The active site of multisubunit RNA polymerases (RNAPs) is highly conserved from humans to bacteria. This single site catalyzes both nucleotide addition required for RNA transcript synthesis and excision of incorrect nucleotides after misincorporation as a proofreading mechanism. Phosphoryl transfer and proofreading hydrolysis are controlled in part by a dynamic RNAP component called the trigger loop (TL), which cycles between an unfolded loop and an α-helical hairpin [trigger helices (TH)] required for rapid nucleotide addition. The precise roles of the TL/TH in RNA synthesis and hydrolysis remain unclear. An invariant histidine residue has been proposed to function in the TH form as a general acid in RNA synthesis and as a general base in RNA hydrolysis. The effects of conservative, nonionizable substitutions of the TL histidine (or a neighboring TL arginine conserved in bacteria) have not yet been rigorously tested. Here, we report that glutamine substitutions of these residues, which preserve polar interactions but are incapable of acid–base chemistry, had little effect on either phosphoryl transfer or proofreading hydrolysis by Escherichia coli RNAP. The TL substitutions did, however, affect the backtracking of RNAP necessary for proofreading and potentially the reactivity of the backtracked nucleotide. We describe a unifying model for the function of the RNAP TL, which reconciles available data and our results for representative RNAPs. This model explains diverse effects of the TL basic residues on catalysis through their effects on positioning reactants for phosphoryl transfer and easing barriers to transcript backtracking, rather than as acid–base catalysts.

Significance

Synthesis of new RNA and removal of incorrect nucleotides during proofreading by RNA polymerase involve the transfer of two protons. Here, we show that a polymerase component, the trigger loop, does not directly mediate proton transfer during these reactions, as previously proposed. Instead, the trigger loop plays a central role in transcription as a positional catalyst by orienting the reactants and promoting the polymerase backtracking necessary for proofreading. This positional-catalyst model of trigger-loop function explains its diverse effects on polymerase catalysis and reconciles contradictory reports in the literature. By establishing that the trigger loop is not an acid–base catalyst, our results also guide the search for alternative proton donors and acceptors for reactions in the active site of polymerase.

Tatiana V. Mishanina, Michael Z. Palo, Dhananjaya Nayaka, Rachel A. Mooney, and Robert Landick

*Department of Biochemistry, University of Wisconsin–Madison, Madison, WI 53706

Edited by Jeff W. Roberts, Cornell University, Ithaca, NY, and approved May 18, 2017 (received for review February 13, 2017)
nucleotide addition. Several structurally distinct nucleic acid polymerases in fact appear to use a basic amino acid (lysine) positioned similarly to the TL His as a general acid in nucleotide addition (14), suggesting a potentially conserved mechanism of phosphodiester bond formation across various polymerases. These interactions of the TL His with the β-phosphate oxygens of a matched NTP have been interrogated by molecular dynamics simulations and shown to be significantly stronger when the TL His is protonated than when it is unprotonated (16). Another set of simulations suggested that the TL His aids the exit of pyrophosphate.

Fig. 1. Activities of RNA polymerase. (A) Nucleotide addition, pyrophosphorolysis, and backtracking events in RNAP active site. RNA synthesis (sequence of events at the top) is accomplished in a NAC consisting of four steps. Step 1: translocation of DNA and RNA to position the 3′-OH of RNA in i site on RNAP for the reaction with an incoming NTP; step 2: NTP binding to i + 1 site; step 3: formation of a new phosphodiester bond; and step 4: pyrophosphate release. Template DNA is shown in black, nontemplate DNA in green, RNA in red, incoming NTP in blue, and catalytic magnesium ions in yellow. One position on template/nontemplate DNA is highlighted in purple, to illustrate translocation. Chemical reactions of the reversible phosphoryl transfer (step 3) and hydrolysis of a backtracked RNA are shown in purple and blue boxes, respectively. The flow of electrons in the reactions is marked with arrows. Protons are supplied by the putative general acid ("A") and base ("B"). (B) Basic residues of the trigger helix as candidates for a general acid-base catalyst (PDB ID code 2O5J). The electron flow in the nucleotide addition reaction is indicated with arrows. (C) Structure of a 1-nt backtracked elongation complex (PDB ID code 4WQS), with the trigger loop in a partially unfolded, "bent" conformation (shown in pink). The folded trigger helix (from PDB ID code 2O5J) is shown for reference in orange.
out of the active site following NMP addition, through hydrogen-bonding contacts to pyrophosphate (17). The recently solved crystal structure of an *E. coli* σ transcription initiation complex indeed places the TL His in coordination with both the pyrophosphate and the phosphate of the 3′-terminal nucleotide of RNA (18). In vivo, several TL His substitutions (Ala, Asn, Asp, and Phe) are lethal to yeast, whereas others (Lys, Arg, Trp, 1yr, Gln, and Leu) cause variable growth defects but are viable (19–22). In bacterial RNAPs, only the effects of a TL His-to-Ala substitution have been studied (9, 10, 23), with inconclusive results. Even less is known about the possible involvement of the TL Arg in catalysis.

