the structure refinement. Final coordinates and structure factor data were listed in the Supplementary Movie 1.
before data acquisition at normal magnification of ×125,000 in order to minimize 2-fold and 3-fold astigmatism, axial coma, and spherical aberration. Automated data collection was performed using EPU software (Thermo Fisher Scientific Inc., USA) at a nominal magnification of ×470,000, which corresponds to a pixel size of 1.4 Å at the specimen scale, and to a defocus range between −1.5 μm and −3.5 μm. Thirty frames per movie data were collected at a dose rate of ~35 e−/Å²/s over a 1.8 s exposure time. A total of 1386 movies were collected for apo-RNAP and 2698 movies for RNAP–TFEα, respectively. For RNAP–TFEα–DNA complex, 1,737 movies and 2,106 movies were collected from the holey grid with and without additional graphene oxide support film, respectively.

Cryo-electron microscopy image processing. MotionCor246 was used for beam-induced movement correction and dose-weighting from the movie data. The motion-corrected micrographs were subjected to contrast transfer function estimation using GcCt47. Relion 2.148 was employed for subsequent image processing procedures, unless otherwise stated. Typically, 10,000–20,000 particles were manually picked from micrographs with high underfocus to aid visibility of the particles. The particles were extracted into 176 × 176 pixels boxes followed by 2D class averaging. Class averages with pronounced structural features were used as a template for automated particle picking. The box size and particle-picking regime were commonly applied to all datasets.

For apo-RNAP, 514,007 particles were extracted from particle auto-picking. Two rounds of 2D class averaging were performed from particles that belonged to class averages with poor structural features were eliminated after each run, leaving 350,867 particles in the dataset (Supplementary Fig. 1b). The X-ray structure of Tko RNAP (PDB: 4QW) was converted to an EM map and low pass filtered to 60 Å, using the x2pdf2mnc-py routine in the EMAN2 software package49. The low pass reference for the subsequent class averages was used for 2D classification, but due to the strong preferential orientation of particles, the initial attempt of 3D reconstruction resulted in a poorly resolved map. To alleviate the effect of the preferred orientation, particle images that belonged to strongly preferred orientations were discarded based on the 2D class averages. The resulting 247,663 particles were used to reconstruct a consensus map without any preceding 3D classification. Then, the particles were classified into four 3D classes without performing any image alignment. The 139,242 particles that belonged in the 3D class that exhibited the best structural detail were selected for subsequent per-particle motion correction and B-factor weighing via particle polishing (Supplementary Fig. 2a). Finally, the corresponding particle images were refined and post-processed to yield a reconstruction at 3.91 Å resolution based on a 0.143 Fourier shell correlation (FSC) cut-off criterion (Supplementary Fig. 1e)50. For RNAP–TFEα, a total of 1,226,339 particles were auto-picked. Then 953,770 particles belonging to the best 2D class averages were selected after two rounds of class averaging (Supplementary Fig. 2b). Using low pass filtered X-ray structure of Tko RNAP as an initial model, the dataset was subjected to 3D classification. Out of six resultant 3D classes, two classes exhibited the best structural details and unambiguously discernible densities for TFEα and stalk domain, and 409,653 particles belonging to those classes were selected for subsequent processing. The particles were re-refined in 3D without adjusting the orientation of the particles to achieve the best possible root mean square deviation (RMSD) value for the density corresponding to stalk and clamp domains of RNAP and TFEα was created from a preliminary atomic model. Focused refinement on this region was performed by continuing the last iteration of the 3D refinement run of the entire structure with the applied local mask. The overall resolution of the resulting map from focused refinement was 4.25 Å with a slight improvement in local resolution of the masked regions (Supplementary Fig. 2e, f).

