

# Time-Resolved Raman and Polyacrylamide Gel Electrophoresis Observations of Nucleotide Incorporation and Misincorporation in RNA within a Bacterial RNA Polymerase Crystal

Ioanna H. Antonopoulos,<sup>†</sup> Yuko Murayama,<sup>‡,§</sup> Brittany A. Warner,<sup>||,¶</sup> Shun-ichi Sekine,<sup>‡,§</sup> Shigeyuki Yokoyama,<sup>‡,¶</sup> and Paul R. Carey<sup>\*,†</sup>

<sup>†</sup>Department of Biochemistry, Case Western Reserve University School of Medicine, Cleveland, Ohio 44106, United States

<sup>‡</sup>RIKEN Systems and Structural Biology Center, Yokohama, Japan

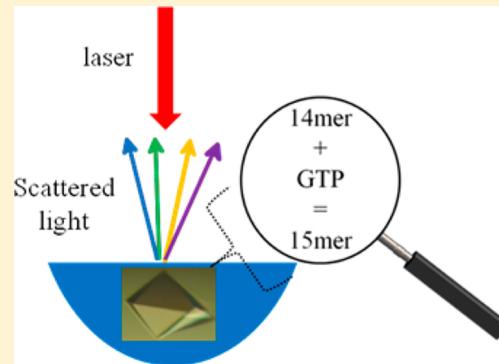
<sup>§</sup>Division of Structural and Synthetic Biology, RIKEN Center for Life Science Technologies, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama 230-0045, Japan

<sup>||</sup>Department of Biochemistry and Molecular Biology, The Center of RNA Molecular Biology, Pennsylvania State University, University Park, Pennsylvania 16802, United States

<sup>¶</sup>RIKEN Structural Biology Laboratory, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama 230-0045, Japan

## Supporting Information

**ABSTRACT:** The bacterial RNA polymerase (RNAP) elongation complex (EC) is highly stable and is able to extend an RNA chain for thousands of nucleotides. Understanding the processive mechanism of nucleotide addition requires detailed structural and temporal data for the EC reaction. Here, a time-resolved Raman spectroscopic analysis is combined with polyacrylamide gel electrophoresis (PAGE) to monitor nucleotide addition in single crystals of the *Thermus thermophilus* EC (TthEC) RNAP. When the cognate base GTP, labeled with <sup>13</sup>C and <sup>15</sup>N (\*GTP), is soaked into crystals of the TthEC, changes in the Raman spectra show evidence of nucleotide incorporation and product formation. The major change is the reduction of \*GTP's triphosphate intensity. Nucleotide incorporation is confirmed by PAGE assays. Both Raman and PAGE methods have a time resolution of minutes. There is also Raman spectroscopic evidence of a second population of \*GTP in the crystal that does not become covalently linked to the nascent RNA chain. When this population is removed by "soaking out" (placing the crystal in a solution that contains no NTP), there are no perturbations to the Raman difference spectra, indicating that conformational changes are not detected in the EC. In contrast, the misincorporation of the noncognate base, <sup>13</sup>C- and <sup>15</sup>N-labeled UTP (\*UTP), gives rise to large spectroscopic changes. As in the GTP experiment, reduction of the triphosphate relative intensity in the Raman soak-in data shows that the incorporation reaction occurs during the first few minutes of our instrumental dead time. This is also confirmed by PAGE analysis. Whereas PAGE data show \*GTP converts 100% of the nascent RNA 14mer to 15mer, the noncognate \*UTP converts only ~50%. During \*UTP soak-in, there is a slow, reversible formation of an  $\alpha$ -helical amide I band in the Raman difference spectra peaking at 40 min. Similar to \*GTP soak-in, \*UTP soak-in shows Raman spectroscopic evidence of a second noncovalently bound \*UTP population in the crystal. Moreover, the second population has a marked effect on the complex's conformational states because removing it by "soaking-out" unreacted \*UTP causes large changes in protein and nucleic acid Raman marker bands in the time range of 10–100 min. The conformational changes observed for noncognate \*UTP may indicate that the enzyme is preparing for proofreading to excise the misincorporated base. This idea is supported by the PAGE results for \*UTP soak-out that show endonuclease activity is occurring.



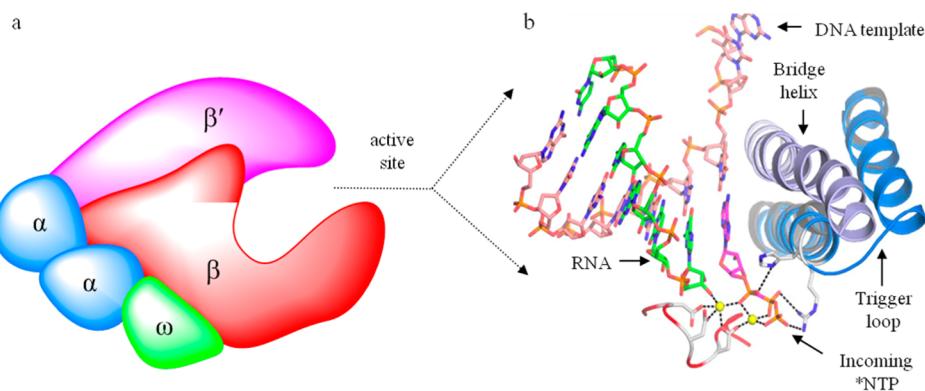
RNA polymerases (RNAPs) are essential enzymes for gene expression in all domains of life, including archaea, eukaryotes, bacteria, and viruses. In the RNAP elongation complex (EC), an incoming ribonucleoside triphosphate (NTP) enters the RNAP active site to form a Watson–Crick base pair with template DNA, and the nucleoside monophosphate portion is covalently linked to the 3' end of the nascent RNA chain. This process is successively repeated as RNAP translocates along the template DNA, thus transcribing the DNA sequence into a RNA

transcript. Pioneering X-ray crystallographic structures of the RNAP complex at different stages on the transcription reaction pathway<sup>1–8</sup> have paved the way for further kinetic, biochemical, biophysical, single-molecule, and structural investigations<sup>9–20</sup>

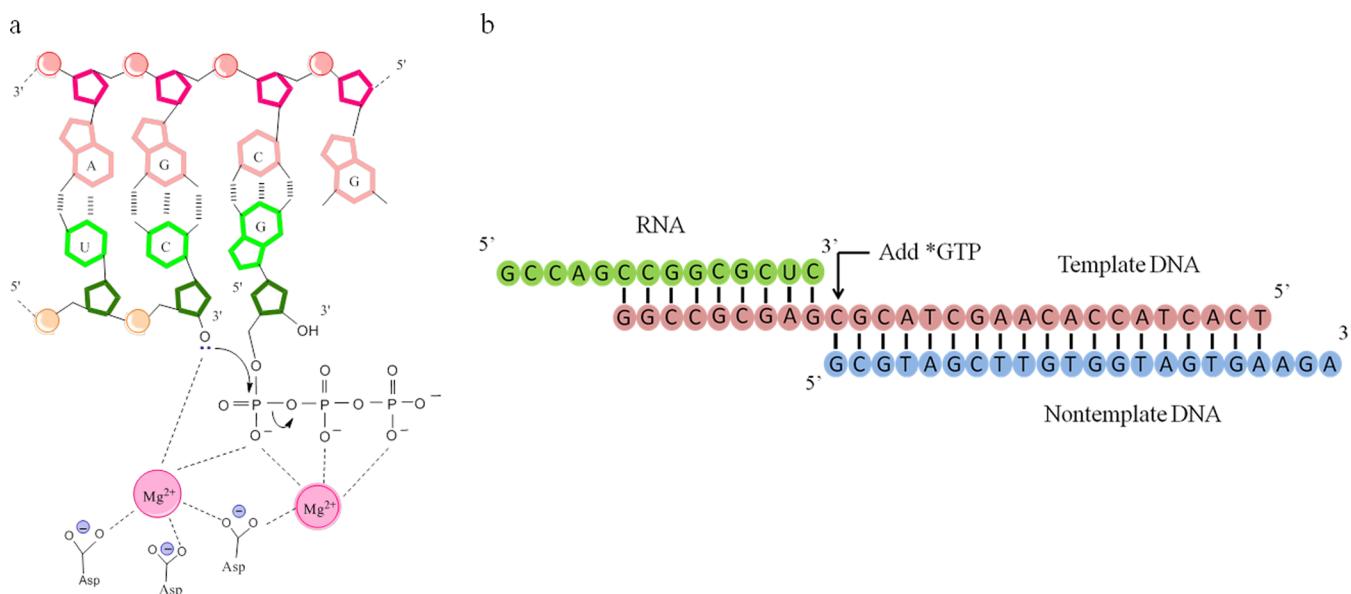
**Received:** September 16, 2014

**Revised:** December 18, 2014

**Published:** December 24, 2014



**Figure 1.** Tth core RNA polymerase active site. (a) Bacterial RNAP is composed of five subunits,  $\alpha_2\beta\beta'\omega$ , where  $\beta$  and  $\beta'$  make up the pincers of the crab claw structure and contain catalytically significant structural elements, including trigger loop and bridge helix. (b) The Tth active site is composed of an 8–9 bp DNA–RNA hybrid upstream from the NTP insertion site, while duplex DNA (not shown) is located downstream of the transcription bubble. Here, only template DNA of the DNA duplex is shown downstream (pink). Incoming NTPs make important contacts with the trigger loop motif (blue) as they are incorporated into the elongating RNA chain (green). Tth RNAP's aspartic acid residues D739, D741, and D743 coordinate with two magnesium metals I and II (yellow spheres) to facilitate catalysis.



**Figure 2.** Chemistry of NTP incorporation. (a) The NTP insertion site is proximal to the nascent cytidine on the RNA chain, whose 3'-OH performs a nucleophilic attack on the  $\alpha$ -phosphate group of an incoming  $^{13}\text{C}$ - and  $^{15}\text{N}$ -labeled GTP, resulting in phosphodiester bond formation and release of pyrophosphate. (b) Nucleic acid sequence used for the Tth RNAP elongation complex (TthEC) crystal. Upstream from the NTP insertion site are a 9 bp RNA–DNA hybrid (green and pink) and a 5-nucleotide single-stranded RNA tail (including a 14mer RNA), while downstream from the NTP insertion site are a 19-nucleotide DNA duplex (template and nontemplate DNA colored pink and blue, respectively) and a 3-nucleotide single-stranded DNA tail.

that provide support for several translocation theories.<sup>16,21–23</sup> However, because of a lack of temporal resolution in X-ray crystallographic studies, controversy yet exists over the choreography of active site events, including the nucleotidyl transfer reaction and ensuing translocation. This study offers the opportunity to begin to address these issues with a time-resolved spectroscopic approach. A bacterial RNAP model, *Thermus thermophilus* (Tth) is employed, which is advantageous for biophysical analysis; it represents the simplest form of a cellular polymerase, therefore reducing the complexity of studying multisubunit polymerases. For example, the bacterial transcription machinery proceeds with the addition of a limited number of transcription factors in contrast to the several eukaryotic factors. This allows for a more facile study of transcription initiation as well as elongation in bacterial systems.