Similarly to nucleotide addition, factor-independent excision of nucleotides from a backtracked RNA transcript involves the transfer of two protons, but in the reverse direction (blue box in Fig. 1A). The same protein residues (the TL His or Arg) could catalyze acid-base chemistry of this factor-independent RNA hydrolysis, although steric effects preclude their participation in Gre-or transcription factor IIS (TFIIS)-dependent hydrolysis. A body of inconclusive and sometimes contradictory data on the involvement of the TL in the intrinsic cleavage of RNA by RNAP has been reported. Fouque et al. found that a TL deletion (∆TL) in an archaeal RNAP from *Pyrococcus furiosus* does not affect transcript hydrolysis (7). Similarly, in *E. coli* RNAP, neither the ∆TL nor a TL His-to-Ala substitution significantly decreases intrinsic cleavage of a 2-nt backtracked RNA “locked” by the mismatched bases at its 3′ end (8). In contrast, another study found that TL His-to-Ala or ∆TL variants of *E. coli* and *Thermus aquaticus* RNAPs are deficient at hydrolyzing both the RNA with a single and multiple 3′ mismatches (10). Uncertainty about the function of the TL in acid-base catalysis is further exacerbated by crystal structures of backtracked elongation complexes (ECs) in which the TL His or Arg is positioned too far from the relevant nucleotides to function as a general base during hydrolysis (Fig. 1C) (24–26).

An issue with studies of the TL His and Arg residues to date is that most rely on Ala substitutions that do not differentiate between polar and acid–base contributions to catalysis. Furthermore, these drastic changes may alter the available space in RNAP active site and thus complicate the interpretation of the results (see Table S1 for calculated side-chain volumes). To address these concerns and provide more rigorous tests of function, we compared several *E. coli* RNAP mutants: H936A, which removes polar and acid–base contributions; H936Q and R933Q, which preserve the potential for polar interactions but not their acid–base capabilities; and a double mutant (R933Q/H936Q). We tested the effect of these TL substitutions on known RNAP activities (nucleotide addition, pyrophosphorylation, and factor-independent exo- and endonucleolytic hydrolysis). Our results suggest that the TL contributes to RNAP catalysis of these reactions by positioning reactants in the active site, rather than directly donating protons to the reaction.

**Results**

**The TL His Is Not Essential for Efficient Transcript Elongation or Pyrophosphorylation.** To assess the TL His contribution to physiologically relevant processive NMP addition, we developed an experimental method that uses an ~2-kb DNA fragment of the naturally transcribed *E. coli* *rpoB* gene as a transcription template to allow ample time for manual rate measurements. The assay builds on the previously reported scaffold ligation approach (27, 28) and offers the advantage of bypassing initiation at a promoter, often compromised by mutations in the polymerase, and measuring nucleotide addition rates at physiological NTP concentrations (1 mM) and temperature (37 °C) for *E. coli*, which are otherwise too fast to measure manually. The ECs were assembled on a short scaffold (~75 bp) with the polymerase, RNA primer, nontemplate DNA, and a template DNA strand bearing a 5′ phosphate for subsequent ligation (Fig. S1). The ECs were radiolabeled by reaction with [α-32P]GTP, extended to an A26 position by incubation with a subset of NTPs, and then ligated to the 2-kb *rpoB* DNA fragment. Transmission of the ligated template was initiated by addition of 1 mM NTPs and its progress was monitored at various time points at 37 °C (see Materials and Methods and Fig. S1 A and B for additional details). A glutamine (Gln) substitution of the TL His did not affect NMP addition rate across tested buffer pH conditions appropriate for an expected histidine pKₐ (Fig. 2A) (7, 8). The alanine (Ala) mutant, on the other hand, was approximately sixfold slower at pH 8; this could be because of the loss of hydrogen-bonding interactions with the incoming NTP, compromised TL folding into TH necessary for efficient transcription, or both.

A caveat with the processive transcription assay described above is the possibility that during the multiround nucleotide addition, postcatalysis steps other than chemistry (e.g., translocation, pyrophosphate release, conformational changes, and pause escape) may limit the overall rate of elongation. To ensure that such kinetic steps did not mask the effect of the H936Q mutation on chemistry, we measured the rates of single-nucleotide addition in the wild-type and H936Q variant (Fig. 2B). ECs were formed on the same unligated nucleic-acid scaffold as above, labeled with [α-32P]GTP, and then reacted with 1 mM ATP in a quench-flow instrument to follow synthesis of the next two phosphodiester bonds (Materials and Methods). No measurable difference was observed between the ability of the wild-type RNAP and that of the H936Q mutant to extend the RNA (Fig. 2B), consistent with the results of the multiround nucleotide addition experiment.