For RNAP–TFEα–DNA, a total of 1,388,666 particles were auto-picked from the combined micrograph datasets obtained from the vitrified grids with and without graphene oxide film. After two rounds of 2D class averaging, 911,020 particles were auto-picked (Supplementary Fig. 3a) and subjected to 3D classification into four classes. A total of 669,840 particles belonging to the 3D classes with pronounced structural details and intact domains, TFEα and DNA, were used to produce a consensus 3D reconstruction. The dataset was further classified into eight 3D classes without image alignment. After careful inspection of the 3D classes, two classes with intact TFEα, stalk domain, DNA, and good structural details were chosen, from which 312,092 particles were extracted (Supplementary Fig. 3b). Among the particles included for the final reconstruction, 121,727 and 190,365 particles were from the micrographs obtained from grids with and without graphene oxide film, respectively. After particle polishing, refinement, and post-processing, a final map at 3.79 Å resolution was produced (Supplementary Fig. 3d). A binary mask that emulates TFEα, DNA as well as stalk and clamp domains was applied for focused refinement that continued from the last iteration of the 3D refinement of the entire complex. The overall resolution of the resulting map from focused refinement was 3.85 Å with a slight improvement in local resolution of the masked regions (Supplementary Fig. 3e).

Local resolution estimation of resulting cryo-EM map was performed using MonoRes by submitting half maps for the final reconstruction to the Scipion web tools51. FSC between the unfiltered half maps was calculated using Relion 2.1. In case of RNAP–TFEα and RNAP–TFEα–DNA, the composite maps were produced by combining the last iteration of the final map with the maps resulting from focused refinements using combine_focusedMaps tool in Phenix. FSC between the composite map and the refined atomic model was calculated using phenoct.miavi routine in Phenix (Supplementary Figs. 2f and 3e)52. EM map display and figure preparation was performed using UCSF Chimera53.

Modeling of cryo-EM structures. We built molecular models of the apo-RNAP, RNAP–TFEα binary complex, and RNAP–TFEα–promoter DNA ternary complex into the cryo-EM maps using an X-ray crystal structure of apo Tko RNAP (PDB: 4QW). The X-ray crystal structure of Pfu TFEαXZRD and that of human TFEα (PDB: 5GPP)23 were used as references. The Tko RNAP crystal structure was fitted to the map as a rigid body, and then subunits and the mobile clamp domain were fitted individually. We modeled and extended TFEα and DNA structures manually using Coot54 and further real-space refined the structures against the B-factor sharpened cryo-EM maps from relion_postprocess with Phenix suite52 and validated using ModProb54 (Supplementary Table 2).

In vitro transcription and cross-linking. Mutant TFEα variants containing site-specific Bpa substitutions were made using the pSUP system25,26. Amber codons were created in an N-terminal for His tagged TFEα gene at desired amino acid positions, by primer extension/DpnI mutagenesis. The mutated TFEα variants were expressed from inducible plasmids alongside pSUP-BpaRS-6TRN, in the presence of p-benzoyl phenylalanine. The TFEα variants were then purified by nickel chromatography. Transcription reactions were conducted using purified recombinant TBP, TFB2, and TFEα and native Pfu RNAP (containing a 6x His tag on subunit D, purified by nickel chromatography55). Proteins were assembled at the gdh promoter for 30 min at 65 °C, with RNAP at 10 nM, TBP at 60 nM, TFB2 at 120 nM, and TFEα at 200 nM. Ribonucleotide triphosphates (500 μM GTP, CTP, and ATP, 10 μM [α32P]UTP at ~40 Ci mmol−1) were added to start transcription. Reactions were incubated for an additional 20 min at 65 °C, and then stopped by adding EDTA to 17 mM, along with a small amount of radiolabeled DNA marker (100 bp). Samples were phenol extracted and ethanol precipitated, and transcripts visualized on 14% polyacrylamide gels. See Supplementary Fig. 7 for full scans of the gels.

