Dynamic positioning of a TFIIE-related winged-helix domain in RNA polymerase elongation complex III underpins its multi-functionality

Jheng-Syong Wu[1], Yu-Chun Lin[2], Yi-Yu Wei[2], Hsin-Hung Lin[1], Yang-Chih Liu[1], Jen-Wei Chang[1], I-Ping Tu[3], Hung-Ta Chen[2,*] and Wei-Hau Chang[1,4,*]

[1] Institute of Chemistry, Academia Sinica, Taipei, 115, Taiwan
[2] Institute of Molecular Biology, Academia Sinica, Taipei, 115, Taiwan
[3] Institute of Statistical Science, Academia Sinica, Taipei, 115, Taiwan
[4] Institute of Physics, Academia Sinica, Taipei, 115, Taiwan

* Correspondence should be addressed to Wei-Hau Chang. Email: weihau@gate.sinica.edu.tw

Author Note:

Present Address of Dr. Jen-Wei Chang: Development Center for Biotechnology, Taipei, 115, Taiwan

Abstract RNA polymerase III (Pol III) is a large protein complex that efficiently synthesizes structural and non-translated RNAs. The tandem-wing-helix (tWH) domain (WH1 and WH2) in the Rpc34 subunit of the TFIIE-related sub-complex plays multiple roles central to Pol III transcription, but the structural basis remains conundrum due to its cryo-EM density missing in enlongation complex (EC). Here, by site-specifically incorporating an azido unnatural amino acid (UAA) into Pol III Rpc34 WH2 and developing a simple protocol to enable selective labeling using strain-promoted azide-alkyne cycloadditio—the Bertozzi click reaction, we probed Rpc34 WH2 dynamics using single-molecule Förster resonance energy transfer (smFRET) between an acceptor in Rpc34 WH2 and a donor in DNA. Our smFRET resolved multiple FRET states with inter-state transitions exhibiting characteristic kinetics. Further nano-positioning analysis localized the docking positions of those FRET states on Pol III EC, including sites related to initiation and transcription bubble management. Our findings of Rpc34 WH2 dynamical positioning suggest the docking interactions are weak and promiscuous, providing a mechanism to rationalize multi-functionality of Pol III Rpc34 tWH. This study expands our understanding of Pol III transcription machinery and our chemical biology means are widely applicable to explore structural dynamics of large protein complexes.

INTRODUCTION

RNA polymerase III (Pol III) is a highly efficient nuclear RNA polymerase that synthesizes structural and non-translated RNAs required for essential processes in eukaryotic cells, necessitating high efficiency. Pol III-synthesized RNAs are short (a few hundred nucleotides), encompassing tRNAs, 5S rRNA, U6 snRNA, and 7SL RNA, among others (1,2). The Pol III machinery is composed of 17 subunits, forming a complex of 0.7 MD in molecular weight. Having presumably evolved from the core machinery of RNA polymerase II (Pol II) (3,4), Pol III acquired five additional subunits. Those Pol III-specific subunits are organized into two sub-complexes, the Rpc82/34/31 heterotrimer and Rpc53/37 heterodimer (2). As these two sub-complexes are related to initiation factors TFIIE and TFIIF in the Pol II system (3,4), they exert initiation functions similar to those of the Pol II initiation factors. However, since both sub-complexes are stable components of Pol III, they may serve additional functions as "built-in" elongation factors (2). To date, the molecular basis for such multi-functionality remains poorly understood.

Among those Pol III-specific subunits, yeast Rpc34 (RPC6 in human) in the TFIIE-related heterotrimeric sub-complex is unique since it has no counterpart in the other two nuclear RNA polymerases (Pol I and II). Rpc34 is a small protein comprised of 317 amino acids. Functional studies *in vivo* using yeast have revealed its importance (5), uncovering dual roles in the initiation stage, including promoter DNA opening (5) and formation of the pre-initiation complex (PIC) (5,6). The amino acids (aa) responsible for those dual roles have been mapped to the protein's N-terminal region (5), now known to adopt a tandem winged-helix (WH) fold based on homology analysis with TFIIE (7). The WH1 (aa 12 to 78) and WH2 (aa 90 to 151) elements are together referred to as the tandem WH (tWH) domain. Recently cold-sensitive mutants of Rpc34 WH1 revealed a previously unidentified function of the tWH in the elongation stage (8). Together, those studies demonstrate that Rpc34 tWH exerts different functions at various stages of transcription.

Early efforts to position those Pol III-specific subunits in Pol III were made by Chen and colleagues using photo-activated crosslinking (PA-XL) (9), prior to the emergence of high-resolution Pol III structures. The resulting PA-XL model (9) provisionally placed Rpc34 tWH above the Pol III DNA-binding cleft (Supplementary Fig. S1), perhaps to best reflect its function in promoter DNA opening at the stage of initiation (5). Interestingly, subsequent high-resolution cryo-electron microscopy (cryo-EM) structures of the Pol III pre-initiation complex (PIC) (10,11) shifted the position of Rpc34 tWH slightly upstream along the DNA-binding cleft towards the Rpc128 protrusion domain (Supplementary Fig. S1). At this position, Rpc34 tWH is mainly stabilized by the promoter-bound initiation factor TFIIIB, conferring a structural basis for its function in PIC formation (5). This site can be alternatively occupied by Maf1 instead of TFIIIB, thereby suppressing Pol III transcription initiation (12). Intriguingly, the density of Rpc34 tWH in the cryo-EM structure of a ternary elongation complex (TEC) is obscure (13). This "loss of signal" indicates that Rpc34 tWH is structurally flexible in the elongation stage, but definitive evidence for such dynamism has been lacking. Notably, tWH structural flexibility is potentially conserved from yeast to human, as the tWH density in RPC6 is also absent from the corresponding cryo-EM structure (14). Nonetheless, despite advances in cryo-EM technology, presumbed dynamics within Pol III remain largely obscure as, essentially, it is an ensemble method.

To probe the dynamics of Rpc34 tWH in Pol III EC, we adopted a single-molecule fluorescence resonance energy transfer (smFRET) strategy. SmFRET is a powerful approach for illuminating the structural dynamics hidden in ensemble measurements (15-17). When smFRET is applied to immobilized molecules, it tracks the time-courses of individual molecules so that kinetics information can be obtained without the need for synchronization (15). However, applying smFRET to a large-sized native protein faces the issue of site-specific fluorophore labeling. The popular labeling method of using nucleophile-reactive maleimides requires that fluorophores be attached to introduced cysteine residues, with the naturally-occurring cysteines removed by means of mutagenesis, so it is thereby impractical or even damaging for Pol III, a large multi-subunit protein with a total of 65 naturally-occurring cysteine residues. Therefore, we sought a bio-orthogonal strategy, ultimately selecting the approach of unnatural amino acid (UAA) mutagenesis (18-20). A yeast genetic system for UAA incorporation has been previously built (20) and utilized for site-specifically incorporating a photoactive UAA into yeast nuclear RNA polymerases (9,21-22). This system could be minimally modified to suit our needs. Nevertheless, successful cases of using UAA for smFRET have been restricted to small test proteins (23-24) or a subunit of an in vitro-reconstituted protein assembly (25). Thus, it remained to be determined if such approach was readily feasible for labelling a native protein assembly as large as Pol III.

To employ UAA-based smFRET and probe Rpc34 tWH dynamics in a transcribing Pol III, we site-specifically incorporated an azido-UAA (26) (see Figure 1) into the part of WH2, and conjugated the UAA with an alkyne-carrying fluorophore via strain-promoted azide-alkyne [3+2] cycloaddition (SpAAC), now best known as Bertozzi's reaction (26). This simple click chemistry is bio-friendly as it eliminates perilous copper usage (26), thus readily suitable for bio-orthogonal applications (27). Surprisingly, our bulk labeling experiments showed extensive non-specific labeling of Pol III subunits other than the targeted UAA-substituted Rpc34, so we developed a chemical protocol for selective labelling of the incorporated UAA by suppressing labelling of those nature-occurring cysteine residues. To monitor smFRET between the two fluorescent probes in the donor-acceptor-labeled Pol III ternary elongation complex (TEC) (see Figure 2), in which the acceptor was conjugated to the UAA in Rpc34 tWH and the donor was placed in the template DNA (T-DNA) slightly downstream of the active site. It is noted that Rpc34 WH2 can be stabilized by Maf1 (12), for which observation of stationary smFRET is anticipated and can thus be used as a control. Our smFRET results of multiple FRET states demonstrated Rpc34 tWH is indeed a mobile structural element in Pol III TEC. We further used nano-positioning analyses to localize those FRET states. Those identified sites have implications in involvement of different functions in Pol III transcription and thus may provide a structural basis for explaining the multi-functionality of Rpc34 tWH in Pol III transcription (5,8).