Tth RNAP has a functional 380 kDa core containing five subunits ( $\alpha_2\beta\beta'\omega$ ), whose architecture resembles a crab claw (Figure 1a).<sup>1</sup> The two largest subunits ( $\beta$  and  $\beta'$ ) make up the enzyme's pincers and form a deep 27 Å wide channel. It is in the  $\beta'$  subunit that two magnesium ions occupy the active site and serve a critical mechanistic function while coordinated with three aspartic acid residues (D739, D741, and D743) (Figure 2a). In the EC, the downstream DNA enters the main channel and makes extensive contacts with amino acid residues in the pincers. During transcription, the DNA duplex is unwound by approximately 12–14 bp to form a transcription bubble. As RNAP moves along the DNA template in the 3' to 5' direction, the downstream DNA is unwound ahead of the polymerase while the DNA behind the polymerase reanneals upstream. As the template strand is directed into the active site at a 90° angle

(Figure 1b) and maintains an 8–9 bp hybrid with the elongated RNA upstream (Figures 1b and 2b), the nontemplate strand exits to the outer surface of the enzyme.<sup>16</sup> Key structural motifs, such as the trigger loop and bridge helix, also play an important role in the mechanism as they change conformations during the nucleotide addition cycle (NAC). Vassylyev and colleagues postulated a nucleotide addition cycle that involves the conformational change of a trigger loop into a trigger helix as well as reversible changes in portions of Tth RNAP's pincers (Figure 1).<sup>24</sup> In a series of recent papers, Tagami et al. also report that Tth RNAP exhibits a large conformational change when bound to a transcription inhibitory factor, Gfh1.<sup>22,25,26</sup>

Here, a spectroscopic technique, Raman microscopy, is employed to illuminate the dynamics of the nucleotidyl transfer reaction and provide a time sequence over which structural changes during RNA synthesis take place in a TthEC crystal. The feasibility of this approach has been demonstrated for the transcription initiation reaction catalyzed by the enterobacteriophage (N4) RNAP.<sup>27</sup> In this work, we employ two methods to follow the time-dependent incorporation of cognate GTP into a RNA chain in crystals of RNAP from *T. thermophilus*. This represents a relatively large, approximately 400 kDa “molecular machine”. One method employs Raman microscopy and follows the kinetics of GTP incorporation via the loss of the triphosphate Raman signature as well as other structural changes related to the complex. The second method, polyacrylamide gel electrophoresis (PAGE), measures changes in the length of the RNA oligomer with time. To initiate the reaction for both experimental approaches, GTP, labeled with <sup>13</sup>C and <sup>15</sup>N (Raman) or unlabeled (PAGE analysis), was soaked into single crystals of RNAP, where the complex is in the post-translocated state. The results from both methods are in general accord; however, the Raman experiments also reveal time-dependent changes in protein and nucleic acid conformations as well as evidence of a possible secondary NTP binding site. Parallel experiments were conducted with the noncognate UTP (PAGE) and \*UTP (Raman). The Raman spectroscopic changes were especially pronounced when the noncognate UTP was misincorporated into the RNA chain.

## EXPERIMENTAL PROCEDURES

**Crystallization of Core Tth RNAP with Nucleic Acids To Form TthEC.** Purification methods of the Tth RNAP core enzyme are as described previously.<sup>4,5,28</sup> The nucleic acid sequences consisting of RNA (Thermo Scientific), template DNA (Integrated DNA Technologies), and nontemplate DNA (Integrated DNA Technologies) are shown in Figure 2b. Before crystal trays are set, nucleic acid components (25 μM each, purified via high-performance liquid chromatography) and RNAP (20 μM) are annealed to form the elongation complex (TthEC). The mixture is incubated in the thermocycler (Eppendorf Mastercycler EP Gradient Thermal Cycler) for 5 min at 70 °C, after which time the temperature is decreased to 20 °C at a rate of 2 °C/min for 30 min. A 1 μL aliquot of the latter elongation complex mixture is mixed with 1 μL of a reservoir solution on a siliconized glass cover slide (Hampton Research). The 2 μL hanging drop is then sealed over a well containing 1 mL of a reservoir solution within a 24-well crystallization tray. The reservoir solution consists of 4% polyethylene glycol 10000, 30% 1,6-hexanediol (Hampton Research), 50 mM Tris (pH 8.5), and 120 mM magnesium acetate (Sigma-Aldrich).<sup>29</sup> The crystal tray is incubated at 22 °C, and bipyramidal-shaped crystals of approximately 100 μm grow within 48 h. The concentration of

the Tth RNAP enzyme in the single crystal is 1.05 mM, while the concentration of the TthEC complex is 1.1 mM (see Scheme 1 of the Supporting Information for the calculation).<sup>27,30</sup>

**Basis of Raman Crystallography.** Raman microscopy employs a laser beam, which is focused inside a single crystal of the TthEC, contained in a crystallization tray on a microscope stage. The path of the beam begins with its emission from a krypton laser source. The laser is then threaded through a fiber-optic cable, after which it passes through a series of optical lenses, filters, and mirrors before being directed through the microscope objective and aimed at the crystal sample. From here, the laser photons interact with the crystal molecules, and backscattered photons are transmitted to a Kaiser Optical spectrograph with a charge-coupled detector (CCD) to provide a Raman spectrum from the focal volume inside the crystal. This spectrum provides unique information about the chemical environment, composition, and conformation of the molecules in the crystal during the reaction, where each peak corresponds to a vibrational mode of the photon-scattering molecules. These peaks contain information about protein and nucleic acid backbones and their side chains.<sup>31,32</sup>

**Increase in Temperature within the Laser Focal Volume in the Crystal.** This work is the first time that time-dependent Raman data for a reaction in a crystal have been compared to data from another technique, in this case PAGE. It has been recognized that the laser beam used to generate the Raman spectra results in crystal warming.<sup>31</sup> By placing a small thermocouple next to a crystal in a hanging drop, where the ambient temperature is 22 °C, we record a temperature of 32–37 °C when a crystal has a 120 mW, 647.1 nm laser beam focused within it. Raman data are collected from the focal volume of the beam inside the crystal, and it is likely that the temperature within the focal volume is higher. Thus, the measured temperature at the crystal–solution interface must be regarded as a lower limit. The challenging task of measuring the temperature in the waist of the focused laser beam within the crystal is a subject for future studies.

**Collection of Raman Data of [<sup>13</sup>C,<sup>15</sup>N]GTP and [<sup>13</sup>C,<sup>15</sup>N]UTP Soaking into the TthEC Crystal.** Using a Raman microscope, Raman spectra from a single crystal of the TthEC are collected. However, because the TthEC crystals are grown in 120 mM magnesium acetate, and high concentrations of metal inhibit RNAP's function, the TthEC crystal is stabilized in a hanging drop (5 μL) containing low-magnesium equilibrium buffer (EB) [4% PEG 10K, 30% 1,6-hexanediol, 50 mM Tris (pH 8.5), and 10 mM magnesium acetate] for 1 h before Raman data are collected to remove excess magnesium ions from the active site. The hanging drop is then placed on a siliconized cover slide, sealed in a well containing 1 mL of reservoir solution [3% PEG 10000, 30% 1,6-hexanediol, 50 mM Tris (pH 8.5), and 120 mM magnesium acetate], and held in a crystallization tray on the microscope stage. The 647.1 nm krypton laser beam, 100 mW, is focused into the TthEC crystal, and each Raman spectrum is collected for 300 s (5 s × 60 accumulations) using Holograms software (Kaiser Optical). In addition to the crystal spectrum, a spectrum of the surrounding equilibrium buffer (EB), in which the crystal is immersed, is also recorded. To calculate a pure TthEC spectrum, a spectral subtraction is performed with GRAMS/AI (Galactic Inc.), where TthEC is the minuend and EB is the subtrahend. A buffer peak at 1475 cm<sup>-1</sup> is used as an internal standard and reduced to zero to guide the latter subtraction. Subsequently, the TthEC crystal is transferred to a new hanging drop (5 μL) containing substrate (\*GTP or \*UTP)

soaking solution [4% PEG 10000, 30% 1,6-hexanediol, 50 mM Tris (pH 8.5), 10 mM magnesium acetate, and 2 mM \*GTP]. The crystal contains the TthEC consisting of a 9 bp DNA–RNA hybrid, a 19 bp DNA duplex, \*GTP (or \*UTP), and magnesium chloride. When the crystal is transferred from EB to a soaking solution, the time is recorded as 0 min. During the \*NTP soak, a Raman spectrum of the crystal undergoing RNA chain extension is collected every 5 min, recorded as TthEC crystal  $x$  min, where  $x$  minutes equals the time when spectrum collection ends relative to the 0 min time point. Also, a spectrum of the soaking buffer is collected every 20–30 min and recorded as TthEC soaking buffer  $t$  min, where  $t$  also equals the time when spectrum collection ends relative to the 0 min time point. To obtain the spectrum of TthEC at  $x$  min throughout the course of substrate (\*GTP or \*UTP) soak-in, a series of subtractions (TthEC crystal at  $x$  min minus TthEC soaking buffer at  $t$  min, where  $t$  is always chosen to be as close as possible to  $x$ ) were performed, again reducing the  $1475\text{ cm}^{-1}$  buffer peak to zero for the purpose of standardization (Figure S1 of the Supporting Information). Crucially, to reveal the state of the \*NTP-bound TthEC, including protein and DNA structural changes as a function of time during RNA synthesis, a series of secondary subtractions of TthEC at  $x$  min \*NTP soak-in minus “substrate-free” TthEC (Figure S2 of the Supporting Information) were calculated using RNAP’s phenylalanine ring feature at  $1002\text{ cm}^{-1}$  as an internal intensity standard. Phenylalanine is the chosen standard because its ring mode can be identified throughout experimentation, and its intensity remains essentially unchanged during the reaction.

**Raman Data Analysis.** After Raman data collection, spectral subtractions (as mentioned above) are conducted to perform optimal Raman data analysis. Careful attention must be paid to focus the laser beam into the crystal to minimize light scattering from surfaces, and considerable care must be exercised to obtain precise overlap between the substrate-free TthEC spectra and the time-dependent Raman spectra taken during \*NTP soak-in. Experiments were conducted on five crystals for \*GTP soaks and four crystals for \*UTP soaks; results are essentially identical within each set, demonstrating consistent Raman difference spectral results. The reproducibility from two different crystals during \*UTP soak-in at similar time points is illustrated in Figure S3 of the Supporting Information. The signal-to-noise ratio is 100/1 (Figure S1 of the Supporting Information) in the spectrum of the TthEC crystal when \*UTP has been soaked for 15 min (before secondary subtraction of substrate-free TthEC). Also, the signal-to-noise ratios are 7/1 and 10/1 in the time-dependent Raman difference spectra of \*GTP- and \*UTP-soaked TthEC crystals, respectively, in Figure S2 of the Supporting Information. The signal-to-noise ratios are also noted in Figures 3 and 6, while an example of a signal-to-noise measurement is provided in Figure S4 of the Supporting Information. The final Raman difference spectra have low signal-to-noise ratios because the relative Raman intensity changes are <3% of the major bands in Figure S1 of the Supporting Information and because of the low concentration ( $\sim 1\text{ mM}$ ) of the complex in the crystal. Therefore, there is a limit to the accuracy of the kinetic values derived from the standardized intensities of major marker bands. Nonetheless, the kinetic data presented in Figures 4a, 5a, 7a,d, 8b, and 9a–c semiquantitatively represent the trends in the various populations we address.