Phosphoryl transfer (step 3 in Fig. 1A) can occur in reverse at high pyrophosphate concentrations, a process termed pyrophosphorylation, during which the TL His participates in acid–base catalysis of this reverse reaction, we tested the ability of the H936Q polymerase to perform pyrophosphorylation. The sequence of the nucleic-acid scaffold was designed such that the EC would favor the pretranslocated register, placing the 3′ ribonucleotides (GU) in a position from which pyrophosphorylation is possible (Fig. 3A) (29). Indeed, pyrophosphorylation produced a single RNA cleavage product, 1-nt shorter than the starting RNA (Fig. 3B). As evident from the overlay of the wild-type and H936Q RNAP pyrophosphorylation kinetic data (Fig. 3B), the mutation did not compromise the pyrophosphorylation reaction of the polymerase.

The TL His Does Not Serve as a General Base for Intrinsic Cleavage of RNA, but Facilitates RNAP Backtracking. To study the role of the TL in RNA hydrolysis, we used nucleic-acid sequences on which RNAP backtracks following incorporation of cognate nucleotides. Kotlajich et al. identified a pause position on a λPr-*bglF* template at which *E. coli* RNAP backtracked by 4–6 nt (30). We took advantage of this natural backtrack pause sequence to assess the effect of the TL His mutations on the cleavage of multinucleotide backtracked RNA. Briefly, ECs were radiolabeled by incorporation of [α-32P]CMP, elongated to the pause site, and allowed to equilibrate among various backtrack registers. The intrinsic cleavage of backtracked RNA was then initiated with 20 mM Mg²⁺ in pH 9.0 buffer (Materials and Methods and Fig. 4A). Both TL His mutants (Ala and Gln) were capable of hydrolyzing the multinucleotide backtracked RNA, at most twofold slower than the wild-type polymerase (Fig. 4A and Fig. S2B). Although the overall rate of hydrolysis of the starting RNA (C21) was unaffected by the TL His substitutions, the extent of the RNA backtracking changed, as evidenced by the difference in the pattern of cleavage products in the final reactions (Fig. S2B). Specifically, the RNA in the wild-type ECs was backtracked by fewer nucleotides than in the TL His
and further extended in the reaction with 1 mM ATP, to form A20 and generated sticky end on the dsDNA (see a phosphorylated overhang (in green on scaffold 1) complementary to a StyI-denedaturing 8% polyacrylamide gel (Fig. 5). The reactions were quenched with HCl at various time points, and RNA products separated by gel electrophoresis (Fig. S3A). Two radioactive cleavage products were observed in the reactions of the wild-type RNAP, 1 and 2 nt in length, resulting from cleavage of RNA in the pretranslocated (exonucleolytic) and 1-nt backtracked (endonucleolytic) states, respectively (Fig. 4B). The 1-nt product accounted for the majority of the cleaved RNA (60% after 10 min), indicating that at the consensus pause the equilibrium lies toward the pretranslocated EC. The TL His mutants, on the other hand, lacked the 2-nt hydrolysis product, yet still efficiently cleaved 1 nt off of the starting RNA (Fig. 4B). Assuming that 1- and 2-nt cleavage reactions occur by the same mechanism, these observations suggest that the TL His side chain is not required for the intrinsic hydrolysis of RNA, in agreement with the findings described above for the multinucleotide backtracked ECs. Importantly, although the ability of RNAP to cleave RNA was unaffected by the TL His substitutions (as evidenced by the similar rates of the 1-nt cleavage product appearance), the TL His variant polymerases failed to stabilize the 1-nt backtracked state (as evidenced by the lack of the 2-nt product formation with these RNAPs), with implications described further in Discussion, below.