RESULTS

Capping native cysteines enables Bertozzi reaction to selectively target the incorporated azido-UAA

To incorporate 4-Azido-L-phenylalanine (AzF) into Rpc34 in native Pol III, we modified our previously developed yeast UAA system for BPA incorporation (9,21) (Figure 1A & **Methods**). We tested various AzF-substituted sites in Rpc34 — none of those single amino-acid mutants exhibited a detectable phenotype. Successful AzF incorporation was evidenced by cell growth dependency on AzF following genetic modification of Rpc34.

We purified wild type and mutant Pol III bearing a single AzF at a Rpc34 residue, and labelled it with DIBO-Alexa-647 (DIBO, dibenzocyclooctyne, a strained alkyne) using a click chemistry based on strain-promoted alkyne cyclooctyne addition (SpAAC), best known by its nickname of "Bertozzi" click chemistry (Figure 1A). So far, limited number of smFRET studies utilized this labeling approach. As those studies are restricted to small test proteins (23-24), or simple protein complexes that can be assembled *in vitro* using recombinant proteins (25), it has remained unclear if this Bertozzi labeling approach is readily feasible for a large or endogenous protein complex such as Pol III.

To our dismay, direct application of the existing protocol for AzF labeling (25) yielded extensive DIBO-labelling of wild type Pol III not substituted with AzF (see *lane*s 1, Figure 1C). Those "non-specifically" labelled subunits included Rpc1 (Rpc160), Rpc2 (Rpc128), Rpc82, Rpc37/AC40, and ABC27. Given that those subunits possess abundant native cysteines (Supplementary Table S1), we hypothesized that those unwanted labelling mainly arose from some of native cysteines reacting with the alkyne moiety on DIBO (Figure 1A). This notion is supported by a previous report that showed azide-dependent protein labelling with a strain-promoted alkyne-dye via cyclooctyne addition was compromised by thiolyne addition reaction (27,28) (Figure 1A). We then sought usage of small molecules (29) that could cap the thio-moiety of Pol III cysteines (Figure 1B). We tested a number of compounds—iodoacetamide (IAM), N-ethylmaleimide (NEM) and methyl-methanethiosulfonate (MMTS)—by evaluating their impact on labelling AzF-substituted Pol III with respect to wild-type Pol III as a control. As shown in lanes 3-8 of Figure 1C, use of all those thio capping molecules at 20 mM virtually abolished those non-specific labelling and enabled selective labelling of Rpc34 subunit. Since IAM also significantly reduced specific Rpc34 labelling but either NEM or MMTS permitted a good level of specific labeling of Rpc34 (Supplementary Figs. S2-S4), we dispensed with further usage of IAM. To use NEM or MMTS, we lowered the concentrations to 20 μM for subsequent experiments as we observed high concentrations had interfered with formation of the Pol III TEC, evidenced by disappearance of the respective band (>720 kD) in the native gel (Supplementary Figs. S2-S4).

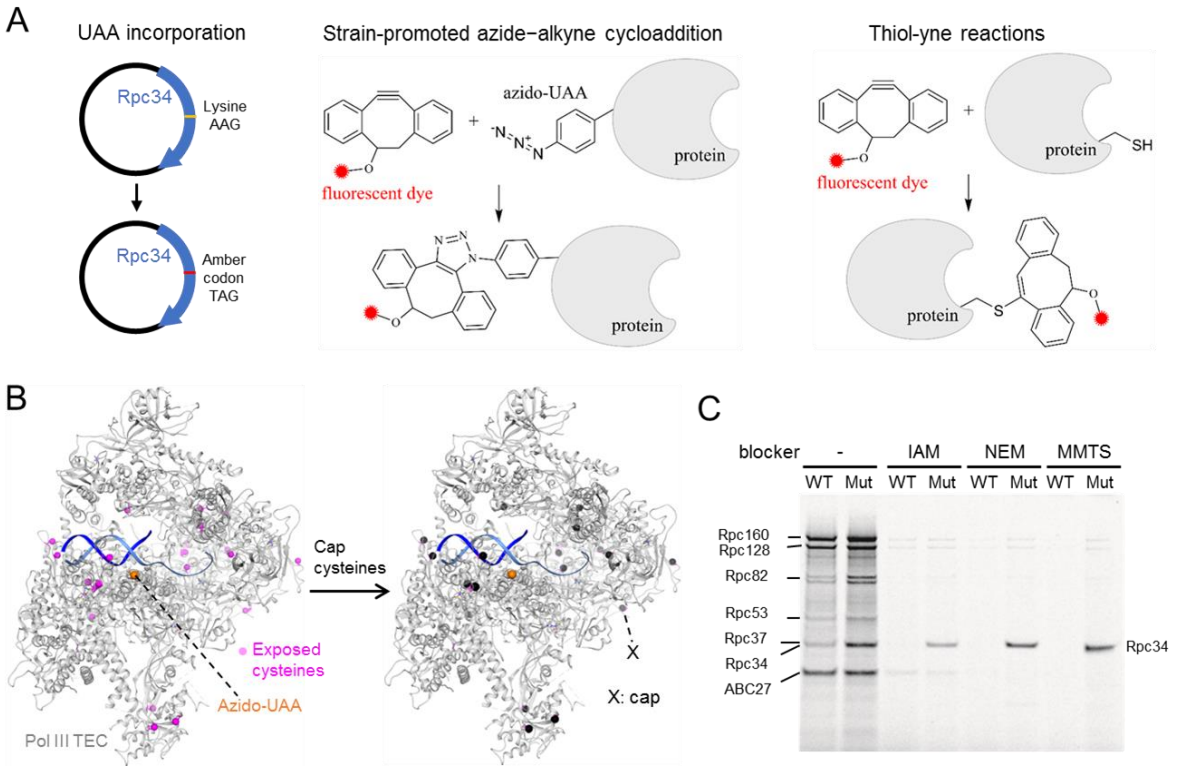

Figure 1. Selective dye labeling of incorporated azido-UAA by capping native cysteines on Pol III TEC. (**A**) *Left*: Schematic of UAA incorporation to Rpc34 WH2 in Pol III. Rpc34/pRS425 was transformed through plasmid shuffling into a yeast strain lacking chromosomal Rpc34 and a specific codon was mutated to amber codon for UAA incorporation at the specific position. *Middle*: Chemical structures of a DIBO fluorophore (DIBO: dibenzocyclooctyne, a strained alkyne) and the azido-phenylalanine UAA moiety on a protein that can be conjugated together using strain-promoted azide-alkyne [3+2] cycloaddition (SpAAC). *Right*: The reaction of thiol-yne addition conjugates the free thiol groups of cysteines and alkyne-dye resulting in non-specific labeling on Pol III. (**B**) The schematic diagram of cysteines blocking illustrates the exposed cysteines (magenta sphere) on Pol III are capped by thio-reactive molecules (black sphere) in order to inhibit non-specific dye conjugation on Pol III. (**C**) Strain-promoted dye labeling and screening of cysteine blockers that suppress non-specific dye labeling. Thio-reactive molecules (iodoacetamide (IAM), N-ethylmaleimide (NEM) and S-methyl methanethiosulfonate (MMTS)) at a concentration of 20 mM were tested. Alexa647 fluorescence image of a SDS-PAGE gel of wild-type Pol III and mutant Pol III for UAA incorporation (Rpc34 L18) shows that NEM and MMTS largely abolished non-specific dye labeling and yet permitted specific labeling of Rpc34 (Supplementary Figs. S2-S4).

SmFRET reveals Rpc34 WH2 is mobile

We first chose to place the donor dye at DNA(+7) downstream of the active center. This location bears high structural order, as evidenced by its well-resolved density in the cryo-EM structure (13) (yellow circle in Figure 2A) and spans a point-to-point distance of approximately 6 nm in known structures (9-11). It is noted that this distance is translated to dye-to-dye distance of 7 nm (30). With the donor dye (TAMRA) at DNA(+7) and acceptor dye labelled at K126 position in Rpc34 WH2, we set out to use smFRET in a Pol III-EC complex to test whether or not Rpc34 WH2 is mobile.

We first considered a control case where Rpc34 WH2 remains still when stabilized by a Maf1 transcription factor. Maf1 binds to pol III near the protrusion domain (12) at a location occupied by TFIIIB during initiation and thereby represses pol III transcription. As the single-molecule experiment was performed at pico-molar concentration range, we added 10 X Maf1 to donor-acceptor-labelled Pol III TEC immobilized on a cover slip. We performed imaging with our custom-built total internal reflection fluorescence microscope equipped with alternating laser excitation (ALEX) capability (see **Methods**). Based on the cryo-EM structure of Pol III-Maf1 (12), we built a Pol III-Maf1 TEC model (Figure 2B), in which the donor dye at DNA(+7) and the acceptor dye at Rpc34 (K126) span a distance of 7 nm (30) to predict a FRET efficiency E_{model} of 0.26 for the dye pair of TAMRA and Alexa-647 with Förster distance of 6 nm (31,32). From this control experiment that included Maf1, we harvested 85 time traces with clear acceptor photo-bleaching signatures (33), and built a frame-wise histogram, which is predominated by single Gaussian mode centered at 0.26 (Figure 2C), agreeing well with the predicted efficiency. It is noted small number of Pol III TEC molecules contributed to higher FRET (E_{mean} > 0.5) as they were not bound by Maf1.This control experiment demonstrates our smFRET system is highly reliable.