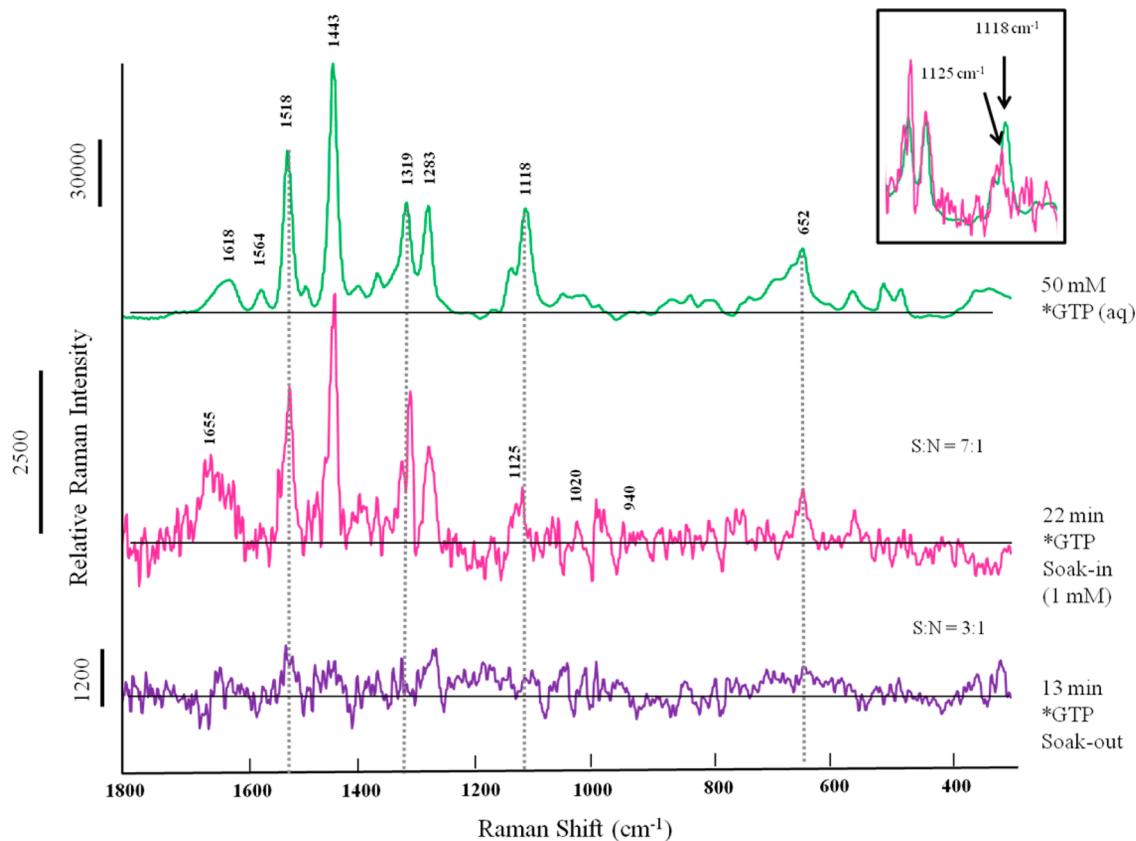
**Spectroscopic Assignments.** Considerable aid in identifying key Raman marker bands in the Raman difference spectra comes from the use of \*GTP, where guanosine is labeled with  $^{13}\text{C}$  and  $^{15}\text{N}$ . Thus, isotopically labeled substrate can be

distinguished from analogous bases in the nucleic acid scaffold crystal complex, which contains unlabeled guanosine, because the \*G ring modes are typically downshifted  $20\text{--}40\text{ cm}^{-1}$  in the Raman spectrum (Figure S5 of the Supporting Information). In addition, assignments are buttressed by an extensive literature,<sup>33–36</sup> as well as our earlier work on the *de novo* transcription initiation reaction of N4 virion RNAP.<sup>27</sup> Quantum mechanical calculations, using Gaussian software, are also used to refine and confirm assignments. Other important and useful tools in aiding spectroscopic assignments include X-ray crystallographic structures that aid in identifying Raman peaks from individual reaction intermediates.<sup>37</sup> Similarly, time-dependent Raman data from crystals can be informed by the results on the kinetic parameters describing the reaction in solution.<sup>10,21</sup> In all of these comparisons, we are sensitive to the possible limitations of comparing crystal and solution data.

**An Approximation of Kinetics by Raman Crystallography.** Raman spectra, recorded as a function of time, are used to track and measure structural changes related to the *in crystallo* transcription elongation reaction. Time course traces of these important reaction features, such as diffusion of the substrate into the crystal, conformational changes in the RNAP enzyme, and changes in nucleic acid structure are obtained by measuring the height of signature Raman features in the Raman difference spectra [TthEC + \*GTP at  $x$  min] minus [substrate-free TthEC]. For example, substrate \*GTP and \*UTP are primarily traced using their intense base ring modes at  $1518$  and  $1203\text{ cm}^{-1}$ , respectively. Importantly, their triphosphate mode near  $1125\text{ cm}^{-1}$  is tracked during the reaction. In addition, conformational changes in protein are tracked by the positive intensity around  $1655\text{ cm}^{-1}$  in the Raman difference spectrum, e.g., [TthEC + \*UTP at  $x$  min] – [substrate-free TthEC crystal] (Figures 6 and 8). The  $1655\text{ cm}^{-1}$  feature is assigned to intensity changes due to amide vibrations of the peptide bonds, with  $\alpha$ -helices making the major contribution.<sup>27,38,39</sup> The integration function of GRAMS/AI (Galactic) was used to measure the peak height intensities. This was accomplished by taking the  $y$  value at the base of the peak and subtracting it from the  $y$  value at the tip of the peak. The resultant difference yields the relative Raman intensity or number of photon events for each peak. For the purpose of standardization, each peak height is divided by the intensity of the  $\text{CH}_2$  stretching mode at  $1450\text{ cm}^{-1}$  in the mother spectrum of (TthEC at  $x$  minutes – buffer) (Figure S1 of the Supporting Information).

**Polyacrylamide Gel Electrophoresis.** These assays were initially performed at  $20\text{ }^\circ\text{C}$ ; however, to provide temperature conditions closer to those of the Raman experiments, where laser beam heating causes the crystal temperature to be in or above the range of  $32\text{--}37\text{ }^\circ\text{C}$ , PAGE assays for UTP were also performed at  $37\text{ }^\circ\text{C}$  (see below). The EC crystals were grown for 5–14 days under conditions that included 4–4.5% PEG 10000, 30% 1,6-hexanediol, 50 mM Tris-HCl (pH 8.5), and 120 mM magnesium acetate. To remove excess magnesium ions, the crystals were equilibrated with EB [4% PEG 10K, 30% 1,6-hexanediol, 50 mM Tris-HCl (pH 8.5), and 10 mM magnesium acetate] for 1 h at  $20\text{ }^\circ\text{C}$  before the soaking assays.

**GTP and UTP Soak-In, Soak-Out, and Pyrophosphorylation Assays at  $20\text{ }^\circ\text{C}$ .** For the soaking-in assays, the reactions were initiated by transferring the EC crystals into  $5\text{ }\mu\text{L}$  hanging drops of EB containing either 1 mM GTP or 2 mM UTP. At the indicated time points, the reactions were quenched by adding an equal volume of electrophoresis sample buffer (89 mM Tris-borate, 50 mM EDTA, 10 M urea, and 0.05% BPB). Before the



**Figure 3.** Raman difference spectrum of 50 mM  $[^{13}\text{C}, ^{15}\text{N}]\text{GTP}$  in an aqueous solution in the absence of magnesium (green). Raman difference spectrum of the TthEC crystal complex 22 min after 1 mM  $[^{13}\text{C}, ^{15}\text{N}]\text{GTP}$  had been soaked into the crystal (pink). Raman difference spectrum 13 min after the crystal has been transferred to a stabilization buffer solution containing no  $[^{13}\text{C}, ^{15}\text{N}]\text{GTP}$  (purple). This is the same crystal used immediately after soak-in conditions. All three Raman difference spectra have baselines added to them. The inset demonstrates the shift in triphosphate mode from an aqueous solution to the crystalline active site environment upon addition of  $*\text{GTP}$ . The signal-to-noise ratio for this series of data is 7/1.

quenching, the crystals were transferred to a fresh drop of the reaction buffer and incubated for 10–30 s, to wash off a potential fraction of the RNAP complex that was dissociated from the crystals and dissolved in the buffer. For the pyrophosphorolysis assays, the reactions were initiated by transferring the EC crystals into 5  $\mu\text{L}$  hanging drops of EB containing 1 mM pyrophosphate. For the soaking-out assays, the crystals were first soaked in EB containing 1 mM GTP or 2 mM UTP. After being soaked for 15 min (for GTP) or 30 min (for UTP), the crystals were then transferred into 5  $\mu\text{L}$  hanging drops of EB without NTPs. These reactions were conducted at 20 °C. The nucleic acids were separated by denaturing PAGE on a 20% gel containing 7 M urea and stained with SYBR Gold (Molecular Probes). Electrophoretic bands were visualized and quantitated with ImageQuant LAS-4000 (GE Healthcare).

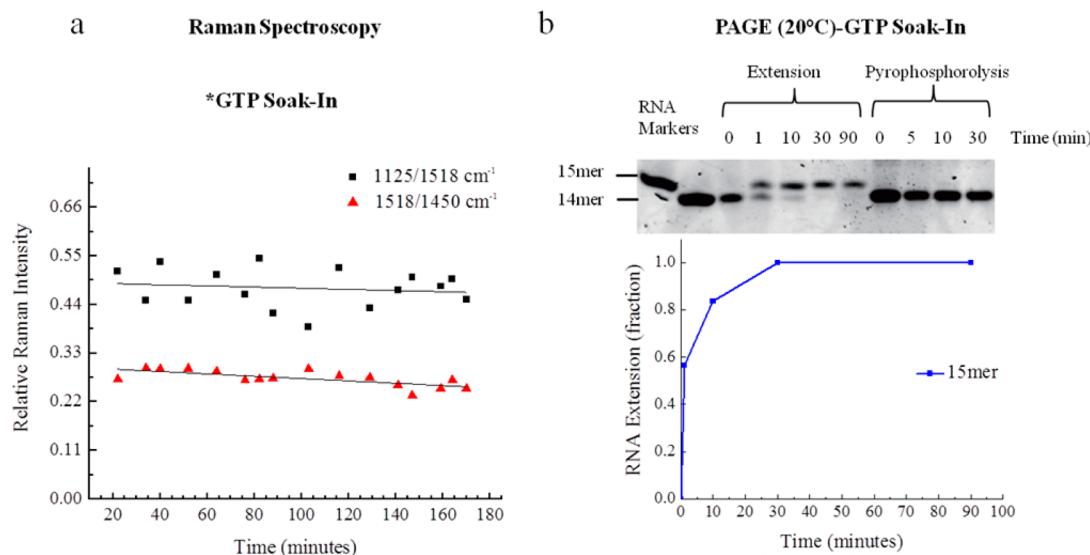
**UTP Soak-In and Soak-Out Assays at 37 °C.** For the soaking-in assays, the crystals were transferred into polypropylene test tubes containing 2  $\mu\text{L}$  of EB and preincubated for 2 min at 37 °C. The reactions were initiated by adding an equal volume of EB containing 4 mM UTP to the tubes and conducted at 37 °C. For the soaking-out assays, the crystals were first soaked in hanging drops of EB containing 2 mM UTP for 90 min at 20 °C. The crystals were then transferred to test tubes containing 4  $\mu\text{L}$  of prewarmed EB and incubated at 37 °C. At the indicated time points, the reactions were quenched by adding equal volumes of the electrophoresis sample buffer to the tubes. The crystals appear to remain intact during incubation (up to 90 min). Because it is technically difficult to wash off the potential fraction

of the RNAP complex dissociated from the crystals, the measured activities might partly include their contribution. The nucleic acids were separated by denaturing PAGE on a 20% gel containing 7 M urea and stained with SYBR Gold (Molecular Probes).

## RESULTS

### Incorporation: RNA Chain Extension following GTP Soak-In.

The primer and template strands are configured such that G is the cognate base for RNA chain extension (Figure 2). When GTP is soaked into single crystals of the *T. thermophilus* RNAP elongation complex (TthEC), this process was followed by time-resolved Raman difference spectroscopy and by PAGE gels. For the Raman experiments,  $^{13}\text{C}$ - and  $^{15}\text{N}$ -labeled GTP, termed  $*\text{GTP}$ , was used to shift the peak positions of G's ring modes into unique positions and prevent confusion with these modes from unlabeled G's already present in the nucleic acid sequences of the crystal co-complex (Figure S5 of the Supporting Information). Figure 3 compares the Raman spectra of aqueous  $*\text{GTP}$  (green) and  $*\text{GTP}$  soaking into (pink) and soaking out of (purple) a TthEC crystal. A key marker band in  $*\text{GTP}$  is the  $*\text{G}$  ring mode near 1518  $\text{cm}^{-1}$  because it is intense and appears in a spectroscopic window. Thus, it can be used to follow the population of  $*\text{G}$  inside the crystal during the reaction. The band near 1118  $\text{cm}^{-1}$  is also highly useful because it represents a molecular vibration of the triphosphate moiety (designated TP) on  $*\text{GTP}$  and can be used to follow the fate of this reacting group in the crystal.