**The TL Arg Is Not the Alternative General Acid–Base Catalyst in Bacterial RNAP.** A TL arginine side chain (β′ R933 in E. coli RNAP), located near the TL His (Fig. 1B) and conserved among bacterial RNAPs, could catalyze acid–base chemistry in place of the His, or potentially rescue the H936Q mutant, provided the TL His and Arg have similar pκ values in the RNAP active site. To evaluate the function of the TL Arg, we tested RNAPs with a Gln substitution of R933 alone, or in combination with the H936Q mutation. The single substitution of R933 with Gln left processive elongation rates largely unaffected across the tested range of buffer pH, whereas the double mutant elongated at about half the wild-type rate at pH 8 and even slower at lower pH (Fig. 5A). These effects of the doubly substituted TL at low pH are readily explained by effects on salt bridges involved in TL folding (SI Discussion). However, the difference between R933Q and R933Q/H936Q RNAPs disappeared in a single-nucleotide addition experiment (Fig. 5B), with both mutants somewhat impaired (two-thirds of the wild-type rate for addition of the first AMP, and approaching half of the wild-type rate for the second AMP incorporation). The discrepancy in the effect of the combined mutations on the processive vs. single-nucleotide transcription at pH 8 could potentially be explained by increased sequence-specific pausing of the R933Q/H936Q mutant on a long DNA template (Fig. 5A). Consistent with the lack of difference in single-nucleotide addition rates, pyrophosphorolysis was not compromised in either R933Q or R933Q/H936Q RNAPs (Fig. 3C).

Similarly to the TL His variants, the TL Arg mutants performed both endo- and exonucleolytic cleavage of RNA (Fig. 6A21).
and Figs. S2C and S3C). Unlike the TL His mutants, the R933Q polymerase was fully capable of achieving the 1-nt backtracked state and of hydrolyzing the resulting backtracked RNA (Fig. 6B). This result suggests the TL Arg does not play a role in proofreading by RNAP, in agreement with the lack of strict conservation of an Arg at this position of the TL, and is consistent with a role of the TL His in 1-nt backtracking.

Discussion

To address the uncertainty about the role of the conserved TL His in phosphoryl transfer and hydrolysis, as well as the contribution of the TL generally to intrinsic transcript cleavage, we studied

Fig. 3. Pyrophosphorylsis by TL mutants. (A) RNA sequence of the nucleic-acid scaffold 2 was designed to promote reconstitution of RNAP in the pretranslocated register from which pyrophosphorylsis of the terminal UMP is possible. (B) Pyrophosphorylsis by the TL H936Q mutant. RNA cleavage data were fit to a single-exponential function. The wild-type and H936Q RNAPs catalyzed pyrophosphorylsis with similar kinetics. Error bars are smaller than data points. (C) Pyrophosphorylsis by the TL Arg mutants. All pyrophosphorylsis reactions were performed in the presence of apyrase, to degrade the released UTP, thereby favoring forward reaction (i.e., cleavage of the starting RNA). Refer to Table S2 for complete nucleic-acid sequences.

Fig. 4. Intrinsic RNA cleavage by TL His mutants. (A) RNA hydrolysis in ECs backtracked by 4-6 nt. The starting RNA in the scaffold 3 was radiolabeled at the 3′-end by incorporation of [α-32P]CMP and extended to the halt position (C21), previously demonstrated to cause RNAP to backtrack by multiple nucleotides (30) (Upper; refer to Table S2 and Fig. S2 for complete sequences of the nucleic acids and Fig. 1 for translocation states). Hydrolysis of the backtracked ECs was initiated with high concentration of Mg2+ in pH 9.0 buffer. Single-exponential kinetics of the C21 RNA disappearance is illustrated in the Lower panel. The TL His substitutions affected hydrolysis only slightly. (B) Exonuclease activity and backtracking. When reconstituted on the shown nucleic-acid scaffold 4 (Upper; see Table S2 for sequences) and walked one position forward to C17, wild-type RNAP partitions between pretranslocated and 1-nt backtracked states, as evidenced by the presence of both the 1- and 2-nt cleavage products in the hydrolysis reactions (Lower and Fig. S3). The TL His mutants preserve the ability to cleave 1 nt off of the RNA, and do so with kinetics similar to the wild-type RNAP, but fail to backtrack, as suggested by the lack of the 2-nt product (open blue and red circles, Lower, and Fig. S3). Thin black lines separate gel portions containing C17 reactant and short cleavage products. See Fig. S3B for an image of the full gel.
RNAPs in which the TL His and Arg residues were substituted with Gln, which preserved the capacity for polar interactions but not acid–base catalysis. We found that the Gln substitutions have little effect on most activities of RNAP. This observation strongly indicates that neither the TL His nor Arg functions as an acid–base catalyst during either phosphorlyl transfer or hydrolysis. The TL His-to-Gln substitution provided additional useful insight into intrinsic RNA cleavage because it inhibited hydrolytic removal of a dinucleotide but not of a single nucleotide from the 3′ end of a consensus-pause RNA. To interpret these findings, we describe below a unified model for TL function as a positional rather than acid–base catalyst. The positional catalyst model of TL function builds on prior descriptions of TL positioning effects and can explain the diverse effects of TL mutants on intrinsic cleavage reported in the literature (7, 8, 10, 33).