We next performed smFRET measurement on the donor-acceptor-labelled TEC molecules by excluding Maf1. We harvested a total of 110 time traces; among them, 48 molecules exhibit stationary FRET, but with three distinct levels---high, middle and low, as shown by three presentative de-noised FRET time traces (34) (Figure 2D). Interestingly, the remaining 62 molecules exhibit inter-level switching transitions among those three FRET levels. Guided by these three levels, we separated the stationary time-traces and the levels in the switching time-traces from the 110 molecules into three classes to build a frame-wise histogram that comprises three sub-histograms, each of which correspond to a distinct level (Figure 2E). This led to precise extraction of three mean FRET efficiencies, high FRET (E_{mean}=0.68), middle FRET (E_{mean}=0.52), and low FRET (E_{mean}=0.26). These results immediately suggest that in a Pol III TEC Rpc34 WH2 can dock on locations with varied proximity to the DNA (+7) site, and explain why Rpc34 WH2 density failed to be detected in an Pol III TEC by single particle cryo-EM reconstruction (13). Those FRET proximities correspond to different locations on Pol III, including a location (E_{mean}=0.26) likely to be similar to that stabilized by TFIIIB in ITC (10,11) or Maf1, and two other locations (E_{mean}=0.52 and 0.68) that shift downstream along the DNA channel.

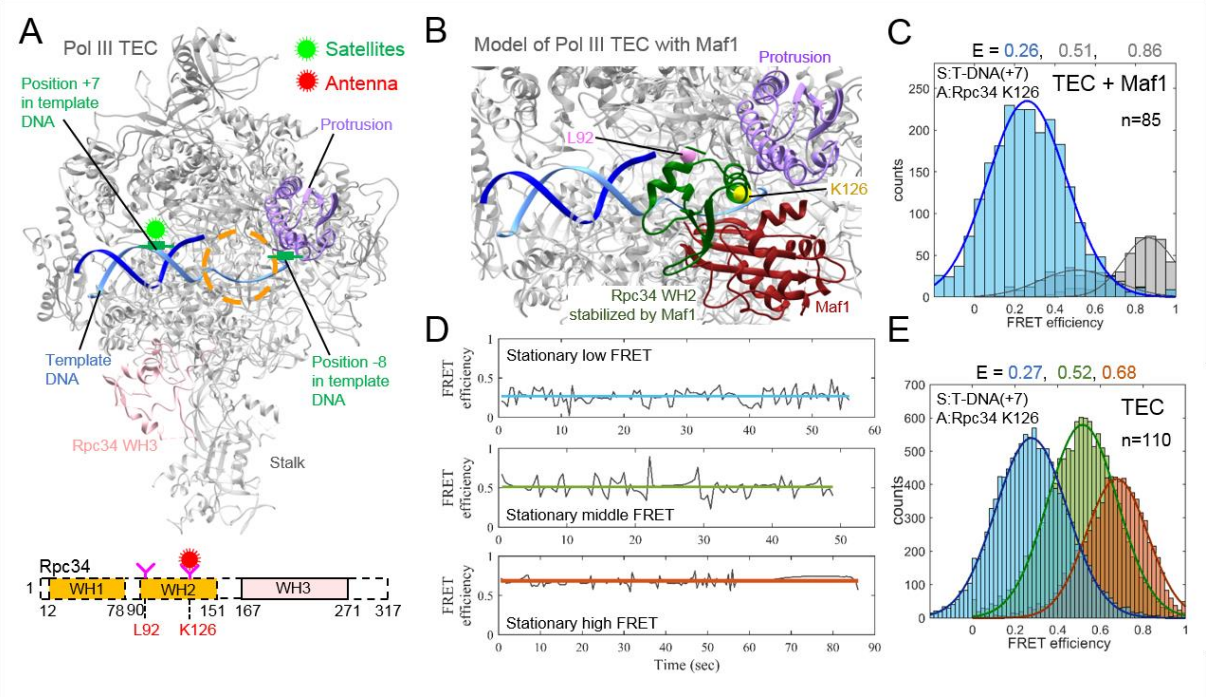

Figure 2. Design, control and measurements of the FRET pair for probing the kinetics and position of Rpc34 WH2 in Pol III. (**A**) The design of the FRET pair based on the yeast Pol III TEC model (PDB: 5fj8, gray), in which the cryo-EM densities of Rpc34 tWH is obscure. The yellow dashed line encompasses the region in the apo Pol III map (PDB: 5fja) that was speculated to correspond to Rpc34 tWH. The satellite dye (TAMRA) and antenna dye (Alexa647) were labeled at the +7 position of the template DNA (T-DNA) (cyan) and Rpc34 K126 (in WH2), respectively (called as AS pair +7/126). The other labeling sites for the satellite and antenna are the -8 position of T-DNA and Rpc34 L92, respectively. For clarity, the Rpc82 and Rpc31 subunits in Pol III TEC model is computationally removed where the Rpc2 (Rpc128) protrusion domain and Rpc34 WH3 domain are represented by light purple ribbons and pink ribbons, respectively. (**B**) The control of smFRET measurement in the same content of Pol III TEC except the repressor Maf1 was added to lock Rpc34 WH2. Rpc34 WH2 (PDB: 6tut, dark green) is stabilized by Maf1 (dark red). The yellow ball and pink balls indicate the residues K126 and L92 of Rpc34. (**C**) The FRET efficiency histogram of the AS pair +7/126 in Pol III TEC including Maf1. The major population in the FRET efficiency histogram is low FRET as that we predicted (Supplementary Fig. 1). E: the center of FRET efficiency distribution; S: Satellite; A: Antenna; n: molecule number. (**D**) Representative time trajectories of stationary low-, middle- and high-FRET states of Rpc34 WH2 in Pol III TEC (AS pair +7/126). The three lines (blue, green and orange) represent the average FRET efficiencies of different states. These traces were smoothed by the denoising algorithm with wavelets and Bayesian inference. (**E**) The FRET efficiency histograms of low- (light blue), middle- (light green) and high- (orange) FRET states of Rpc34 WH2 in Pol III TEC (AS pair +7/126). All frame data of 110 smFRET time trajectories (including stationary and state-changed trajectories) were classified to low-, middle- and high-FRET based on Hidden Markov modeling (HMM) analysis of traces and FRET efficiencies. The FRET efficiency histograms of three states were built up separately. Each histogram was fitted well with single Gaussian function, and the center of low-, middle- and high-FRET efficiency (E) distribution is 0.27, 0.52 and 0.68, respectively. The R^2 value of curve fitting is 0.996, 0.991 and 0.985 for the low-, middle- and high-FRET states.

Dwell-time analysis characterize FRET state life-times

Those inter-converting smFRET traces directly captured the movement of the Rpc34 WH2 domain. Nonetheless, whether or not the three distinct FRET levels reflect three FRET states remained unclear. To address this question, we deployed dwell-time analysis to characterize the life time of each FRET level.

To do so, we used hidden Markov modeling (HMM) analysis (35) to detect the time-point when a FRET level change occurs along a time trace and obtain the time duration of the FRET level between changes (see **Methods**). Assisted by ALEX imaging, we were able to eliminate artificial FRET level changing due to acceptor quenching (Supplementary Fig. S5). Our HMM analysis led to de-noising of FRET levels in those inter-level switching time traces, and independently gave rise to the estimates for low, middle and high FRET efficiencies (Figure 3A) consistent with the mean FRET efficiencies extracted from the aforementioned frame-wise histogram analysis (Figure 2E). By plotting histogram of the dwell times for each level, we obtained three individual dwell-time distributions. Remarkably, each distribution could be fitted well to a single exponential decay with one characteristic life time (Figure 3B), indicating there are virtually no apparent sub-states. The single exponential decay behaviour shows the transitions follow a simple first-order kinetics of bimolecular interactions. We thereby dubbed these FRET levels as three FRET states. However, those states are observed from a particular site, e.g. DNA(+7), and there could be other states that are missed.

To test whether or not those life-times do reflect molecular characteristics, we performed repeatedly measurement on those dwell times with three sets of independent experiments. We found the single exponentials and life-times were remarkably reproducible, where the high FRET state was found to have the longest life-time (9.37 sec), the middle FRET state the intermediate life-time (3.12 sec), and the low FRET state the shortest life-time (1.47 sec). To further understand if some inter-state transitions are preferred, we plotted a two-dimensional histogram (Figure 3C). Intriguingly, there is no direct transition between the low and the high FRET states while every transition must involve the middle FRET state. By plotting the dwell-time for the middle to high and the middle to the low separately, we were able to solve kinetics equations to obtain inter-state transition rates (see Figure 5).