**Figure 4.** (a) Raman kinetics of GTP incorporation. The height intensities of the triphosphate peak at 1125 cm<sup>-1</sup> (black) and guanosine ring mode vibration at 1518 cm<sup>-1</sup> (red) are measured as a function of time. The triphosphate peak is standardized to the 1518 cm<sup>-1</sup> mode within each difference spectrum as an internal standard, while the guanosine ring mode is standardized to the 1450 cm<sup>-1</sup> peak, which is present in the mother spectrum before subtraction. (b) Polyacrylamide gel electrophoresis experiments at 20 °C demonstrate that 2 mM GTP is incorporated into the TthEC crystal (left, PAGE). According to the blue kinetic trace (bottom), approximately 80% of the GTP is incorporated into the RNA chain at 10 min, and by 30 min, mostly all \*GTP has been converted to \*GMP. Pyrophosphorylation experiments at 20 °C demonstrate that the TthEC crystal is in the post-translocated state because adding pyrophosphate to the crystal does not promote cleavage (right, PAGE).

In this series of time-resolved experiments, data collection began at 10 min, but it took 20 min before optical stability allowed reliable Raman spectra to be recorded. The Raman difference spectrum, [TthEC crystal + \*GTP] – [TthEC crystal], 22 min after soak-in commenced is shown in Figure 3 (pink) and has features mostly from \*GTP. The standardized kinetic traces for the \*G/internal standard (1518/1450) and TP/\*G (1125/1518) marker bands from 22 to 170 min in Figure 4a (black and red) show that the populations of the \*G and TP groups are essentially invariant, within the limits of experimental error, over this period. \*G is ratioed to the 1450 cm<sup>-1</sup> peak (from protein -CH<sub>2</sub> deformation modes) in the mother spectrum as an internal standard. Because TP and \*G occur in the same trace, they do not require standardization. Compared to that of aqueous \*GTP in Figure 3 (green), the ratio of the TP to \*G intensity (measured by peak heights) has decreased by ~28% in the kinetic trace. This is seen in Table 1 and Figure 3 (pink) where the average TP/\*G ratio is 0.48 ± 0.04 in the enzyme complex compared to 0.67 (TP/\*G) for aqueous \*GTP [Figure 3 (green)]. These values indicate that approximately 28% of the \*GTP has reacted. Together, these results suggest that by 22 min \*GTP has reached an equilibrium population in the crystal, but

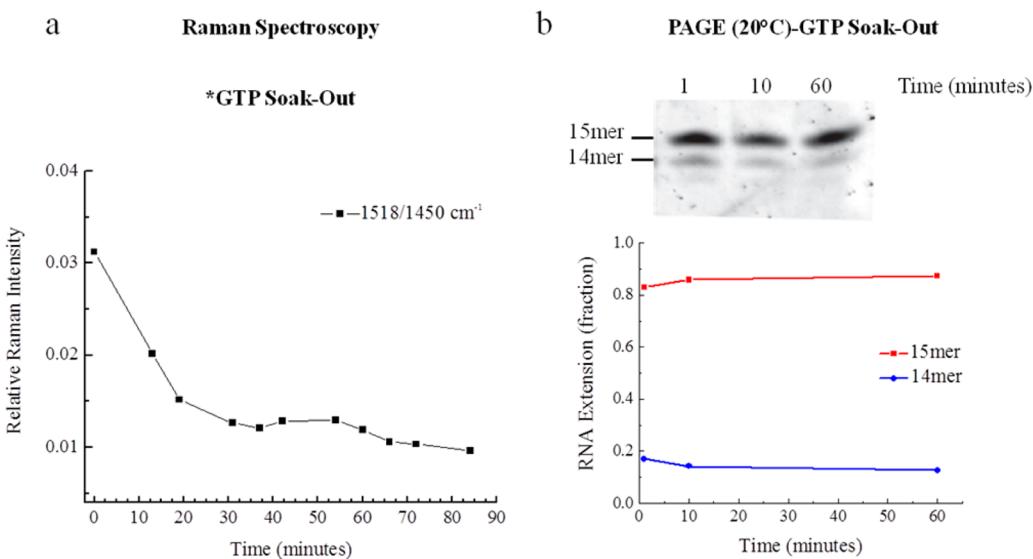
the relative TP population has declined by approximately one-quarter within this time. Thus, a population of unreacted \*GTP remains in the crystal. This is discussed further below. The shift of the TP peak from 1118 cm<sup>-1</sup> (aqueous solution) to 1125 cm<sup>-1</sup> (bound to enzyme) indicates that the conformation of the bound TP changes upon binding to the TthEC (Figure 3, inset).<sup>40</sup>

The notion that G\*MP incorporation has occurred is further supported by PAGE data. Figure 4b confirms that GTP soaks into the crystal and reacts to form the 15mer RNA. The reaction appears to be more than 80% complete at 10 min soak-in, well within the dead time of the Raman data. Thus, it appears that by the time we begin to record the Raman data, the reaction is complete and a depleted, static TP population from the portion of \*GTP that has not reacted is observed. The PAGE results show that the 14mer has been completely converted to the 15mer after 30 min. In addition to confirming \*GTP incorporation, Figure 4b (right) also shows the PAGE results confirming that the EC is in the post-translocated state. The pyrophosphorylation experiment indicates that when pyrophosphate is added to the complex, no cleavage is observed. This means that no nucleotide is occupying the insertion site. Thus, the complex is not likely in the pretranslocated state but rather the post-translocated state. Other notable features of the Raman data in Figure 3 (pink) are the lack of intense bands in the protein's amide I region (1630–1680 cm<sup>-1</sup>). However, there is a weak, broad peak at 1655 cm<sup>-1</sup> that suggests a minor increase in the level of  $\alpha$ -helical structure.<sup>38,41,42</sup>

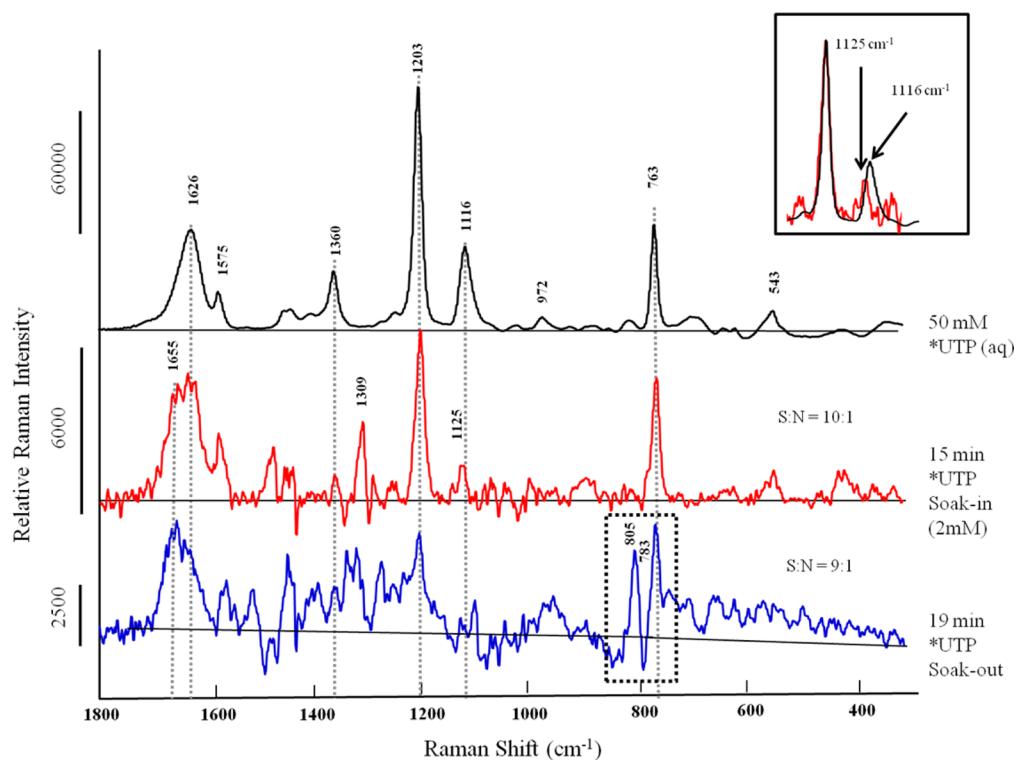
**Removal of Unreacted GTP during Soak-Out.** The Raman data after \*GTP soak-in for 60 min (Figure S2a of the Supporting Information) are very similar to the spectrum seen at 22 min in Figure 3 (pink). When this crystal is relocated to a holding solution containing no \*GTP, this “soaking-out” process removes any noncovalently bound \*GTP. The Raman data at 13 min “soak-out” (Figure 3, purple) show a marked decrease in the intensity of the \*G ring mode at 1518 cm<sup>-1</sup>. The plot of the 1518 cm<sup>-1</sup> mode standardized intensity as a function of time in Figure

**Table 1. Comparison of the TP/\*G and TP/\*U Values to Those Same Ratios in an Aqueous Solution for Calculation of the Percentage of \*GTP and \*UTP Incorporated as \*GMP and \*UMP into the RNA Chain of the TthEC Crystal**

| Raman spectral ratio   | Raman height intensity average |
|--|--------------------------------|
| TP/*G (1125 cm <sup>-1</sup> /1518 cm <sup>-1</sup> ), soak-in | 0.48 ± 0.04                    |
| TP/*U (1125 cm <sup>-1</sup> /1203 cm <sup>-1</sup> ), soak-in | 0.24 ± 0.03                    |
| TP/*G (1125 cm <sup>-1</sup> /1518 cm <sup>-1</sup> ), aqueous | 0.67                           |
| TP/*U (1125 cm <sup>-1</sup> /1203 cm <sup>-1</sup> ), aqueous | 0.33                           |
| % *GTP reacted   | 28% (0.48/0.67)                |
| % *UTP reacted   | 28% (0.24/0.33)                |