The TL/TH may serve as a positional—rather than an acid–base—catalyst in RNAP in two distinct ways during: (i) phosphorlyl transfer and (ii) backtracking and intrinsic RNA hydrolysis. During phosphorlyl transfer, the TH His (and Arg in bacteria) makes contacts to the phosphorylated reactants that are indispensable for optimal positioning of the electrophile (α-phosphate of the incoming NTP during addition or the 3′ phosphodiester bond of RNA during pyrophosphorolysis) for the SN2 attack of the nucleophile (the nascent RNA’s 3′ hydroxyl or a pyrophosphate’s oxygen) (Fig. 7A, Insert). These proposed positioning contacts are well documented structurally (5, 12, 18) and are likely both electrostatic and steric in nature: the TH side chains form salt bridges (i.e., hydrogen bonds and charge–charge...
interactions) with the substrates and push them together to achieve a reactive configuration. Such dual nature of protein–substrate interactions can explain deficient in vitro elongation by the His-to-Ala mutants (Fig. 24; see also ref. 8), and can also explain the robust in vitro activity of E. coli RNAP mutants (Fig. 2) and viability of yeast carrying a His-to-Gln TL mutation (20). The reactive positioning may potentially even be satisfied solely by steric effects, which could explain the recently reported viability of yeast containing a sterically conservative His-to-Leu substitution in the TL (22) (see calculated side-chain volumes in Table S1). Additionally, these contacts may aid folding of the TL into the TH necessary for catalysis. In contrast, a role for the TL residues in proton transfer to or from pyrophosphate in the phosphoryl transfer reaction is not supported by our finding that the single-turnover reactions are uncompromised by Gln substitutions in the TL. At most, such a role must be dispensable rather than obligate and be easily assumed by other proton donors or acceptors in the active site. This conclusion is consistent with previous negative

Mishanina et al. PNAS | Published online June 12, 2017 | E5109

Fig. 7. Positional catalyst model of TL/TH function. The TL/TH serves as a positional, not an acid–base, catalyst in transcription and proofreading by RNAP. During the bond formation step of the nucleotide addition cycle (A), the TL is folded into the active site, thus allowing the His and Arg to orient the NTP phosphates for the substitution reaction through hydrogen-bonding contacts (Inset, PDB ID code 2O5J). Upon incorporation of an incorrect nucleotide, RNAP reverse threads the nucleic-acid template (backtracks) and the TL assumes a partially folded conformation, such that its invariant His can form ionic and hydrogen-bonding contacts with the backtracked RNA nucleotide (B and C). The TL/TH His could then contribute to proofreading RNA hydrolysis by directly positioning the backtracked ribose for excision, thereby lowering the activation barrier for the reaction (B, green dashed line). Alternatively, the interactions with the His could stabilize 3′ nucleotide in the backtracked state (C, blue dashed line) or reduce the energy of the transition states/intermediates leading to that state (C, purple dashed line), consequently promoting RNAP passage into the 1-nt backtracked register. The stabilization of the 1-nt backtracked state in turn eases the path of RNAP to subsequent, multinucleotide backtracked registers.
tests for a role of the TL His in proton-transfer chemistry during nucleotide addition (e.g., ref. 8).

The second way in which the TL functions as a positional catalyst occurs during interruptions to the NAC that lead to backtracking and transcript hydrolysis, which together underlie transcriptional proofreading. The TL can act as a positional catalyst of proofreading RNA hydrolysis in at least three different scenarios. First, the dynamic fluctuations of the TL could shift a backtracked nucleotide into the reactive alignment with water, effectively lowering the activation barrier to hydrolysis (effect 1 in Fig. 7B). This would be analogous to the TL contribution to phosphorly transfer (Fig. 7A), except that the TL would remain in one or multiple partially folded conformations to avoid steric clash with the backtracked nucleotide (Fig. 1C) (26). A role for the TL His in orienting the backtracked nucleotide for cleavage has been proposed previously (10), although without distinguishing between potential TL His functions in positioning vs. as a general base. Crystal structures reveal that a single backtracked nucleotide, such as would arise after misincorporation, is shifted to a different position for hydrolysis upon binding of the GreA/B or TFIIH cleavage factors (24, 26). Thus, it is plausible that a partially folded TL aids transient occupancy of a similarly shifted state required to achieve the trigonal bipyramidal geometry for cleavage in the absence of GreA/B or TFIIH, even though the TL would remain in the backtracked base directly (Fig. 1C) (24, 26).

In the second and third scenarios, the TL could aid entry into or increase occupancy of the initially backtracked state or further backtracked states, either by lowering the barrier to backtracking or by stabilizing a 1-nt backtracked complex (effects 2 and 3, respectively, in Fig. 7C). The polymorphous nature of the TL lends itself to either possibility because the TL could adopt a conformation that positions the TL His, Arg, or other residues to accomplish stabilization of a ground or transition state along the translocation coordinate. Indeed, a molecular dynamics simulation of pyrophosphate release observed transient interactions of the TL His with departing pyrophosphate (17). A similar interaction of the TL His with the 3′-most RNA phosphodiester bond could aid RNA backtracking.