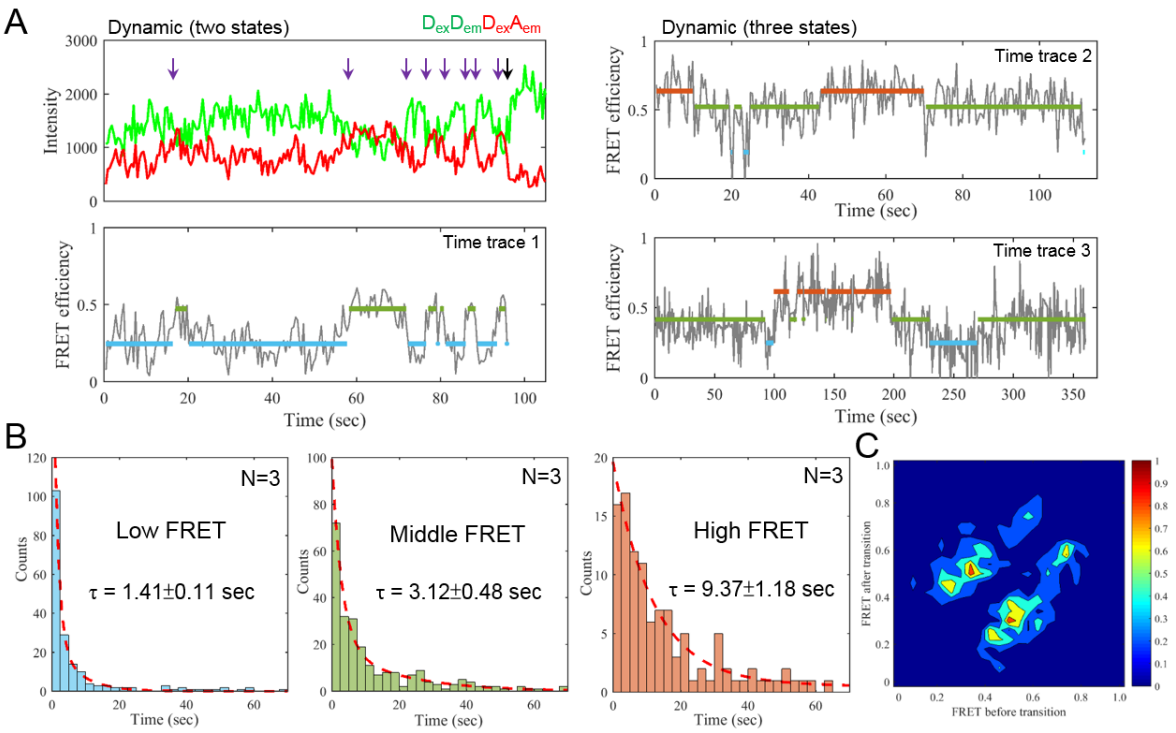

Figure 3. Representative time trajectories of inter-converting FRET and dwell time analysis of Rpc34 WH2 in Pol III TEC. (**A**) The satellite dye (TAMRA) and antenna dye (Alexa647) were labeled at the +7 position of the template DNA and Rpc34 K126, respectively (AS pair +7/126). *Left*: One representative two-states smFRET time trajectories of Pol III TEC is validated with the anti-correlated signals of donor and acceptor intensities. The green and red traces represent the detected donor (D_{em}) and acceptor (A_{em}) intensities, respectively, with alternating laser excitation (ALEX) using 532 nm (D_{ex}) and 638 nm (A_{ex}) wavelength laser excitations. The additional acceptor excitation was used to verify that the acceptor is active, thereby yielding meaningful anti-correlated signals of the donor and acceptor when acceptor photo-bleaching occurs. The corresponding time traces of FRET efficiencies (gray trace) are calculated from the intensities of the donor and acceptor based on individual γ normalization. Hidden Markov modeling (HMM) was applied to the intensity time traces to identify the level of each FRET state (low-FRET: light blue, middle-FRET: light green) and their respective durations. Purple and black arrows denote the anti-correlated signals when the FRET state changes and the event of acceptor photo-bleaching, respectively. *Right*: Two representative smFRET time trajectories of Pol III TEC exhibit vast fluctuations among different FRET levels. The frame data of each trajectory are classified to low-, middle- and high- (orange) FRET state based on the HMM analysis and average FRET efficiency. (**B**) The dwell time histograms for low-, middle-, and high-FRET efficiencies were built from the HMM analysis of 61 state-changed traces. The average dwell time and the standard deviation are calculated from three individual smFRET measurements and dwell time analysis. The average dwell times are 1.41±0.11 sec in the low-FRET state, 3.12±0.48 sec in the middle-FRET state and 9.37±1.18 sec in the high-FRET state. (**C**) The transition density plot built up from 633 transition data extracted from 82 state-change traces illustrate that the major and minor populations are the transitions between middle-FRET state and low-FRET state, and between middle-FRET state and high-FRET state, respectively.

Nano-positioning analysis localizes FRET state anchoring sites

Those FRET states with different mean FRET efficiencies reflect different proximity of Rpc34 K126 to the DNA(+7) position, as the low FRET is referred to as the distant site, the high FRET is the proximal site, and the middle FRET situated in between as the middle site. To determine the location of those sites in Pol III TEC, we employed nano-positioning system (NPS) analysis (36) a low-resolution FRET-based structure determination method. This method requires multiple antenna-and-satellite (AS) pairs to generate a set of FRET-derived distance constraints to suit triangulation analysis for predicting the position of an unknown site (antenna) in reference to those of known sites (satellites). This method has proven effective in mapping disordered structural elements in a protein complex (32).

Our design of the AS pairs is described in Figure 4A. We utilized satellites at DNA(+7) and DNA(-8) to locate the antenna at K126 or L92 of Rpc34 WH2. We chose to place another antenna at L92, which is on the opposite site of K126 in Rpc34 WH2. By knowing the positions of these two antennas, we should be able to more precisely locate the center of Rpc34 WH2 domain. To establish a reliable NPS system, we considered the case of antenna at K126 of Rpc34 WH2 stabilized by Maf1. With (S/A: +7/126 = 0.26, S/A: -8/126 = 0.55) (see Figure 2C for +7/126 and Supplementary Fig. S6 for -8/126), we determined the credible volume determined for a dye attached to Rpc34 K126 as shown in Figure 4C. This volume matches well with that obtained by modelling (Supplementary Fig. S7) where the error is only 0.69 nm along the long axis of the volume as anticipated from the choice of particular satellite positions.

We next performed FRET measurements for multiple satellite-antenna pairs as described in Figure 4A. The smFRET results are summarized in the FRET histograms, each of them exhibit multiple modes (Figure 4B). With satellites on DNA(+7) and (-8), we first derived three meaningful NPS credible volumes for the dye attached to K126 (yellow in Figure 4D) based on the following combinations of Satellite(DNA)/Antenna(C34) and FRET efficiency (S/A: +7/126 = 0.27, S/A: -8/126 = 0.78), (S/A: +7/126 = 0.52, S/A: -8/126 = 0.48), and (S/A: +7/126 = 0.68, S/A: -8/126 = 0.23). We next derived three meaningful NPS credible volumes for the dye attached to L92 (pink in Figure 4D) based on the following combinations (S/A:+7/92 = 0.53, S/A: -8/126 = 0.59), (S/A: +7/92 = 0.72, S/A: -8/92 = 0.39), and (S/A: +7/92 = 0.86, S/A: -8/92 = 0.25).

With the pair of L92 and K126 Antenna, we were able to precisly localize Rpc34 WH2. As a result, the aforementioned low FRET state of Rpc34 WH2 in reference to DNA(+7) is found to be located to a region next to the protrusion domain and potentially overlap with the site during transcription initiation when Rpc34 WH2 is stabilized by TFIIIB (10). We therefore renamed this site as an initiation-prone site (Figure 5). Compared to the initiation-prone site, the high and middle FRET states in reference to DNA(+7) are situated more downstream along the DNA channel. Intriguingly, the middle FRET state in reference to DNA(+7) is situated at the head of the transcription bubble, for which we renamed this site as the bubble-head site. This finding suggests a role of Rpc34 WH2 in transcription bubble opening or managing its open state (Figure 5), providing a structural basis to support the notion of Rpc34 WH2's capability in promoter opening (5) or stabilizing the transcription bubble during elongation. The high FRET state in reference to DNA(+7) is situated further downstream of the transcription bubble to cover dsDNA, for which we renamed this site as the

downstream site where Rpc34 WH2 rest for the longest time. Together with the FRET state changes, our data and analysis have shown that any transition between the downstream site and the initiation-prone site will always transit through the bubble-head site as an intermediate site (Figure 5).