TFEα cross-linking in preinitiation complexes was tested using t1 site specifically radioabeled gdh promoter DNA, at 10 nM RNAP, 60 nM TBP, 120 nM TFB1, and 300 nM TFEα. Radiolabeled gdh promoter templates were made from biotinylated PCR products attached to paramagnetic streptavidin beads. The non-biotinylated strand was dissociated, and primers annealed and extended with radiolabeled dATP, washed, and fully extended with 500 μM unlabeled dNTPs. The fully extended products were released by restriction enzyme digestion, phenol extracted and ethanol precipitated, and then dissolved and quantified by scintillation counting. Reactions were assembled in 25 microliters overlaid with mineral oil, incubated for 30 min at 65 °C, and cross-links were formed by direct illumination with a handheld UV lamp (365 nm) at 2 cm distance for 60 min. DNA was digested with DNasel and micrococcal nuclease (10 units of each), and samples were loaded on 4–20% gradient polyacrylamide gels55. See Supplementary Fig. 8 for full scans of the gels.

Reporting summary. Further information on experimental design is available in the Nature Research Reporting Summary linked to this paper.

Data availability

The atomic structures of native and ScMet substituted Pfu TFEαXZRD and their structure factors were deposited in the Protein Data Bank with the accession code 6PILN and 6XLIF, respectively. The cryo-EM maps of Tko apo-RNAP, RNAP–TFEα complex, and RNAP–TFEα–DNA complex were deposited in the Electron Microscopy Data Bank with the accession codes EMD-9960, EMD-9961, and EMD-9962, respectively. Their corresponding models were deposited in the Protein Data Bank with the accession codes 6KF3, 6KF4, and 6KF9, respectively. Other data are available upon request. PDB files used for structure comparisons (Supplementary Figs. 5, 6, and 9: 5GPP, 3P8W, 5ONX, 5IYA, 5EZ5, and 5FMF) are from the Protein Data Bank (https://www.rcsb.org).

Received: 27 April 2020; Accepted: 29 October 2020; Published online: 30 November 2020

References

1. Hirata, A., Klein, B. J. & Murakami, K. S. The X-ray crystal structure of RNA polymerase from Archaea. Nature 451, 851–854 (2008).
2. Jun, S. H. et al. The X-ray crystal structure of the euryarchaeal RNA polymerase in an open-clip configuration. Nat. Commun. 5, 5132 (2014).

NATURE COMMUNICATIONS | https://doi.org/10.1038/s41467-020-19998-x | www.nature.com/naturecommunications
3. Korkhin, Y. et al. Evolution of complex RNA polymerases: the complete archaeal RNA polymerase structure. *PLoS Biol.* 7, e1000102 (2009).

4. Hahn, S. Structure and mechanism of the RNA polymerase II transcription machinery. *Nat. Struct. Mol. Biol.* 11, 394–403 (2004).

5. Jun, S. H., Reichlen, M. J., Tajiri, M. & Murakami, K. S. Archaeal RNA polymerase and transcription regulation. *Crit. Rev. Biochem. Mol. Biol.* 46, 27–40 (2011).

6. Grunberg, S. & Hahn, S. Structural insights into transcription initiation by RNA polymerase II. *Trends Biochem.* Sci. 38, 603–611 (2013).

7. Plaschka, C. et al. Transcription initiation complex structures elucidate DNA opening. *Nature* 533, 353–358 (2016).

8. He, Y. et al. Near-atomic resolution visualization of human transcription promoter opening. *Nature* 533, 359–365 (2016).

9. Murakami, K. et al. Structure of an RNA polymerase II preinitiation complex. *Proc. Natl Acad. Sci. USA* 112, 13543–13548 (2015).

10. Dienemann, C., Schwab, W., Schibach, S. & Cramer, P. Promoter distortion and opening in the RNA polymerase II clef. *Mol. Cell* 73, 97–106.e9 (2019).

11. Schulz, S. et al. TFE and Sp4t5 open and close the RNA polymerase clamp during the transcription cycle. *Proc. Natl Acad. Sci. USA* 113, E1816–E1825 (2016).