Our results strongly support backtracking assistance by the TL His. We found that Gln substitution for the TL His showed no effect on the rate of intrinsic cleavage of 1 nt from a pretranslocated EC, but eliminated cleavage of 2 nt from the backtracked conformation (Fig. 4B), consistent with the TL His aiding entry into the initial backtracked state (Fig. 7C). Furthermore, the same substitution altered the population distribution of ECs among the 4-, 5-, and 6-nt backtracked states, consistent with preferential stabilization of the earlier, 4-nt backtracked state by the TL His, a preferential stabilization that is lost in the TL His mutants (Fig. S2B). Taken together, these effects of the TL His are most readily explained by the positional catalyst model rather than by direct participation of the TL His in the acid–base chemistry of the cleavage reaction.

The positional catalyst model for the TL function in intrinsic RNA hydrolysis can explain diverse and sometimes seemingly contradictory results published to date. Yuzenkova et al. (10) observed large effects of the TL deletion or Ala substitution of the TL His on RNA cleavage in ECs containing a single mismatched 3′ U or A. These effects could reflect either a role of the TL His in repositioning the mismatched nucleotide into the hydrolysis-competent backtracked state (Fig. 7B) or the TL His favoring the backtracked over the hydrolysis-resistant pretranslocated register (Fig. 7C), because the mismatch could be accommodated in the pretranslocated register by 3′-nt fraying. The short downstream duplex on the scaffolds used in Yuzenkova et al. (10), the downstream end of which lies within the RNAP DNA binding cleft, may perturb the energetic landscape and make the pretranslocated state more accessible to a mismatched 3′ nt. Such TL effects on translocation register bias could be altered by TL His protonation, which may provide an explanation for pH effects on cleavage that depend on the TL His (10), and could also underlie some of the effects of 3′-nt structure on cleavage rates observed on short scaffolds elsewhere (34). In yeast RNAPII, substitution of the TL His with tyrosine (Tyr) stimulates intrinsic transcript cleavage, also in a pH-dependent fashion. Tyr can interact with the RNA bases in multiple modes altered by its deprotonation (planar stacking, edge stacking, hydrogen bonding, and so forth), possibly facilitating positional catalysis by the TL Tyr and explaining the diverse ways this substitution alters transcript cleavage (33). In γ-proteobacteria, the TL contains a large insertion (sequence insertion 3, SI3) (35, 36) that may modulate TL dynamics, thereby affecting positional catalysis of intrinsic RNA cleavage by the TL. Indeed, partial deletion of SI3 or binding of a monoclonal antibody to it inhibited RNA hydrolysis (37), whereas complete removal of SI3 boosted intrinsic cleavage rates of certain mismatched scaffolds (38). Finally, although the TL is generally thought to play minimal role in GreA/B- or TFIIH-mediated transcript cleavage because of steric clash that displaces the TL away from the active site upon cleavage factor binding, we note that TL effects on factor-stimulated transcript hydrolysis are still possible via positional catalysis because the TL could modulate the location of the transcript before binding of the cleavage factors.

Conclusion
The unifying model of the TL as a positional catalyst, rather than an acid–base catalyst, provides a parsimonious explanation for varied effects of TL substitutions on phosphoryl transfer and transcript hydrolysis. Nevertheless, significant questions remain. The participants in proton transfer during phosphoryl transfer and hydrolysis catalyzed by RNAP remain to be identified, and both non-TL RNAP side chains and a metal–ion–activated water molecule are possibilities that remain to be tested rigorously. Furthermore, we cannot exclude the possibility that the roles of TL side chains change in different conditions, are altered upon interaction of transcription factors with the EC, or differ among diverse RNAPs, as demonstrated by the effects of lineage-specific variations in bacterial RNAPs on RNA hydrolysis (38) and the effects of the TL mutations on transcription start site selection in yeast (39). The plasticity of the RNAP active site and the polymorphous nature of the TL afford intriguing opportunities for regulation via positional catalysis that provide fertile ground for future study.