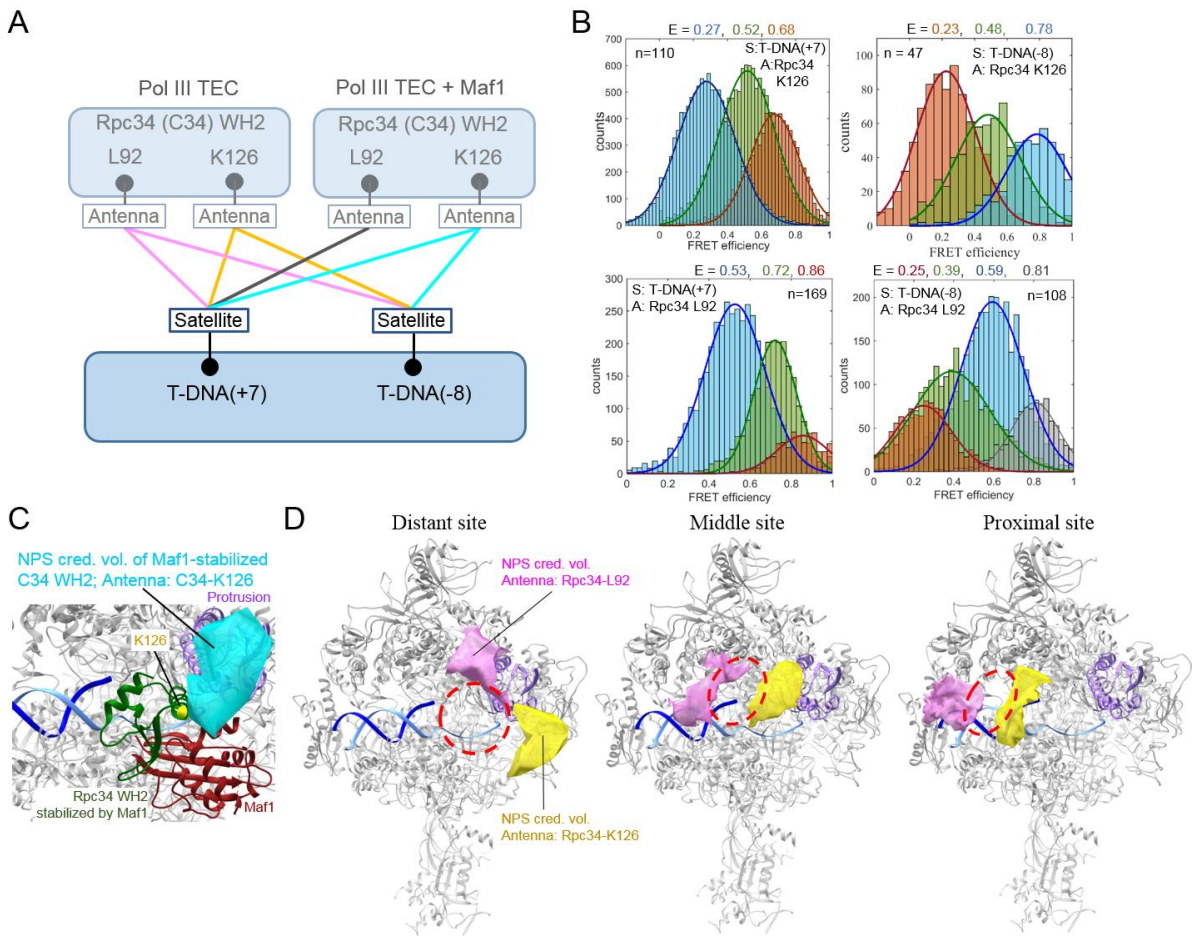

Figure 4. The FRET efficiency histograms and nano-positioning localization of low-, middle- and high-FRET states of Rpc34 WH2 in Pol III TEC. (**A**) The schematic representation of FRET network for localizing Rpc34 WH2 in Pol III TEC. (**B**) The FRET efficiency histograms of four antenna-and-satellite (AS) pairs (S/A: +7/126, -8/126, +7/92 and -8/92) for localizing Rpc34 WH2 in Pol III TEC. In the AS pair -8/126, the satellite dye (TAMRA) and antenna dye (Alexa647) were labeled at -8 position in T-DNA and Rpc34 K126. Similarly, the FRET pairs +7/92 and -8/92 means the satellite dye and antenna dye were labeled at +7 or -8 position in T-DNA and Rpc34 L92. The colors light blue, light green and orange in FRET efficiency histograms correspond to the low-, middle- and high-FRET states of Rpc34 WH2 based on the AS pairs +7/126 (Figure 2A). E: the center of FRET efficiency distribution; S: Satellite; A: Antenna; n: molecule number. (**C**) Nano-positioning localization of Rpc34 WH2 stabilized by Maf1 on Pol III by using nano-positioning system (NPS) analysis (38,39) (see **Methods**) and the smFRET measurements (the major peak in FRET efficiency histogram) of two AS pairs (+7/126 and -8/126) (Fig. 2C and Supplementary Fig. S6). The NPS credible volume of the antenna dye colored in cyan positioned on the protrusion (light purple) and beside the Maf1-locking Rpc34 WH2 is similar to the simulated dye volume based on the position and the orientation of Maf1-locking Rpc34 WH2 in Pol III. (**D**) Nano-positioning localization of Rpc34 WH2 on Pol III by using two NPS credible volumes of which the antenna dye were labeled at Rpc34 L92 or K126. The NPS analysis of antenna dye at Rpc34 K126 (yellow volume) utilize smFRET measurements of AS pair +7/126 and -8/126. The NPS credible volume of antenna dye at Rpc34 L92 was analyzed by using smFRET measurements of AS pair +7/92 and -8/92. The red dashed circles between the two NPS credible volumes denote the locations of Rpc34 WH2 at the three different states because L92 and K126 are at the opposite site of Rpc34 WH2.

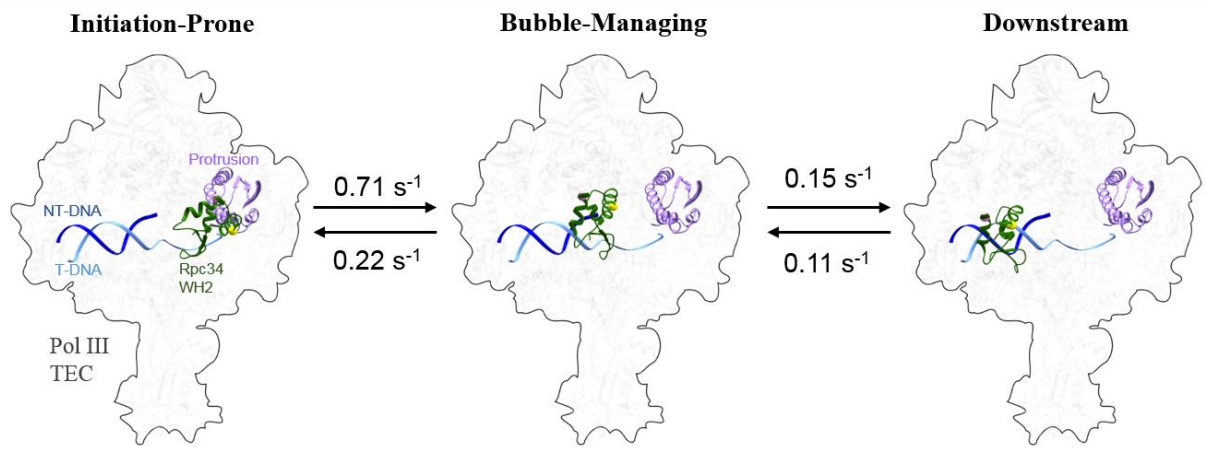

Figure 5. A model of dynamic positioning of Rpc34 WH2 in Pol III elongation complex. Based on nano-positioning analysis with various FRET states (Figure 4D), a model that Rpc34 WH2 (dark green) can dock on various sites of Pol III is proposed. In a dynamically positioning scenario that reflects inter-FRET state transitions, Rpc34 WH2 can switch among those positions. The location of the initiation-prone site corresponding to the low-FRET state is localized at the cleft entrance to lie next to the Rpc2 protrusion domain (light purple), similar to the position of Rpc34 WH2 in Pol III PIC. The bubble-managing site based on the middle-FRET state resides near the cleft and is close to DNA bubble. The downstream site based on the high-FRET state is in the DNA-binding cleft and close to downstream DNA. The yellow ball and pink ball indicate the residues K126 and L92 of Rpc34. The template-DNA (T-DNA) and non-template DNA (NT-DNA) are represented by cyan and blue, respectively. The transition rates are obtained from the dwell time analysis of state-change traces (Figure 3).

DISCUSSION

The TFIIE-related sub-complexes of Pol III remain tightly associated with Pol III during elongaiton, in contrast to TFIIE that dissociates from Pol II when transiting from initiation to elongation. It is possible that the stably associated sub-complex contributing to Pol III elongation function or serves as a "built-in" elongation factor (2). The density and thereby positioning of this sub-complex in the initiation stage has been nicely revealed by high-resolution cryo-EM (10,11). However, it remained unclear if this stably-associated sub-complex occupies the same position as transcription proceeds. Intriguingly, a recent analysis of puzzling crosslinking data from holo- and transcribing Pol III by Nilges and colleagues (37) suggests the TFIIE-related sub-complex is potentially mobile as their results indicated that the largest subunit of this sub-complex (Rpc82) could undergo large-scale re-arrangement. Nonetheless, potential movements of other smaller components in this sub-complex (Rpc34 and Rpc31) of similarly large scale have evaded detection (37) (Supplementary Figure S1B). Capturing the movement of a flexible module within a macromolecule has proven extremely challenging for conventional structural methods, despite advances in cryo-EM heterogeneity analysis (38) and crosslink (37) analyses, and this difficulty is exacerbated when an intermediate is excessively short-lived or a structural component is too flexible, as is the case for Rpc34 tWH, where the flexibility might be correlated with its multi-functionality. Here, we employed smFRET analysis to reveal the dynamical positioning of Rpc34 WH2 in a Pol III EC, conferring the first glimpse of Pol III as a dynamic machine (Figure 5). Those findings reveal the spatial and temporal behaviour of Rpc34 WH2 in pol III EC with recovery of the localization information of Rpc34 WH2 in Pol III EC missed by cryo-EM studies (13) to support a hypothesis of a different function at a different location.