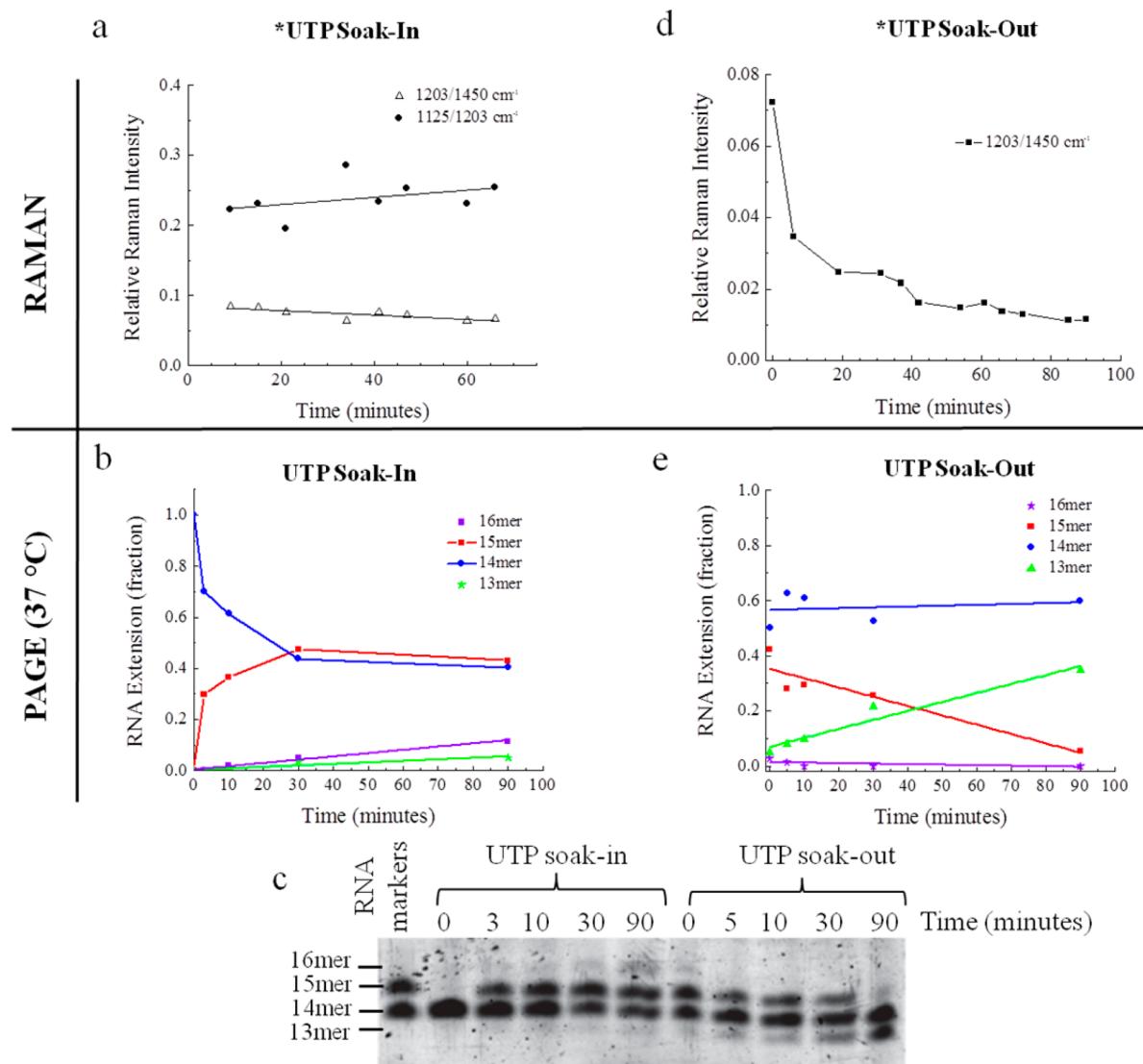
**Figure 5.** (a) Raman kinetic plot that reflects the height intensity of the guanosine ring mode as a function of time during \*GTP soak-out. Soak-out refers to the process by which the crystal, after \*GTP soak-in for several hours, is transferred to a buffer solution containing no \*GTP. In this manner, all noncovalently bound \*GTP can be removed from the crystal. The first point on the plot at time zero represents the \*G height intensity at the last time point of soak-in (167 min). (b) When the \*GTP-soaked crystals were subjected to the soaking-out experiments and analyzed via PAGE at 20 °C, the extended (15mer) and unextended (14mer) band intensities remain similar, and no cleavage product is observed. The 15mer and 14mer RNA markers (left) are shown for comparison. This is confirmed for PAGE kinetics in red and blue (bottom).



**Figure 6.** Raman difference spectrum of 50 mM [<sup>13</sup>C,<sup>15</sup>N]UTP in an aqueous solution (black). Raman difference spectrum of the TthEC crystal complex 21 min after 2 mM [<sup>13</sup>C,<sup>15</sup>N]UTP is soaked into the crystal (red). Raman difference spectrum 19 min after the crystal is transferred to a stabilization buffer solution containing no [<sup>13</sup>C,<sup>15</sup>N]UTP (blue). This is the same crystal used immediately after soak-in conditions. All three Raman difference spectra have baselines added to them. The inset demonstrates the shift in triphosphate mode from an aqueous solution to the crystalline active site environment upon addition of \*UTP. The signal-to-noise ratio for this series of data is 10/1.

5a shows this decrease reaching a plateau after ~30 min. This is the point at which most of the noncovalently bound \*GTP has soaked out. The plateau has a standardized residual intensity of ~0.01 unit compared to 0.03 unit for the last data point of \*GTP

soak-in. These results indicate that ~33% of the \*G population is covalently linked to the RNA chain, whereas 66% is present as noncovalently bound \*GTP. This result is consistent with the soak-in Raman data that indicate 28% \*G is covalently bound



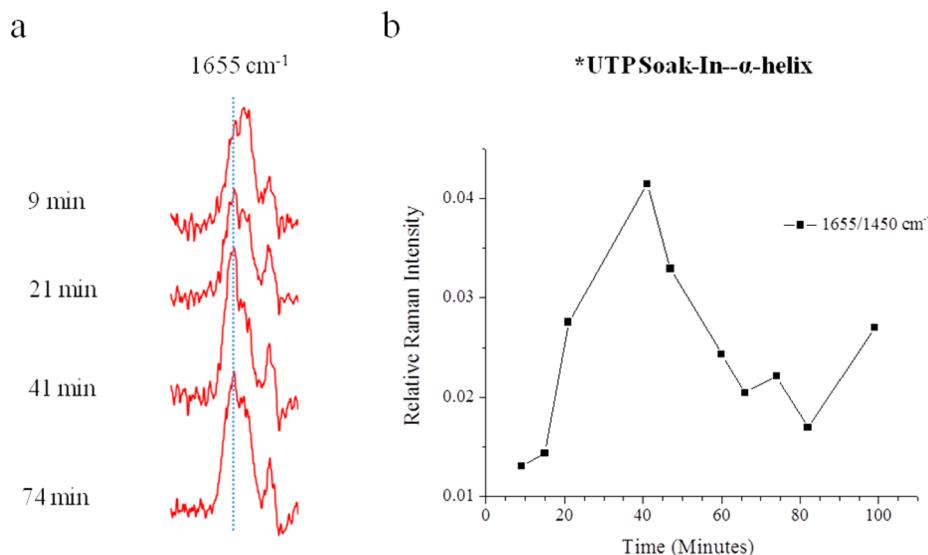
**Figure 7.** (a) Raman kinetics of \*UTP misincorporation. The height intensities of the triphosphate peak at  $1125\text{ cm}^{-1}$  (red) and uridine ring mode vibrations at  $1203\text{ cm}^{-1}$  (black) are plotted as a function of time. The triphosphate peak is standardized to the  $1203\text{ cm}^{-1}$  mode within each difference spectrum as an internal standard, while the uridine ring mode is standardized to the  $1450\text{ cm}^{-1}$  peak, which is present in the mother spectrum before subtraction. (b) Misincorporation of UTP does occur in the crystal according to PAGE kinetic experiments conducted at  $37\text{ }^{\circ}\text{C}$ . It occurs over a 30 min time period, which is slower than the incorporation of the correct substrate (GTP). After a long period of incubation, the second misincorporation also takes place to produce the 16mer RNA. The 13mer RNA is likely to be produced by the cleavage of the 15mer RNA. (c) PAGE results of 2 mM UTP soak-in and soak-out at  $37\text{ }^{\circ}\text{C}$ . The time course for UTP soak-in shows the 14mer RNA chain being elongated to a 15mer within 3 min of the reaction. Upon soak-out for 90 min, the 15mer is cleaved into a 14mer (right). (d) The Raman kinetic plot reflects the height intensity of the \*U ring mode as a function of time during \*UTP soak-out (buffer solution containing no \*UTP). All noncovalently bound \*UTP is removed from the crystal. The first point on the plot at time zero represents the \*U height intensity at the last time point of soak-in. (e) PAGE kinetics upon UTP soak-out at  $37\text{ }^{\circ}\text{C}$ . The level of 15mer RNA (red) decreases, and the 13mer RNA (green) accumulates instead. This indicates that the misincorporated UTP was removed by the RNAP RNA cleavage activity.

(refer to Table 1). The kinetics of \*GTP soak-out seen via PAGE are shown in Figure 5b. These demonstrate the 14mer and 15mer populations remain essentially invariant; the ratio of the extended (15mer) and unextended (14mer) RNAs remains similar, and no cleavage product was observed. This is consistent with the Raman kinetics (Figure 5a) that show covalently bound \*G also remains essentially constant from 30 to 80 min. The presence of noncovalently bound \*GTP raises the possibility of a second NTP binding site. As shown below for \*UTP misincorporation, a second NTP binding site may play a role in the enzyme's conformational changes.

#### Misincorporation: RNA Chain Extension following UTP Soak-In.

Although \*UTP is a noncognate base, under the conditions of our *in crystallo* experiments, it appears to be misincorporated into the RNA chain. Moreover, in contrast to cognate \*GTP incorporation, the \*UTP causes major perturbations to the conformational landscape of both the protein and the nucleic acid secondary structures.

Figure 6 compares the Raman spectrum of aqueous \*UTP (black) with the difference spectra for \*UTP soaking into (red) and soaking out of (blue) TthEC crystals. \*UTP has intense ring modes at  $1203$  and  $763\text{ cm}^{-1}$ , which provide good marker bands for the population of \*U. Likewise, the intensity of the  $1116\text{ cm}^{-1}$



**Figure 8.** (a) Raman intensity at 1655 cm<sup>-1</sup> throughout \*UTP soak-in (because of amide I vibrations of proteins and  $\alpha$ -helix contributions) that increases from 9 to 41 min, subsequently decreases, and reaches a plateau at 60 min. (b) Raman kinetics of the height of the 1655 cm<sup>-1</sup> peak from 9 to 100 min during \*UTP soak-in.