Materials and Methods
Nucleic Acids and Nucleotides. All RNA and DNA oligonucleotides were acquired from Integrated DNA Technologies and were gel-purified on 8 M urea, 8% (for DNA) or 15% (for RNA) polyacrylamide (PA: 19:1 acrylamide:bisacrylamide) gels in 1.25 M Na2EDTA and 44 mM Tris-borate, pH 8.3 (TBE). The gel pieces containing desired oligonucleotides were excised and soaked overnight in buffer (10 mM Tris–pH 7.5, 1 mM EDTA, 100 mM NaCl) to extract the nucleic acids. The extracted oligos were then purified using DEAE Sepharose resin (Bio-Rad) and ethanol-precipitated. For the ligated-scaffolding transcription assay, a 2.3-kb double-stranded DNA (dsDNA) fragment of an rpoB gene from pRL785 plasmid (Table S2) was ligated to the scaffold downstream of RNAP. The fragment was PCR-amplified using forward primer (#10551) (Table S2), coding for a StyI cleavage site, and reverse primer #10242 (Table S2). The amplified fragment was then spermine-precipitated and digested with restriction enzyme StyI to generate a sticky end for ligation to the A26 EC. The longer, 2.3-kb DNA fragment of the digest was spermine-precipitated and stored in TB (20 mM Tris-acetate, pH 8.0, 40 mM potassium acetate, 5 mM magnesium acetate, 1 mM DTT) before use. Unlabeled NTPs were from GE Healthcare Life Sciences. 32P-labeled NTPs were obtained from PerkinElmer.

Proteins. T4 DNA ligase, high-fidelity (HF) Sty, Sds, BsmI, Q5 HF DNA polymerase, T4 polynucleotide kinase, and apyrase were obtained from New England Biolabs. Wild-type E. coli RNA pol was prepared by isopropyl-β-D-thiogalactopyranoside (IPTG)-induced expression from plasmid pRM756 in E. coli strain #224S (Table S2), and purified by polyethylenimine (PEI) and ammonium sulfate precipitation, anion-exchange FPLC on a GE HiTrap HP 5-mL column, and
mL) at 16 °C for 1 h in A26 position by incubation with ATP and GTP (100 μM hybrids of scaffolds 1, 3, and 4, 100 μM restriction enzymes SacII and BsmI. pRL663 and then moved into the pRM756 RNAP overexpression plasmid using KOH, pH 8.0, 50 mM KCl, 5 mM MgCl2, 1 mM DTT, 5% glycerol, 25 μM acetylated BSA), for the transcription and hydrolysis assays, respectively. The scaffold for pyrophosphorolysis assay (scaffold 2 in Fig. 3A) contained 5 μM RNA, 7.5 μM T-DNA, and 9.4 μM nontemplate DNA (NT-DNA), with all three oligonucleotides annealed simultaneously. Each RNAP (1.5 μM) was incubated at 37 °C with the appropriate scaffold (0.5 μM) for 15 min in the corresponding buffer. To complete EC reconstitution on scaffolds 1, 3, and 4 (Table S2), 1.5 μM NT-DNA was added at 37 °C for another 15 min. The ECs (150 nM) with scaffold N 1 (Fig. S1A and B), were incubated with heparin (100 μg/mL) for 10 min and repeatedly washed with RB. Pyrophosphorolysis was commenced by resuspension of the beads in RB containing 5 mM magnesium acetate, 0.5 mM sodium pyrophosphate, and 1.5 units/mL apyrase. Apyrase was included in the assay to break down released UTP, and in doing so to push the pyrophosphorolysis reaction forward, as previously described (29). Reaction progress was monitored by withdrawing aliquots at indicated time points, quenching with 2× stop buffer, and visualizing RNA species on a 20% denaturing PAGE. Decay kinetics of the starting RNA was determined by fitting the quantified data to a single-exponential function.

### Multinucleotide (5–7 nt) Cleavage Assay
Labeled C17 ECs formed on scaffold 3 (Fig. S2A and Table S2) were immobilized on 15-μL Ni beads as described above. The beads were washed with H8 three times to remove unincorporated [α-32P]CTP and then resuspended in 100 μL H8. The “C17” sample was withdrawn, and the remaining buffer was removed. The beads were resuspended in H8 prewarmed to 37 °C and split into two 50-μL samples for the cleavage reaction. Cleavage of each sample was initiated by resuspension of the bead-bound C21 complexes in Cleavage Buffer (CB; 25 mM Tris-HCl pH 9.0, 50 mM KCl, 20 mM MgCl2, 1 mM DTT, 5% glycerol, and 25 μg acetylated BSA per milliliter) preheated to 37 °C. Samples were withdrawn at 3, 6, 10, 20, 40, 60, and 120 min and mixed with 2× stop buffer as described above. Samples were heated, separated in an 8 M urea, 20% PA, TBE gel, and scanned as described above. The amount of the starting C21 RNA and cleavage product (5, 6, and 7 nt in length) bands was quantified with ImageQuant software and used to calculate the fraction of each RNA species for each time point. The C21 fractions were plotted vs. time for each polymere and fitted to a single-exponential decay model using Igor graphing software to obtain rate constants. Time courses from at least three independent cleavage experiments were fitted.