Considering tWH motif, it is a common module in three nuclear RNA polymerases (39,40). For Pol I (39), tWH is hosted by the Rpa49 subunit of the stably-associated TFIIF-related sub-complex. Compared to Rpc34 tWH in Pol III, Rpa49 tWH density was resolved by cryo-EM at different stages during early transcription transitions (41). In those snapshot structures, this tWH domain moves to different positions for serving different functions. In an open complex (OC2), Rpa49 tWH is bound to the upstream DNA at a TFIIE-like position for promoter recognition or promoter opening. As transcription proceeds to elongation with RNA reaching a critical length, the tWH shifts to a new position where it displaces Rrn3, enabling Pol I to escape from the promoter. Thus, the structural findings for the tWH domain of Pol I (41) together with that presented in this study (Figure 5) suggest that the dynamical property of tWH is perhaps a shared feature among nuclear RNA polymerases.

In conclusion, our study has revealed the dynamic positioning of Rpc34 WH2 in Pol III EC, providing the first observation of Pol III's structural dynamics in real-time. Given that each identified position may correspond to a distinct function, our results confer a structural framework that underpins the multi-functionality of Rpc34 tWH. Importantly, the chemical biology protocol we developed here has overcome a previously overlooked issue on the usage of Bertozzi reaction for selective bio-orthogonal labelling, and thus render smFRET amenable to large native protein assemblies (for instance, the Pol III assembly comprises ~6,000 amino acids with 1% to be cysteine residues).

REFERENCES

1. Dieci,G., Fiorino,G., Castelnuovo,M., Teichmann,M. and Pagano,A. (2007) The expanding RNA polymerase III transcriptome. Trends Genet., 23, 614-622.
2. Kessler,A.C. and Maraia,R.J. (2021) The nuclear and cytoplasmic activities of RNA polymerase III, and an evolving transcriptome for surveillance. Nucleic Acids Res., 49, 12017–12034.
3. Vannini,A. and Cramer,P. (2012) Conservation between the RNA polymerase I, II, and III transcription initiation machineries. Mol. Cell, 45, 439-446.
4. Carter,R. and Drouin,G. (2010) The increase in the number of subunits in eukaryotic RNA polymerase III relative to RNA polymerase II is due to the permanent recruitment of general transcription factors. Mol. Biol. Evol., 27, 1035-1043.
5. Brun,I., Sentenac,A. and Werner,M. (1997) Dual role of the C34 subunit of RNA polymerase III in transcription initiation. EMBO J., 16, 5730-5741.
6. Andrau,J.C., Sentenac,A. and Werner,M. (1999) Mutagenesis of yeast TFIIIB70 reveals C-terminal residues critical for interaction with TBP and C34. J. Mol. Biol., 288, 511-520.
7. Plaschka,C., Hantsche,M., Dienemann,C., Burzinski,C., Plitzko,J. and Cramer,P. (2016) Transcription initiation complex structures elucidate DNA opening. Nature, 533, 353-358.
8. Wei,YY and Chen,HT (2018) Functions of the TFIIE-related tandem winged-helix domain of Rpc34 in RNA polymerase III initiation and elongation. Mol. Cell. Biol., 38, e00105-17.
9. Wu,CC, Herzog,F., Jennebach,S., Lin,YC, Pai,CY, Aebersold,R., Cramer,P. and Chen,HT (2012) RNA polymerase III subunit architecture and implications for open promoter complex formation. Proc. Natl. Acad. Sci. USA 109, 19232–19237.
10. Vorländer,M.K., Khatter,H., Wetzel,R., Hagen,W.J.H. and Müller,C.W. (2018) Molecular mechanism of promoter opening by RNA polymerase III. Nature, 553, 295-300.
11. Abascal-Palacios,G., Ramsay,E.P., Beuron,F., Morris,E. and Vannini,A. (2018) Structural basis of RNA polymerase III transcription initiation. Nature, 553, 301-306.
12. Vorländer, MK, Baudin, F, Moir, RD, Wetzel, R, Hagen, WJH, Willis, IM, Müller, C.W. (2020). Structural basis for RNA polymerase III transcription repression by Maf1. Nat Struct Mol Biol., 27,: 229-232.
13. Hoffmann,N.A., Jakobi,A.J., Moreno-Morcillo,M., Glatt,S., Kosinski,J., Hagen,W.J.H., Sachse,C. and Müller,C.W. (2015) Molecular structures of unbound and transcribing RNA polymerase III. Nature, 528, 231-236.
14. Girbig,M., Misiaszek,A.D., Vorländer,M.K., Lafita,A., Grötsch,H., Baudin,F., Bateman,A. and Müller, C.W. (2021) Cryo-EM structures of human RNA polymerase III in its unbound and transcribing states. Nat. Struct. Mol. Biol., 28, 210-219.
15. Kim,H. and Ha,T. (2013) Single-molecule nanometry for biological physics. Rep. Prog. Phys., 76, 016601.
16. Banerjee,P.R. and Deniz,A.A. (2014) Shedding light on protein folding landscapes by single-molecule fluorescence. Chem. Soc. Rev., 43, 1172–1188.

17. Schuler,B. and Eaton,W.A. (2008) Protein folding studied by single-molecule FRET. Curr. Opin. Struct. Biol. 18, 16–26.
18. Chin,J.W., Cropp,T.A., Anderson,J.C., Mukherji,M., Zhang,Z. and Schultz,P.G. (2003) An expanded eukaryotic genetic code. Science, 301, 964-967.
19. Chatterjee,A., Sun,S.B., Furman,J.L., Xiao,H. and Schultz,P.G. (2013) A versatile platform for single- and multiple-unnatural amino acid mutagenesis in Escherichia coli. Biochemistry, 52, 1828-1837.
20. Chen,S., Schultz,P.G. and Brock, A. (2007) An improved system for the generation and analysis of mutant proteins containing unnatural amino acids in Saccharomyces cerevisiae. J. Mol. Biol., 371, 112-122.
21. Chen,HT, Warfield,L. and Hahn,S. (2007) The positions of TFIIF and TFIIE in the RNA polymerase II transcription preinitiation complex. Nat. Struct. Mol. Biol., 14, 696-703.
22. Wu,CC, Lin,YC and Chen,HT (2011) The TFIIF-like Rpc37/53 dimer lies at the center of a protein network to connect TFIIIC, Bdp1, and the RNA Polymerase III active center. Mol. Cell. Biol., 31, 2715-2728.
23. Brustad,E.M., Lemke,E.A., Schultz,P.G. and Deniz,A.A. (2008) A general and efficient method for the site-specific dual-labeling of proteins for single molecule fluorescence resonance energy transfer. J. Am. Chem. Soc., 130, 17664–17665.
24. Kim,J., Seo,M.H., Lee,S., Cho,K., Yang,A., Woo,K., Kim,H.S. and Park, H.S. (2013) Simple and efficient strategy for site-specific dual labeling of proteins for single-molecule fluorescence resonance energy transfer analysis. Anal. Chem., 85, 1468–1474.
25. Lee,TC, Kang,M., Kim,C.H., Schultz,P.G., Chapman,E. and Deniz,A.A. (2016) Dual unnatural amino acid incorporation and click-chemistry labeling to enable single-molecule FRET studies of p97 folding. ChemBioChem, 17, 981-984.
26. Agard,N.J., Prescher,J.A. and Bertozzi,C.R. (2004) A strain-promoted [3+2] Azide-Alkyne cycloaddition for covalent modification of biomolecules in living systems. J. Am. Chem. Soc., 126, 15046-15047.
27. Tian,H., Sakmar,T.P. and Huber, T. (2016) A simple method for enhancing the bioorthogonality of cyclooctyne reagent. Chem. Commun., 52, 5451-5454.
28. van Geel,R., Pruijn,G.J.M., van Delft,F.L. and Boelens,W.C. (2012) Preventing Thiol-Yne addition improves the specificity of stain-promoted azide-alkyne cycloaddition. Bioconjugate Chem., 23, 392-398.
29. Karala,A.R. and Ruddock,L.W. (2007) Does S-methyl methanethiosulfonate trap the thiol–disulfide state of proteins? Antioxid. Redox Signal, 9, 527-531.
30. Kalinin, S., Peulen, T., Sindbert, S., Rothwell, P. J., Berger, S., Restle, T., Goody, R. S., Gohlke, H., Seidel, C. A. M. (2012) A toolkit and benchmark study for FRET-restrained high-precision structural modeling. Nat. Methods 9, 1218-1225.
31. Sass,L.E., Lanyi,C., Weninger,K. and Erie,D.A. (2010) Single-molecule FRET TACKLE reveals highly dynamic mismatched DNA-MutS complexes. Biochemistry, 49, 3174–3190.