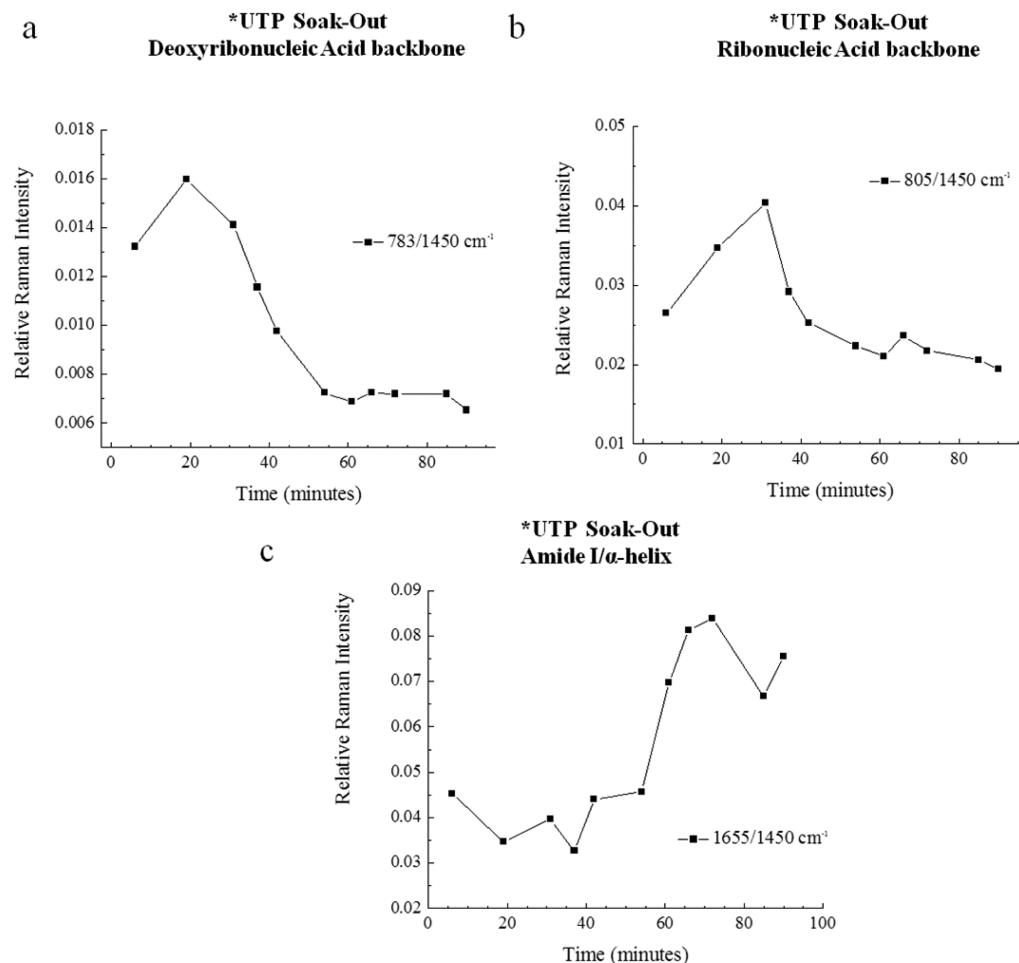
band provides a monitor for the triphosphate group accompanying the \*U moiety. \*UTP also has an intense feature near 1626 cm<sup>-1</sup> that is due to modes that include contributions from the uracil C=O groups.<sup>43</sup> The time dependence of \*U soak-in depicted in Figure 7a (empty triangles) shows that the standardized intensity of the \*U ring mode at 1203 cm<sup>-1</sup> is essentially unchanged from 10 to 60 min, and thus, the \*U ring population has reached equilibrium within 10 min. Evidence that the reaction occurs in the first 10 min is observed in Figure 6 where the intensity of the 1116 cm<sup>-1</sup> TP peak of aqueous \*G (black) relative to that of TP (red, soak-in) exhibits a notable decrease upon soak-in. Figure 7a (circles) plots the ratio of the intensities of the 1125 cm<sup>-1</sup>/1203 cm<sup>-1</sup> (TP/\*U) modes with time (using the \*U ring mode as an internal intensity standard), and it is also essentially invariant between 10 and 60 min. However, the average value for TP/\*U during \*UTP soak-in is approximately 0.24 ± 0.03 (Figure 7a, 1125/1203 cm<sup>-1</sup>) compared to the value of 0.33 (TP/\*U) for aqueous \*UTP seen in Figure 6 (black). Together, these data suggest that approximately 25% of the TP population in the crystal has reacted and been incorporated and that the reaction is finished at 10 min. Nucleotide incorporation is confirmed by the PAGE analysis shown in Figure 7c. During soak-in, the 15mer chain begins to appear at 3 min and clearly by 10 min. Interestingly at 90 min, double \*U incorporation also appears as a smaller population (Figure 7c). Evidence of incorporation is seen more quickly in the Raman data where ~25% of the \*UTP available in the crystal appears to have been incorporated within the dead time of the experiments, 0–10 min. The PAGE data show that at 37 °C this UTP is incorporated in approximately 50% of the RNA chain by 40 min and remains relatively constant through 90 min.

The Raman kinetic data in Figure 7a suggest that the \*U is fully incorporated at 10 min because the TP (1125 cm<sup>-1</sup>)/\*U (1203 cm<sup>-1</sup>) ratio is 0.33 for aqueous \*UTP and decreases to 0.24 (Table 1) during soak-in and remains constant within experimental error. However, the PAGE data (Figure 7b) show that the extent of incorporation is approximately 50% at 10 min. The more rapid incorporation seen in the Raman data is due to

the heating effect of the laser beam focused in the crystal. The lower limit of the temperature is in the range of 32–37 °C (see Experimental Procedures).

Soaking in also produces a slow nearly reversible change in protein conformation. A shoulder appears on \*U's carbonyl (C=O) profile at 1655 cm<sup>-1</sup>; its intensity increases with time and is maximal at 40 min (see Figure 8a) and nearly reverses by 74 min (Figure 8b). This is the spectral region in which amide I vibrations from  $\alpha$ -helical peptide bonds occur, and thus, we are observing an increase in  $\alpha$ -helical secondary structure content.<sup>27,39</sup> It is possible that we are observing the formation of the  $\alpha$ -helical trigger helix, but this is occurring long after catalysis is complete and does not precede the phosphodiester bond formation step as proposed for the mechanism. However, it does underscore the plasticity of the conformational events and indicates the finely balanced conformational states of the trigger helix.

**UTP Soak-Out.** \*UTP soak-out produces substantial changes in both protein and nucleic acid conformational markers. Using the same crystal after a 100 min soak-in, Figure 6 (blue) shows the Raman difference spectrum 19 min after soak-out begins. The \*U ring population at 1203, 1360, and 1626 cm<sup>-1</sup> is diminished but not eliminated. Figure 7d plots the standardized intensity of the intense \*U mode at 1203 cm<sup>-1</sup> as a function of time. It decreases steadily from 20 to 40 min, and the intensity plateaus from 40 to 100 min. The 1203 cm<sup>-1</sup>/1450 cm<sup>-1</sup> intensity ratio at 30 min indicates that ~35% of \*U remains. This is \*U incorporated in the 15mer and is in reasonable accord with the value derived from \*UTP soak-in. After a 90 min soak-out, this figure is reduced to ~15%, suggesting that the RNA chain may be undergoing endonuclease cleavage. This is supported by the PAGE data (Figure 7e) that show a reduction in the level of the 15mer and an increase in the level of the 13mer over 100 min. The population of the unextended 14mer does not change (Figure 7e). This population likely represents the 14mer population that did not react in the soak-in. Together, the Raman and PAGE data reveal that a significant portion of misincorporated \*UTP has been removed by RNAP's cleavage activity.



**Figure 9.** (a) Raman kinetics of the peak at  $805\text{ cm}^{-1}$  in the spectra demonstrate the time-dependent changes in ribonucleic acid backbone vibrations during  ${}^*\text{UTP}$  soak-out. (b) The Raman kinetic plot demonstrates the time dependence of the nucleic acid backbone vibrations at  $783\text{ cm}^{-1}$  during  ${}^*\text{UTP}$  soak-out. (c) The Raman kinetic plot, which follows the change in  $1655\text{ cm}^{-1}$  intensity throughout  ${}^*\text{U}$  soak-out, begins to increase starting at approximately 55 min, before reaching a maximum at 70 min.

A unique finding from the Raman data is that removing noncovalently bound  ${}^*\text{UTP}$  by soak-out has marked effects on nucleic acid and protein conformations. The striking new features in Figure 6 (blue, soak-out) at  $805$  and  $783\text{ cm}^{-1}$  are due to changes in RNA and DNA backbone conformations,<sup>33,44</sup> respectively. They are maximal at 20–25 min and decrease to a plateau after 1 h (Figure 9a,b). A relatively intense amide I feature near  $1655\text{ cm}^{-1}$  builds in slowly and plateaus at ~60 min (Figure 9c). This suggests that a significant change in  $\alpha$ -helix population has occurred. It is notable that the nucleic acid and protein conformational changes brought about by soaking out  ${}^*\text{UTP}$  (Figure 9a–c) occur on the same time scale as the slow decrease in the  ${}^*\text{U}$  Raman intensity of  ${}^*\text{UTP}$  (Figure 7d) and 15mer decay detected in the PAGE experiments (Figure 7e). Thus, we propose that we are observing conformational changes in the enzyme complex that are needed to bring the 3' end of the RNA chain into the endonuclease region of the complex.

## DISCUSSION

**Factors Determining Raman Instrumental Dead Time.** Obtaining data from crystals at early time points is technically challenging because of the change in the orientation of the crystal immediately after exposure to reaction conditions. First, the Raman spectrum of the apo crystal scattering in the hanging drop is obtained. Then the substrate solution is transferred into the

hanging drop, or a loop is used to transfer the apo crystal into a hanging drop containing substrate. In either case, it takes several minutes for the crystal position to stabilize before Raman data for a reliable series of time points can be collected. For the crystal used to generate the data in Figures 3 and 4, we have time points at 10 and 16 min; however, they are not included in the analysis because they may have been compromised by crystal movement. Obtaining data requires patience, dexterity, and an element of good luck. For the smaller N4 phage RNAP (~110 kDa) and Dpo4 DNAP (~40 kDa), time points have been successfully obtained as early as 5 min after soak-in has begun.<sup>27,45</sup> However, TthEC is much larger (~400 kDa), making Raman difference spectra more challenging because the required signal is a smaller fraction of the total vibrational spectrum of the Tth complex. Thus, precise Raman difference spectra are needed, and it takes the Tth crystals longer to reach the required level of stability. In practice, 10 min will be the shortest dead time.

The dead time may be reduced by holding the crystal in a loop or in a capillary against a wad of porous glass fiber as described by Petsko<sup>46</sup> for X-ray analysis. However, the method will likely lead to an increase in the level of deleterious scattering from the interfaces. Our lab has had success in slowing reactions down by mounting a DNA polymerase crystal in a cold cell and accessing earlier events in the reaction. Unfortunately, all these approaches

will likely reduce the signal-to-noise ratio, and the data would be below the acceptable limit.

**Incorporation of GTP and UTP.** The Raman experiments indicate that soak-in of  $^*\text{GTP}$  and  $^*\text{UTP}$  and their subsequent incorporation into the nascent RNA chain are rapid on the experimental Raman time scale (Figures 4a and 7a). This conclusion is supported by PAGE (Figures 4b and 7c). Because the dead time before Raman data could be collected in the  $^*\text{GTP}$  soak-in reaction run was 20 min, the PAGE data provide the most accurate time definition. As Figure 4b shows for  $^*\text{GTP}$ , the reaction is 80% complete after soak-in for 10 min at 20 °C. Thus, it was not possible to follow the time course of soak-in and incorporation by Raman. The kinetic Raman crystallographic plots for ring modes and triphosphate features show an essentially straight line parallel to the  $x$ -axis (see Figure 4a and Table 1 for standard deviations), confirming that the reaction in the crystal was mostly complete at ~20 min. Kinetics in the Raman experiments are likely accelerated by the heating effect of the laser beam. For  $^*\text{UTP}$  soak-in, the relative Raman intensity of the  $^*\text{U}$  ring mode ( $1203\text{ cm}^{-1}$ ) appears to be maximal at 10 min (Figure 7a, empty triangles) and remains approximately constant. The  $1125\text{ cm}^{-1}/1203\text{ cm}^{-1}$  intensity ratio (circles) shows a similar behavior with apparent incorporation being complete by the first time point, which was ~10 min. Again, this is faster than the rate of incorporation measured by PAGE (Figure 7c), which is not complete until 30 min (Figure 7b). Interestingly, whereas  $^*\text{GTP}$  incorporation goes to 100% completion,  $^*\text{UTP}$  misincorporation plateaus at ~50% of the original 14mer population. After a long period of incubation, the second misincorporation also takes place to produce the 16mer RNA, although this is a minor population (Figure 7b,c). The 2-nucleotide-extended product was not observed in the GTP soaking-in experiments. Because the +2 template is dG, the less bulky U may be more easily misincorporated.

When PAGE analyses were initially conducted for both GTP and UTP, the temperature of the assays was 20 °C. The PAGE results and Raman results agreed for GTP soak-in and soak-out, however, did not agree very well for UTP; the Raman data indicated that the UTP misincorporation reaction occurred very fast [within the first 10 min (Figure 7a)], while the PAGE data showed slow incorporation of UTP occurring over a period of 90 min (Figure S7 of the Supporting Information). This is in contrast to the 37 °C results previously mentioned, where UTP becomes incorporated almost 3 times as fast in a period of 30 min (Figure 7b). Thus, the PAGE analyses at higher temperatures, although they do not perfectly match the Raman kinetics, agree with each other more closely than the results of PAGE analysis at 20 °C for UTP soak-in. The faster rate seen in the Raman data may be due, in part, to the heating effect on the crystal of the excitation laser beam, especially within the focal volume.