12. Werner, F. & Grohmann, D. Evolution of multisubunit RNA polymerases in the three domains of life. *Nat. Rev. Microbiol.* 9, 85–98 (2011).

13. Vannini, A. & Cramer, P. Conservation between the RNA polymerase I, II, and III transcription initiation machineries. *Mol. Cell* 45, 439–446 (2012).

14. Brodazzo, F. E., Meinhart, A., Blobel, J. & Whitman, W. B. Genotype-scale analysis of gene function in the hydrogenotrophic methanogenic archaeon *Methanococcus maripaludis*. *Proc. Natl Acad. Sci. USA* 110, 4726–4731 (2013).

15. Meinhart, A., Blobel, J. & Cramer, P. An extended winged helix domain in the archaeal RNA polymerase structure. *Acta Crystallogr D. Biol. Crystallogr.* 76, 2028–2038 (2020).

16. Grunberg, S. et al. The initiation factor TFE and the elongation factor Sp4t5 compete for the RNA clamp during transcription initiation and elongation. *Mol. Cell* 43, 263–274 (2011).

17. Naji, S., Grunberg, S. & Thomm, M. The RPB7 orthologue E is required for transcription in vitro, but a Thermococcus kodakarensis mutant lacking subunit F is temperature-sensitive. *Mol. Microbiol.* 70, 623–633 (2008).

18. He, Y., Fang, J., Taatjes, D. J. & Nogales, E. Structural visualization of key steps in human transcription initiation. *Nature* 495, 481–486 (2013).

19. Hoffmann, N. A. et al. Th1/Th2 cytokine-stimulated complexes and mechanism of transcription initiation. *Nature* 436, 323–330 (2005).

20. Kettenberger, H., Armache, K. J. & Cramer, P. Complete RNA polymerase II elongation complex structure and its interactions with NTP and TFIIH. *Mol. Cell* 16, 955–965 (2004).

21. Feklistov, A. et al. RNA polymerase motions during promoter melting. *Science* 356, 863–866 (2017).

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

23. Chakraborty, A. et al. Opening and closing of the bacterial RNA polymerase core. *Science* 337, 591–595 (2012).

24. Bae, B. et al. CarD uses a minor groove wedge mechanism to stabilize the RNA polymerase open promoter complex. *Elife* 4, e08505 (2015).

25. Davis, E., Chen, J., Leon, K., Darst, S. A. & Campbell, E. A. Mycobacterial RNA polymerase forms unstable open promoter complexes that are stabilized by CarD. *Nucleic Acids Res.* 43, 443–445 (2015).

26. Double, S. Preparation of selenomethionyl proteins for phase determination. *Methods Enzymol.* 276, 523–530 (1997).

27. Ortwonowski, Z. & Minor, W. Processing of X-ray diffraction data collected in oscillation mode. *Methods Enzymol.* 370, 307–326 (1997).

28. Adams, P. D. et al. PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Crystallogr D. Biol. Crystallogr.* 66, 213–221 (2010).

29. Emsley, P. & Cowtan, K. Coot: model-building tools for molecular graphics. *Acta Crystallogr D. Biol. Crystallogr.* 60, 2126–2132 (2004).

30. Zheng, S. Q. et al. MotionCor2: anisotropic correction of beam-induced motion for improved cryo-electron microscopy. *Nat. Methods* 14, 331–332 (2017).

31. Zhang, K. GetF Real-time CTF determination and correction. *J. Struct. Biol.* 193, 1–12 (2016).

32. Kumanius, D., Forsberg, B. O., Scheres, S. H. & Lindahl, E. Accelerated cryo-EM structure determination with parallelisation using GPUs in RELION-2. *Elife* 5, e18722 (2016).

33. Tang, G. et al. EMAN2: an extensible image processing suite for electron microscopy. *J. Struct. Biol.* 157, 38–46 (2007).