32. Treutlein,B., Muschielok,A., Andrecka,J., Jawhari,A., Buchen,C., Kostrewa,D., Hög,F., Cramer,P. and Michaelis,J. (2012) Dynamic architecture of a minimal RNA polymerase II open promoter complex. Mol. Cell, 46, 136-146.
33. McCann,J.J., Choi,U.B., Zheng,L., Weninger,K. and Bowen,M.E. (2010) Optimizing methods to recover absolute FRET efficiency from immobilized single molecules. Biophys. J., 99, 961-970.
34. Taylor JN, Makarov DE, Landes CF. Denoising single-molecule FRET trajectories with wavelets and Bayesian inference (2010) Biophys J. 98:164-73.
35. McKinney,S.A., Joo,C. and Ha,T. (2006) Analysis of single-molecule FRET trajectories using hidden Markov modeling. Biophys J., 91, 1941-1951.
36. Muschielok,A., Andrecka,J., Jawhari,A., Bruckner,F., Cramer,P. and Michaelis,J. (2008) A nano-positioning system for macromolecular structural analysis. Nat. Methods, 5, 965-971.
37. Ferber,M., Kosinski,J., Ori,A., Rashid,U.J., Moreno-Morcillo,M., Simon,B., Bouvier,G., Batista,P.R., Müller,C.W., Beck,M. and Nilges,M. (2016) Automated structure modeling of large protein assemblies using crosslinks as distance restraints. Nat. Methods, 13, 515-520.
38. Zhang,S., Zou,S., Yin,D., Zhao,L., Finley,D., Wu,Z. and Mao,Y. (2022) USP14-regulated allostery of the human proteasome by time-resolved cryo-EM. Nature, 605, 567–574.
39. Khatter,H., Vorlander,M.K. and Müller,C.W. (2017) RNA polymerase I and III: similar yet unique. Curr. Opin. Struct. Biol., 47, 88-94.
40. Engel,C., Neyer,S. and Cramer,P. (2018) Distinct mechanisms of transcription initiation by RNA polymerases I and II. Annu. Rev. Biophys., 47, 425-446.
41. Sadian,Y., Baudin,F., Tafur,L., Murciano,B., Wetzel,R., Weis,F. and Müller,C.W. (2019) Molecular insight into RNA polymerase I promoter recognition and promoter melting. Nat. Commun., 10, 5543.

MATERIAL AND METHODS

Plasmids, yeast strains and cell growth

Incorporation of UAA into yeast was achieved as described previously (20). In brief, to generate a yeast strain in which UAA had been incorporated into Rpc34, the coding gene sequence of Rpc34 with a V5 epitope tag at the N terminus was cloned into the 2-micron vector pRS425 (LEU2+), as shown in Figure 1A. The TAG (amber) nonsense codon was introduced by in vitro mutagenesis into Rpc34 of the Rpc34/pRS425 plasmid to express a series of Rpc34 mutants in which UAA was incorporated at specific single positions. To incorporate an azido-UAA (4-Azido-L-phenylalanine, AzF), the plasmid pLH157 (tRNACUA/BPA-tRNA synthetase) was modified to become pLH157-AzF (tRNACUA/AzF-tRNA synthetase plasmid), in which the EcTyrRS gene was mutated to host amino acid changes Tyr37Leu, Asp182Ser, Phe183Met, and Leu186Ala according to the original gene mutations as previously described (18). The Rpc34 mutant/pRS425 plasmid and pLH157-AzF were co-transfected into the Saccharomyces cerevisiae YLy3 yeast strain [MATα ade2::his3G his3Δ200 leu2Δ met15Δ lys2Δ trp1Δ63 ura3Δ (rpc34::KanMX4) Rpc34/pRS316 (Ura3+)] modified from the BY4705 strain. In this modified yeast strain, the chromosomal Rpc34 gene is disrupted by a KanMx gene cassette resistant to the drug G418. Additionally, the C terminus of the Rpc2 (Rpc128) protein hosts a TAP tag (FLAG3-His6-TEV cleavage sequence-protein A).

Saccharomyces cerevisiae Maf1 gene was condon optimized and synthesized (GenScript Biotech.), and then subcloned into pDuet2 vector with the BamHI and XhoI restriction sites to produce the pDuet2-Maf1 plasmid. In this plasmid, the N-terminus of Maf1 had His6-HA-SUMO tag and that at C-terminus was a TAP tag (FLAG-Twin Strep Tag-HRV3C cleavage site).

Protein purification

The Pol III was purified as described previously (8). In brief, yeast cell cultures were grown in YPD medium with 0.2 mM AzF to OD= 2 and then harvested. Cells from 3-liter cell cultures were pooled and resuspended in the TAP-tag purification buffer containing 40 mM HEPES (pH 7.5), 350 mM NaCl, 10% glycerol, 0.1% Tween-20, 0.5 mM ethylenediaminetetraacetic acid (EDTA) and 1x protease inhibitors (1 μM pepstatin A, 1 mM phenylmethylsulfonyl fluoride (PMSF), 2.58 mM benzamidine and 0.7 mM leupeptin) and then lysed by glass bead beating. The cell lysate was centrifuged at 10,000 rpm for 10 min and ultra-centrifuged at 35,000 rpm for 1 hour to collect the clarified lysate. The resulting supernatant was incubated overnight with 2 ml IgG-Sepharose beads (GE Healthcare) at 4ºC. Subsequently, the bound proteins on the beads were washed with purification buffer and then resuspended in 1 mL TEV cleavage buffer containing 10 mM Tris (pH 8.0), 150 mM NaCl, 10% glycerol, 0.1% NP-40, 0.5 mM EDTA, and 50 μg TEV protease for reaction at 4 ºC for 2 hours. The eluted protein was verified by Coomassie-blue stained SDS-PAGE and Western blotting for the FLAG and V5 tags (Supplementary Figure S8).

For overexpression of recombinant *S. cerevisiae* Maf1, the pDuet2-Maf1 palsmid was transformed into Rosetta 2(DE3) competent cells (Merck). Cells grew in 2 liter LB medium containing 30 mg/ml of chloramphenicol and 100 mg/ml Ampicillin at 37ºC to OD_{600} = 0.5 and then protein expression was induced with 0.1 mM IPTG overnight at 18ºC. Cells were harvested by centrifugation and

resuspended in the lysis buffer (25 mM Tris:HCl pH8.0, 1x protease inhibitor (Roche) and 5 mM beta-mercaptoethanol), and then lysed by the constant cell disruption system (TS 2.2Kw, Constant Systems). The cell lysate was ultra-centrifuged at 30,000 g for 30 minutes and then collected the supernatant to incubated with SUMO protease (ThermoFisher) overnight at 4ºC to remove N-terminal tag. After SUMO cleavage, the lysate was incubated with the Strep-Tactin resin (Cytiva) for 2 hours at 4ºC and the beads were washed by 100 column volume of the high-salt washing buffer (25 mM Tris:HCl pH8.0, 500 mM NaCl, 1% NP-40 and 5 mM beta-mercaptoethanol) and equilibrated with the HRV3C cleavage buffer (25mM Tris:HCl pH8.0, 150 mM NaCl, 1 mM DTT and 5% glycerol). The beads bound with Maf1 were incubated with human rhinovirus 3C (HRV3C) protease overnight at 4ºC and then the Maf1 protein was eluted. In final, the eluted protein sample was incubated with NTA beads to remove HRV3C protease. The eluted Maf1 protein was verified by Coomassie-blue stained SDS-PAGE and by mass spectrometry.

Pol III TEC formation

Purified Pol III was incubated with a DNA/RNA oligonucleotide scaffold—composed of TAMRA dye-labeled template DNA (T-DNA), biotinylated non-template DNA (NT-DNA) and 10-nucleotide (nt) RNA—to form the Pol III TEC complex in buffer (100 mM HEPES pH 7.9, 400 mM KCl, 10% glycerol, 25 mM $MgCl_2$, 5 mM EDTA). The oligonucleotide sequences and modifications are: (i) template strand with TAMRA dye labeling at the +7 position (5'-CATAAAAAACCCAAAAAAAGAGAG TATT-TAMRA dye-AATTGTTGAAGAAAGAGTATACTACATA); (ii) non-template strand (5'-biotin-TATGTAGATATGAGAAAGAAGTACAATTAAATACTCTCTTTTTTTGGGTTTTTTATG); (iii) RNA strand (5'-UCUUUCUUCA-3'). The three oligonucleotide strands were mixed and annealed to form the DNA/RNA scaffold with a 15-nucleotide bubble. After Pol III TEC formation, free sulfhydryls of cysteines in the Pol III TEC were capped with 1 mM iodoacetamide (IAM), 20 μM N-ethylmaleimide (NEM), or 20 μM S-methyl methanethiosulfonate (MMTS) to suppress non-specific dye labeling. Subsequently, 10 μM Alexa647-DIBO (C10408, Thermo-Fisher) was added into 3 μg of Pol III TEC complex to specifically label AzF by copper-free click chemistry reaction at 4 ºC overnight. Following the reaction, a desalting spin column (Roche) was used to remove excessive dye-DIBO reagent. TEC formation was verified by fluorescence and Coomassie-blue-stained imaging of native PAGE (Thermo Fisher), as shown in Supplementary Figs. 2 and 3.