However, comparison of PAGE results at 20 and 37 °C for UTP soak-out (Figure 7e and Figure S7 of the Supporting Information) shows that even at higher temperatures, soak-out and RNA cleavage take >1 h to complete (Figure 7c,d). The 13mer RNA (Figure 7c) is likely produced by the cleavage of the 15mer RNA. Therefore, the Raman kinetics of  $^*\text{UTP}$  soak-out match those of the reported PAGE kinetics for UTP soak-out.

Upon comparison of the PAGE and Raman kinetic traces, it must be kept in mind that the laser beam heats the crystal by 15 °C or more (see Experimental Procedures).<sup>31</sup> Thus, on the basis of the van't Hoff equation, we expect the reaction to occur 3 times faster during Raman experiments at 37 °C compared to the PAGE time courses at 20 °C (Figure S7 of the Supporting

Information). This may be an oversimplification because other factors may come into play. For example, increasing the crystal temperature may reduce solvent channel access as the protein becomes more mobile. Also, the thermophilic enzyme may perform more efficiently when it is closer to its preferred reaction temperature of 65–70 °C.

**Minor Protein and Nucleic Acid Conformational Changes Occur upon G Base Incorporation.** There are two regions of the Raman spectra that are highly useful for following changes in protein and nucleic acid secondary structure. For proteins, the region from  $1620$  to  $1700\text{ cm}^{-1}$  is dominated by vibrational modes associated with the peptide stretching motion involving the carbonyl ( $\text{C}=\text{O}$ ) groups. This is the amide I region, and different secondary structures give rise to characteristic peaks within the region. For example,  $\alpha$ -helices have Raman amide I signatures in the  $1645$ – $1660\text{ cm}^{-1}$  range, and  $\beta$ -sheet amide I modes are found between  $1665$  and  $1680\text{ cm}^{-1}$ .<sup>38</sup> For the nucleic acids, the phosphodiester backbones have stretching motions that give rise to peaks in the  $800\text{ cm}^{-1}$  region.<sup>44</sup> In this context, it is noteworthy that A and B forms of DNA have a relatively intense Raman backbone mode near  $785\text{ cm}^{-1}$ ,<sup>33,36,47</sup> while the A form of RNA has an intense backbone mode near  $810\text{ cm}^{-1}$ .<sup>44</sup>

Using these criteria, it is apparent that  $^*\text{G}$  incorporation is not accompanied by major changes in protein or nucleic acid secondary structure after the experimental dead time (Figure 3, pink). With the exception of the poorly defined feature near  $1655\text{ cm}^{-1}$ , which may be due to an increase in  $\alpha$ -helical content following incorporation, there appears to be little change in secondary structure. The TEC crystals grown have been confirmed to be in the post-translocated state because pyrophosphorolysis experiments indicate no cleavage of the 14mer RNA upon addition of pyrophosphate (Figure 4b). Given the latter knowledge, the *Nature* papers by Vassylyev et al.<sup>24,48</sup> that report on the trigger loop to trigger helix formation associated with the nucleotide addition cycle, and the fact that we do not observe translocation in the crystal, we conclude that it is reasonable to detect just a low intensity from a small increase in  $\alpha$ -helical content in the amide I region of the Raman spectrum ( $1655\text{ cm}^{-1}$ ).

In the soak-out data, the  $^*\text{G}$  ring modes lose a major portion of intensity because of unreacted  $^*\text{GTP}$  soaking-out from the crystal (Figure 3, purple). Thus, there is no evidence of concurrent protein or nucleic acid secondary structure changes. This could mean that reversible conformational changes have occurred before we begin recording Raman data or that  $^*\text{GTP}$  incorporation can occur for this construct without large movements of protein or nucleic acid. This is in contrast to UTP, for which we do detect large conformational changes.

**Substantial Protein and Nucleic Acid Changes Occurring upon U Base Misincorporation.** The contrast with  $^*\text{UTP}$  misincorporation is striking. During  $^*\text{U}$  soak-in, a well-defined shoulder near  $1655\text{ cm}^{-1}$  appears on uracil's "carbonyl" profile (Figure 8a) at 9 min, and its intensity is maximal at 40 min before slowly declining to a plateau at 65 min (Figure 8b). The shoulder is reliably assigned to an increase in  $\alpha$ -helical content, suggesting that  $^*\text{UTP}$  soak-in may have led to the near reversible formation of the trigger helix. In addition, soaking-out (after  $^*\text{UTP}$  soak-in) leads to complex conformational changes. The changes include the appearance of a striking differential at  $783$  and  $805\text{ cm}^{-1}$ . The latter marker bands, which are due to DNA and RNA backbone changes, appear within the experimental dead time, and their intensities are maximal at 20–25 min and

decline to a finite plateau at ~50 min (Figure 9a,b). After little change, the intensity of the  $\alpha$ -helix marker at  $1655\text{ cm}^{-1}$  increases from 55 to 70 min (Figure 9c).

The protein and nucleic acid backbone changes seen upon \*UTP soak-in and soak-out are slow (tens of minutes) and are likely not directly related to the catalytic act that occurs much more quickly (<10 min). However, these observations suggest that noncognate base misincorporation drives slow conformational changes to a greater degree compared to cognate base incorporation. We speculate that we may be glimpsing conformational changes that prepare the enzyme for proofreading. This is supported by a decline in the \*U ring mode intensities after 50 min seen in Figure 7d that may be due to slow endonuclease activity. Importantly, the PAGE data (Figure 7c,e, right) also confirm slow cleavage activity during \*UTP soak-out experiments.

As a control experiment, r\*ATP ( $^{13}\text{C}$ - and  $^{15}\text{N}$ -labeled) was employed for a soak-in experiment to test for misincorporation. Adenine (\*A) ring modes are apparent; however, the relative intensity of the triphosphate band at  $1125\text{ cm}^{-1}$  does not decrease, indicating that incorporation has not occurred (Raman data not shown). When the crystal is transferred to a holding solution with no ATP (soak-out) for 1 h, the \*A ring mode bands disappear, confirming that significant misincorporation has not occurred. In addition, major bands associated near the  $800\text{ cm}^{-1}$  region in the difference spectrum after soak-out are not observed. This shows that, in contrast to the case for \*UTP misincorporation, major changes are not occurring in the nucleic acid phosphodiester backbone upon addition of the noncognate nucleotide ATP. Additionally, r\*ATP does not trigger significant protein conformational changes.

**A Second NTP Binding Site?** In addition to the NTP that is incorporated, the Raman experiments demonstrate that for both \*GTP and \*UTP, an excess of unreacted NTP remains in the crystal after soak-in (see the values in Table 1). For the cognate base, \*G soak-in leads to a small increase in the  $\alpha$ -helix population. The excess \*GTP can be soaked out with no apparent change in the protein or nucleic acid conformations. In contrast, soaking-in noncognate \*UTP brings about a well-defined, partially reversible formation of  $\alpha$ -helix. Moreover, subsequently soaking-out \*UTP leads to marked changes in amide modes (likely from  $\alpha$ -helices) and the backbone modes of the DNA and RNA chains (Figure 6, blue). The observations for noncognate \*UTP suggest that there may be a second NTP binding site, separate from the active site, that affects both protein and nucleic acid conformations. This is speculative, and our data cannot identify possible locations for this site. However, the data do demonstrate that a noncognate NTP being added or removed from the crystal complex can trigger quite remarkable conformational changes, which raises the possibility that a second conformationally active binding site distinct from the active site exists. However, it is entirely possible that some of the nonspecifically bound \*UTPs reside in solvent channels and are conformationally neutral. We cannot distinguish between separate populations.

## CONCLUSION

The Raman and PAGE data show that RNA chain extension occurs following soak-in of cognate GTP into the crystal. At present, the extension reaction in the crystal is too fast for us to follow in real time by Raman microscopy; however, PAGE experiments reveal it is 80% complete in 10 min. However, after the dead time for soak-in of \*GTP, only minor conformational

changes could be seen in the Raman data for the protein in the TthEC complex. In the future, the use of a cold cell may slow the reaction and provide Raman access to the extension reaction itself. Noncognate \*UTP could also be incorporated into the RNA chain, although in the PAGE experiment after soak-in for 30 min at  $37^\circ\text{C}$ , the reaction appears to stop with ~50% of the 14mer in the crystal converted to 15mer. In the Raman experiments, the nucleotidyl transfer reaction appears to occur faster. It is nearly completed within approximately 10–15 min of the instrumental dead time after soak-in commences. This is likely due to a higher temperature within the focal volume of the laser beam. During \*UTP soak-in, a well-defined and reversible increase in  $\alpha$ -helical content in the EC is indicated by the Raman data. The Raman experiments for both \*GTP and \*UTP soak-ins reveal the presence of major populations of unreacted NTPs in the crystals, in addition to the \*GMPs and \*UMPs that have been incorporated by RNA chain extension. Soaking out the unreacted \*GTPs has no effect on the Raman difference spectra from the crystals, but soaking out unreacted \*UTPs brings about substantial changes in protein and nucleic acid conformations. These changes are tentatively assigned to the TthEC preparing for proofreading. This concept is supported by the PAGE data that reveal some RNA cleavage does occur during \*UTP soak-out from the crystal.

This study provides some insight into the utility of following reactions in macromolecular crystals by Raman crystallography. Our bacterial model system is a large EC complex, and the data are near the threshold in terms of signal-to-noise ratio. This can be improved by using a directly coupled microscope, which should improve the signal by 1 order of magnitude. The second issue associated with large “molecular machines” comes with interpreting the Raman data. This is best addressed by supplementing Raman with another technique, such as PAGE analysis.

## ASSOCIATED CONTENT

### Supporting Information

Calculation of Tth elongation complex and *T. thermophilus* protein concentrations in the active site of a single crystal, a Raman spectrum of the *T. thermophilus* elongation complex with [ $^{13}\text{C},^{15}\text{N}$ ]UTP before secondary subtraction, a complete series of Raman spectra for \*GTP soak-in, a complete series of Raman spectra for \*UTP soak-in, a comparison of two TthEC crystals soaked with \*UTP at similar time points, an example of a signal-to-noise ratio calculation, Raman spectra of labeled and unlabeled GTP, Raman spectra of labeled and unlabeled UTP, and PAGE analysis of UTP soak-in and soak-out at  $20^\circ\text{C}$ . This material is available free of charge via the Internet at <http://pubs.acs.org>.

## AUTHOR INFORMATION

### Corresponding Author

\*E-mail: prc5@case.edu.

### Present Address

<sup>†</sup>B.A.W.: Janssen Research & Development, LLC, Spring House, PA 19477.

### Funding

This work was supported in part by a Japan Society for the Promotion of Science (JSPS) Grant-in-Aid for Scientific Research (to S.S. and S.Y.), the Targeted Proteins Research Program (TPRP), the Ministry of Education, Culture, Sports, Science and Technology (MEXT) of Japan (to S.Y.), and

National Institute of General Medical Sciences Grant R01 GM54072 (to P.R.C.).

## Notes

The authors declare no competing financial interest.

## ACKNOWLEDGMENTS

We thank Dr. Kastuhiko S. Murakami at Pennsylvania State University for providing purified *T. thermophilus* protein as well as nucleic acid templates.

## ABBREVIATIONS

RNAP, RNA polymerase; Tth, *T. thermophilus*; EC, elongation complex; TthEC, *T. thermophilus* elongation complex; \*GTP, [ $^{13}\text{C}, ^{15}\text{N}$ ]guanosine triphosphate; \*UTP, [ $^{13}\text{C}, ^{15}\text{N}$ ]uridine triphosphate; \*G, [ $^{13}\text{C}, ^{15}\text{N}$ ]guanosine portion of guanosine triphosphate; \*U, [ $^{13}\text{C}, ^{15}\text{N}$ ]uridine portion of uridine triphosphate; TP, triphosphate; NTP, ribonucleotide triphosphate; NMP, ribonucleotide monophosphate; EB, equilibrium buffer.