34. Scheres, S. H. & Chen, S. Prevention of overfitting in cryo-EM structure determination. *Nat. Methods* 9, 853–854 (2012).

35. Vilas, J. L. et al. MonoRes: automatic and accurate estimation of local resolution for electron microscopy maps. *Structure* 26, 337–344.e4 (2018).

36. Afonine, P. V. et al. New tools for the analysis and validation of cryo-EM maps and atomic models. *Acta Crystallogr D. Struct. Biol.* 74, 814–840 (2018).

37. Petterson, E. F. et al. UCSF Chimera—a visualization system for exploratory research and analysis. *J. Comput. Chem.* 25, 1605–1612 (2004).

38. Chen, V. B. et al. MolProbity: all-atom structure validation for macromolecular crystallography. *Acta Crystallogr D. Biol. Crystallogr.* 66, 12–21 (2010).

39. Dezl, S. et al. Displacement of the transcription factor B reader domain during transcription initiation. *Nature* 359, 10086–10081 (2018).

40. Wang, L., Schmidt, G., Thomm, M. & Hauser, W. Shuttle vector-based transformation system for *Pyrococcus furiosus*. *Appl. Environ. Microbiol.* 76, 3308–3313 (2010).

Acknowledgements

We thank the staff at the Jinju Bio-industry Foundation for supporting Tko cell fermentation and Macromolecular X-ray science at the Cornell High Energy Synchrotron Source (MacCHESS) for supporting the crystallographic data collection. The cryo-EM experiment was performed at the Cryo-EM facility in Electron Microscopy Research Center, Korea Basic Science Institute (KBSI) and supported by KBSI grant C0130221 and C021820. We thank Carmen V. Amoah-Kusi for assistance with cross-linking and crystallography. The cryo-EM experiment was performed at the Jinju Bio-industry Foundation for supporting Tko cell fermentation and Macromolecular X-ray science at the Cornell High Energy Synchrotron Source (MacCHESS) for supporting the crystallographic data collection. The cryo-EM experiment was performed at the Cryo-EM facility in Electron Microscopy Research Center, Korea Basic Science Institute (KBSI) and supported by KBSI grant C0130221 and C021820. We thank Carmen V. Amoah-Kusi for assistance with cross-linking and crystallography. This work was supported by the Basic Science Research Program through the National Research Foundation of Korea funded by the Ministry of Education (NRF-2015R1D1A1A01059907 to S.-H.J.) and by the Ministry of Science, ICT & Future Planning (NRF-2016R1A5A10101674 and NRF-2017M3AF6029755 to H.-S.C. and NRF-2020R1F1A1072050 to S.-H.J.) and NIH grants (R01 GM087350 and R35 GM131860 to K.S.M. and R15 GM083306 to M.S.B.).

Author contributions

S.-H.J., M.S.B., and H.-S.C. conceived and designed the experiments. S.-H.J. and K.S.M. crystallized and solved TFEa crystal structure. S.-H.J., J.S.C., and H.K. prepared Tko RNA and its complexes for cryo-EM. J.H. conducted cryo-EM sample preparation, data collection, and image processing. S.-H.J. and K.S.M. carried out cryo-EM model building. M.S.B. performed cross-linking and transcription assays. K.S.M.,
S.-H.J., H.-S.C., J.H., and M.S.B. wrote the paper. All authors discussed the results and commented on the paper.

Competing interests
The authors declare no competing interests.

Additional information
Supplementary information is available for this paper at https://doi.org/10.1038/s41467-020-19998-x.

Correspondence and requests for materials should be addressed to S.-H.J., H.-S.C. or K.S.M.

Peer review information Nature Communications thanks Finn Werner and the other, anonymous, reviewer(s) for their contribution to the peer review of this work. Peer reviewer reports are available.

Reprints and permission information is available at http://www.nature.com/reprints

Publisher’s note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.