SmFRET system equipped with alternating laser excitation

Our smFRET experiments and analytical methodology were conducted as described previously (42). In brief, approximately 100 pM Pol III TEC complex labeled with the donor-acceptor dye pair was immobilized on a clean cover-glass that was surface-coated with Neutravidin (Thermo Fisher) and biotin-PEG (Laysan Bio) in a sample chamber (42,43). The oxygen scavenger imaging buffer (175 nM protocatechuate-3,4-dioxygenase, 7.8 mM protocatechuic acid, 2 mM Trolox, 50 mM Tris:HCl pH7.8, 150 mM NaCl, 1 mM PMSF and 0.1 mg/ml BSA) was injected into the chamber to increase fluorophore photo-stability. Wide-field fluorescence images were acquired using a custom-built dual-view total internal reflection fluorescence (TIRF) microscope in which a 532 nm-wavelength laser

(DPGL-2100, Photop Technologies) and a 638 nm-wavelength laser (LuxX 638–100, Omicron) were used to excite the donor and acceptor dyes; an oil-immersion objective (UPLSAPO100XO, Olympus, NA 1.49) was used to focus the laser beam onto the sample and collect fluorescence. The dichromatic mirror (FF640-FDi01, Semrock) separated the collected fluorescence photons into the donor and the acceptor emission arms, with the two different color images projected onto different areas of the EMCCD camera (DV887DCS-BV, Andor Technology). The movies were captured with a frame rate of 5fps, namely 200 milliseconds per frame. In addition, we implemented an alternating laser excitation (ALEX) (44) scheme using a shutter system (LS3ZM2, Uniblitz Electronic) and relay circuit for switching between the excitation of the donor and that of the acceptor as described (45). ALEX imaging filters out molecules with the donor and acceptor both present, and also time periods when the donor or acceptor is absent due to switching to unwanted photo-physical states.

Simulation of dye accessible volume and the predicted FRET efficiency

In order to compare different structural models of Rpc34 WH2 and estimate FRET efficiencies of FRET pairs in which the acceptor was labeled at Rpc34 WH2, the Rpc34 tWH domain of different models was docked into Pol III TEC (PDB: 5fj8) based on alignment of Rpc1 between models (Supplementary Fig. 1). Accessible volumes of dyes were simulated by FPS software (30) based on these Pol III TEC docked with Rpc34 tWH and dye parameters (size of dye and the length and width of linker), and then the distances R_{MP} between the mean dye positions of two accessible volumes of FRET pair in different models were obtained. We calculated the predicted FRET efficiencies of different models based on the Forster radius of dye pair.

SmFRET data analysis and dwell-time analysis

We used a custom-built two-channel TIRF system to record fluorescence images of the donor and acceptor (42). Fluorescence images were analyzed to identify the co-localized donor/acceptor pair, before extracting the donor and acceptor time trajectories using algorithms and programs as described previously (46). The extracted signal intensities of the donor and acceptor were used to calculate FRET efficiencies according to the following equation:

$$E = \frac{I_A - \beta I_D}{\left(\left(I_A - \beta I_D\right) + \gamma I_D\right)}$$

according to the method of individual γ normalization (33). I_A and I_D represent the measured intensities of the acceptor and donor, respectively; βI_D corrects for leakage of donor emission into the acceptor channel, and γ corrects for the difference in detection efficiency and quantum yield between the donor and acceptor. We adopted acceptor-photo-bleaching FRET (33) to validate each candidate FRET trace and also to obtain the absolute FRET efficiencies. This method assesses a time trace based on anti-correlated donor/acceptor signals at the time-point when the acceptor is photo-bleached. Using the stringent criterion of acceptor-photo-bleaching, typically we only identifed a few “authentic” smFRET traces from hundreds of candidate donor/acceptor co-localization pairs extracted from tens of movies. Moreover, considering that the acceptor may not always be active in the time trace of a single molecule, we excited both channels using ALEX (44) to continuously survey the

presence of an acceptor so as to detect the period when it is not in the active state, namely in the dark state, prior to photo-bleaching. Using this approach, we filtered out 5 traces from the initially selected 14 traces for the Rpc34 WH2 (K126) and DNA (+7 site) pair that exhibited apparent transitions, as those traces represented transitions due to acceptor quenching rather than FRET as exemplified in Supplementary Figure S5. After verification of time traces, we used the frame-wise FRET efficiency to build a histogram and fitted it using multiple Gaussian distributions. Based on the fitting result (the number of states and the center of populations), we classified the stationary time traces into different states. For state-change traces, the period of different states were identified by hidden Markov modeling (HMM) (35) and the frame data were classified into different states. After traces classification, FRET efficiency histograms of different states were built up and the FRET efficiency histogram of each state was fitted with single Gaussian distribution. Furthermore, the duration of each state determined by HMM was used for the dwell-time histogram analysis fitted with exponential decay curve. The transitions between two states in the FRET time traces were used to build up the transition density plot. In addition, the representative traces were smoothed by the denoising algorithm with wavelets and Bayesian inference (34).

Nano-positioning system analysis

Nano-positioning system (NPS) analysis (36) was performed by using the Fast-NPS method recently developed by the Michaelis group (47). The credible volume accessible by the antenna dye (acceptor) was calculated using the Fast-NPS software (47), with the parameters of the mean and standard deviation of FRET efficiencies derived from the smFRET histogram, the coordinates of the satellite dye (donor) attached to the protein, the Förster radius of the dye pair, fluorescence anisotropy of both dyes, and the properties of the dye and linker (dye diameter: 13 Å; linker length: 13 Å; linker diameter: 4.5 Å, as suggested by Ref. 47). The coordinates of the satellite dye attachment point were based on the atomic model of Pol III TEC (PDB: 5fj8). The Förster radius of the dye pair is 6.0 nm (32). The anisotropy of the donor dye on DNA was measured as ~0.3 using a lifetime technique by means of a time-resolved laser scanning confocal microscope (Q2, ISS, USA). The measured value is in agreement with those of previous studies. For the anisotropy of the acceptor dye on UAA in Pol III, we adopted the value of ~0.3 reported by the Deniz group (25).

42. Wu,JS, Chen,TY, Lin,S.SY, Lin,SY, Hung,CY, Tu,IP, Chen,HT and Chang,WH (2019) Deriving a sub-nanomolar affinity peptide from TAP to enable smFRET analysis of RNA polymerase II complexes. Methods, 159-160, 59-69.
43. Roy,R., Hohng,S. and Ha,T. (2008) A practical guide to single-molecule FRET, Nat. Methods, 5, 507-516.
44. Kapanidis,A.N., Lee,N.K., Laurence,T.A., Doose,S., Margeat,E. and Weiss,S. (2004) Fluorescence-aided molecule sorting: Analysis of structure and interactions by alternating-laser excitation of single molecules. Proc. Natl. Acad. Sci. USA, 101, 8936-8941.
45. Lee,S., Lee,J. and Hohng,S. (2010) Single-molecule three-color FRET with both negligible spectral overlap and long observation time. PLoS One, 5, e12270.

46. Lee,HC, Lin,BL, Chang,WH and Tu,IP (2012) Toward automated denoising of single molecular Förster resonance energy transfer data. J. Biomed. Opt., 17, 011007.
47. Nagy,J., Eilert,T. and Michaelis,J. (2018) Precision and accuracy in smFRET based structural studies — A benchmark study of the Fast-Nano-Positioning System. J. Chem. Phys., 148, 123308.

DATA AVAILABILITY

The traces data in this study are avalible from the corresponding authors upon resonable request.

SUPPLEMENTARY DATA

Supplementary Data are available in a separate file.

ACKNOWLEDGEMENTS

W.H.C. is grateful to Dr. Jiun-Jie Shie of Institute of Chemistry for commenting on the issues of Bertozzi reaction on bio-orthogonal applications and suggesting the usage of thio-blocking agents. The authors thank John O'Brien for English editing.

FUNDING

This work is supported by Academia Sinica [AS-TP-110-L11], and Ministry of Science and Technology (MOST) of Taiwan [105-2627-M-001-010, 106-2627-M-001-009 and 107-2627-M-001-005] to W.H.C. and H.T.C. J.S.W. has been supported by an Academia Sinica Postdoc Fellowship.

CONFLICT OF INTEREST

The authors declare there is no conflict of interest.