## REFERENCES

- (1) Zhang, G. Y., Campbell, E. A., Minakhin, L., Richter, C., Severinov, K., and Darst, S. A. (1999) Crystal structure of *Thermus aquaticus* core RNA polymerase at 3.3 Å resolution. *Cell* 98, 811–824.
- (2) Cramer, P., Bushnell, D. A., and Kornberg, R. D. (2001) Structural basis of transcription: RNA polymerase II at 2.8 Å resolution. *Science* 292, 1863–1876.
- (3) Gnatt, A. L., Cramer, P., Fu, J. H., Bushnell, D. A., and Kornberg, R. D. (2001) Structural basis of transcription: An RNA polymerase II elongation complex at 3.3 Å resolution. *Science* 292, 1876–1882.
- (4) Vassylyeva, M. N., Lee, J., Sekine, S., Laptenko, O., Kuramitsu, S., Shibata, T., Inoue, Y., Borukhov, S., Vassylyev, D. G., and Yokoyama, S. (2002) Purification, crystallization and initial crystallographic analysis of RNA polymerase holoenzyme from *Thermus thermophilus*. *Acta Crystallogr. D58*, 1497–1500.
- (5) Vassylyev, D. G., Sekine, S., Laptenko, O., Lee, J., Vassylyeva, M. N., Borukhov, S., and Yokoyama, S. (2002) Crystal structure of a bacterial RNA polymerase holoenzyme at 2.6 Å resolution. *Nature* 417, 712–719.
- (6) Murakami, K. S., Masuda, S., and Darst, S. A. (2003) Crystallographic analysis of *Thermus aquaticus* RNA polymerase holoenzyme and a holoenzyme/promoter DNA complex. *Methods Enzymol.* 370, 42–53.
- (7) Cheetham, G. M. T., and Steitz, T. A. (1999) Structure of a transcribing T7 RNA polymerase initiation complex. *Science* 286, 2305–2309.
- (8) Yin, Y. W., and Steitz, T. A. (2002) Structural basis for the transition from initiation to elongation transcription in T7 RNA polymerase. *Science* 298, 1387–1395.
- (9) Kennedy, S. R., and Erie, D. A. (2011) Templated nucleoside triphosphate binding to a noncatalytic site on RNA polymerase regulates transcription. *Proc. Natl. Acad. Sci. U.S.A.* 108, 6079–6084.
- (10) Holmes, S. F., Foster, J. E., and Erie, D. A. (2003) Kinetics of multisubunit RNA polymerases: Experimental methods and data analysis. *Methods Enzymol.* 371, 71–81.
- (11) Kettenberger, H., Armache, K. J., and Cramer, P. (2004) Complete RNA polymerase II elongation complex structure and its interactions with NTP and TFIIS. *Mol. Cell* 16, 955–965.
- (12) Westover, K. D., Bushnell, D. A., and Kornberg, R. D. (2004) Structural basis of transcription: Nucleotide selection by rotation in the RNA polymerase II active center. *Cell* 119, 481–489.
- (13) Westover, K. D., Bushnell, D. A., and Kornberg, R. D. (2004) Structural basis of transcription: Separation of RNA from DNA by RNA polymerase II. *Science* 303, 1014–1016.
- (14) Wang, D., Bushnell, D. A., Westover, K. D., Kaplan, C. D., and Kornberg, R. D. (2006) Structural basis of transcription: Role of the trigger loop in substrate specificity and catalysis. *Cell* 127, 941–954.
- (15) Toulokhonov, I., Zhang, J., Palangat, M., and Landick, R. (2007) A central role of the RNA polymerase trigger loop in active-site rearrangement during transcriptional pausing. *Mol. Cell* 27, 406–419.
- (16) Erie, D. A., and Kennedy, S. R. (2009) Forks, pincers, and triggers: The tools for nucleotide incorporation and translocation in multi-subunit RNA polymerases. *Curr. Opin. Struct. Biol.* 19, 708–714.
- (17) Zhang, J., Palangat, M., and Landick, R. (2010) Role of the RNA polymerase trigger loop in catalysis and pausing. *Nat. Struct. Mol. Biol.* 17, 99–104.
- (18) Sosunov, V., Sosunova, E., Mustaev, A., Bass, I., Nikiforov, V., and Goldfarb, A. (2003) Unified two-metal mechanism of RNA synthesis and degradation by RNA polymerase. *EMBO J.* 22, 2234–2244.
- (19) Abbondanzieri, E. A., Greenleaf, W. J., Shaevitz, J. W., Landick, R., and Block, S. M. (2005) Direct observation of base-pair stepping by RNA polymerase. *Nature* 438, 460–465.
- (20) Larson, M. H., Landick, R., and Block, S. M. (2011) Single-Molecule Studies of RNA Polymerase: One Singular Sensation, Every Little Step It Takes. *Mol. Cell* 41, 249–262.
- (21) Burton, Z. F., Feig, M., Gong, X. Q., Zhang, C. F., Nedialkov, Y. A., and Xiong, Y. L. (2005) NTP-driven translocation and regulation of downstream template opening by multi-subunit RNA polymerases. *Biochem. Cell Biol.* 83, 486–496.
- (22) Tagami, S., Sekine, S., and Yokoyama, S. (2011) A novel conformation of RNA polymerase sheds light on the mechanism of transcription. *Transcription* 2, 162–167.
- (23) Kireeva, M., Kashlev, M., and Burton, Z. F. (2010) Translocation by multi-subunit RNA polymerases. *Biochim. Biophys. Acta* 1799, 389–401.
- (24) Vassylyev, D. G., Vassylyeva, M. N., Zhang, J., Palangat, M., Artsimovitch, I., and Landick, R. (2007) Structural basis for substrate loading in bacterial RNA polymerase. *Nature* 448, 163–168.
- (25) Tagami, S., Sekine, S.-i., Kumarevel, T., Hino, N., Murayama, Y., Kamagamori, S., Yamamoto, M., Sakamoto, K., and Yokoyama, S. (2010) Crystal structure of bacterial RNA polymerase bound with a transcription inhibitor protein. *Nature* 468, 978–982.
- (26) Sekine, S.-i., Tagami, S., and Yokoyama, S. (2012) Structural basis of transcription by bacterial and eukaryotic RNA polymerases. *Curr. Opin. Struct. Biol.* 22, 110–118.
- (27) Chen, Y., Basu, R., Gleghorn, M. L., Murakami, K. S., and Carey, P. R. (2011) Time-Resolved Events on the Reaction Pathway of Transcript Initiation by a Single-Subunit RNA Polymerase: Raman Crystallographic Evidence. *J. Am. Chem. Soc.* 133, 12544–12555.
- (28) Warner, B. (2012) Structural Study of Bacterial RNA Polymerase Conformational Changes During Transcription Elongation. M.S. Thesis, Pennsylvania State University, University Park, PA.
- (29) Murayama, Y., Sekine, S.-i., and Yokoyama, S. (2013) Crystallization and preliminary X-ray crystallographic analyses of *Thermus thermophilus* backtracked RNA polymerase. *Acta Crystallogr. F69*, 174–177.
- (30) Basu, R. S., and Murakami, K. S. (2013) Watching the Bacteriophage N4 RNA Polymerase Transcription by Time-dependent Soak-trigger-freeze X-ray Crystallography. *J. Biol. Chem.* 288, 3305–3311.
- (31) Carey, P. R. (2006) Raman crystallography and other biochemical applications of Raman microscopy. *Annu. Rev. Phys. Chem.* 57, 527–554.
- (32) Carey, P. R., Chen, Y., Gong, B., and Kalp, M. (2011) Kinetic crystallography by Raman microscopy. *Biochim. Biophys. Acta* 1814, 742–749.
- (33) Duguid, J., Bloomfield, V. A., Benevides, J., and Thomas, G. J. (1993) Raman spectral studies of nucleic acids. Raman spectroscopy of DNA-metal complexes. I. Interactions and conformational effects of the divalent cations: Mg, Ca, Sr, Ba, Mn, Co, Ni, Cu, Pd, and Cd. *Biophys. J.* 65, 1916–1928.
- (34) Benevides, J. M., Wang, A. H. J., Rich, A., Kyogoku, Y., Vandermarel, G. A., Vanboom, J. H., and Thomas, G. J. (1986) Raman spectral studies of nucleic acids. Raman-spectra of single crystals of R(GCG)D(CGC) and D(CCCCGGGG) as models for A-DNA, their structure transitions in aqueous solutions, and comparison with double-helical poly(DG)-poly(DC). *Biochemistry* 25, 41–50.

- (35) Benevides, J. M., Overman, S. A., and Thomas, G. J. (2005) Raman, polarized Raman and ultraviolet resonance Raman spectroscopy of nucleic acids and their complexes. *J. Raman Spectrosc.* 36, 279–299.
- (36) Deng, H., Bloomfield, V. A., Benevides, J. M., and Thomas, G. J. (1999) Dependence of the Raman signature of genomic B-DNA on nucleotide base sequence. *Biopolymers* 50, 656–666.
- (37) Gleghorn, M. L., Davydova, E. K., Basu, R., Rothman-Denes, L. B., and Murakami, K. S. (2011) X-ray crystal structures elucidate the nucleotidyl transfer reaction of transcript initiation using two nucleotides. *Proc. Natl. Acad. Sci. U.S.A.* 108, 3566–3571.
- (38) Carey, P. R., Ed. (1982) *Biochemical Applications of Raman and Resonance Raman Spectroscopies*, Academic Press, Ottawa, ON.
- (39) Pezolet, M., Pigeon, M., Menard, D., and Caille, J. P. (1988) Raman spectroscopy of cytoplasmic muscle fiber proteins: Orientational order. *Biophys. J.* 53, 319–325.
- (40) Takeuchi, H., Murata, H., and Harada, I. (1988) Interaction of adenosine 5'-triphosphate with Mg<sup>2+</sup>: Vibrational study of coordination sites by use of O-18-labeled triphosphates. *J. Am. Chem. Soc.* 110, 392–397.
- (41) Zheng, R., Zheng, X. J., Dong, J., and Carey, P. R. (2004) Proteins can convert to  $\beta$ -sheet in single crystals. *Protein Sci.* 13, 1288–1294.
- (42) Yu, N. T., Chang, R. C. C., and Huber, J. D. (1974) Single crystal Raman spectra of native insulin: Structures of insulin fibrils, glucagon fibrils, and intact calf lens. *Arch. Biochem. Biophys.* 160, 614–622.
- (43) Bowman, W. D., and Spiro, T. G. (1980) MNDO-MOCIC evaluation of the uracil force-field: Application to the interpretation of flavin vibrational spectra. *J. Chem. Phys.* 73, 5482–5492.
- (44) Chen, Y., Eldho, N. V., Dayie, T. K., and Carey, P. R. (2010) Probing Adenine Rings and Backbone Linkages Using Base Specific Isotope-Edited Raman Spectroscopy: Application to Group II Intron Ribozyme Domain V. *Biochemistry* 49, 3427–3435.
- (45) Espinoza-Herrera, S. J., Gaur, V., Suo, Z., and Carey, P. R. (2013) Following DNA Chain Extension and Protein Conformational Changes in Crystals of a Y-Family DNA Polymerase via Raman Crystallography. *Biochemistry* 52, 4881–4890.
- (46) Petsko, G. A. (1985) Flow cell construction and use. *Methods Enzymol.* 114, 141–146.
- (47) Thomas, G. J., Benevides, J. M., Overman, S. A., Ueda, T., Ushizawa, K., Saitoh, M., and Tsuboi, M. (1995) Polarized Raman spectra of oriented fibers of A-DNA and B-DNA: Anisotropic and isotropic local Raman tensors of base and backbone vibrations. *Biophys. J.* 68, 1073–1088.
- (48) Vassylyev, D. G., Vassylyeva, M. N., Perederina, A., Tahirov, T. H., and Artsimovitch, I. (2007) Structural basis for transcription elongation by bacterial RNA polymerase. *Nature* 448, 157–153.