### One- vs. Two-Nucleotide Cleavage Assay
Similarly to the multinucleotide cleavage experiment described above, labeled C17 ECs formed on scaffold 4 (Fig. S3A and Table S2) were immobilized on 15-μL Ni beads. The beads were washed with WB four times to remove unincorporated [α-32P]CTP and resuspended in 105 μL WB, from which the “C17” sample was withdrawn. The C17 EC mixture was split into two 50-μL samples for replicates. Remaining WB was removed and transcript cleavage was initiated by resuspension in 50 μL CB preheated to 37 °C. Samples were withdrawn at 0.5, 1, 2, 5, 10, 30, and 60 min and quenched with 2× stop buffer. Samples were heated, separated in a 25% urea-PAGE gel, and scanned. The amounts of each C17, 2-nt, and 1-nt RNA band were quantified with ImageQuant software and used to obtain rate constants for both the formation of the 1- and 2-nt RNA products, as described above. Time courses from at least three independent cleavage experiments were used in the fitting.

### Side-Chain Volume Calculations
To determine the relative steric bulk of the relevant side chains (Table S1), we calculated their molar volumes with density functional theory (42, 43) using Gaussian 09 (44). Each side chain was truncated from at least three independent cleavage experiments were used in the fitting.

### Acknowledgments
We thank the members of the R.L. laboratory for fruitful discussions and valuable input throughout the course of this work, and Daniel Rostock (University of Wisconsin–Madison) for performing side-chain volume calculations. This work was supported by NIH Grant GM038660 (to R.L.).
Supporting Information

Mishanina et al. 10.1073/pnas.1702383114

SI Discussion

Polymerases bearing Gln substitutions at both the TL His and Arg extended RNA with nearly the wild-type kinetics at alkaline pH (8.0 and 8.4). At acidic pH, however, the R933Q/H936Q RNAP showed impaired transcription (Fig. 5A). This defect could indicate that: (i) the TL Arg indeed rescues the TL His-to-Gln mutant in an acid–base reaction, and vice versa; (ii) at least one ion-pair interaction between the TH His or Arg and NTP phosphates is necessary for optimal NTP positioning; or (iii) at least one ion-pair interaction of the His or Arg with other TL residues, or other parts of RNAP, is required for efficient TL folding and unfolding during processive elongation by bacterial RNAP. A requirement for the formation of an ion-pair involving either the His or Arg (models ii and iii) should be most obvious at lower pH, where the majority of the His or Arg side chains, but not Gln, are protonated (i.e., positively charged) and thus able to form charge–charge interactions. Indeed, the defect in processive elongation caused by the His/Arg double mutation was exacerbated by acidic pH (Fig. 5A). Because the difference from wild-type of the multinucleotide addition rate for the R933Q/H936Q and R933Q enzymes disappeared when a single-nucleotide addition was monitored (Fig. 5B), we favor model iii. The folding of the TL into TH indeed involves the formation and breakage of a series of salt bridges, as observed in the simulated TL–TH transition in pol II (46). In particular, a salt bridge between the TL His and an aspartate (Saccharomyces cerevisiae Rpb1 D1359) stabilizes an open conformation of the TL. The simulations suggest that the transition from an open TL to a closed TH occurs in two steps, with a partially folded TL as an intermediate. Existence of the intermediate conformer raised the possibility that full opening of the TL may not be required during processive elongation, and may only take place during pausing or backtracking. This possibility would explain the pronounced effect of the TL His/Arg double substitution on multiround nucleotide addition, which is clearly accompanied by pauses (see transient RNA bands on transcription gels in Fig. S1C), and not on pause-free single-nucleotide addition. Taken together, our findings argue against the TL His or Arg functioning as a general acid or a base in either transcription or factor-independent proofreading and can be readily explained by roles in ionic interactions.
**Fig. S1.** Ligated-scaffold transcription assay. (A) Nucleic-acid scaffold 1 used in elongation rate measurements. The starting 17mer RNA is italicized and bolded, with the sequence complementary to the template DNA shown in uppercase letters. Positions of halted 26-nt RNA and 66-nt run-off transcript are marked as A26 and C66, respectively. The 4-nt overhang and 3′ phosphate of the template DNA, necessary for ligation, are boxed. (B) Schematic of the experimental set-up for EC reconstitution, labeling, and ligation to a StyI-treated dsDNA fragment of rpoB (see Materials and Methods for details). RNAP and nucleic acids at various stages are represented as a cartoon. (C) Elongation of ligated A26 complexes by the wild-type and mutant RNAPs at pH 8.0. Indicated are the positions of halted A26 complexes, run-off transcript of unligated ECs (C66), and elongation products of successfully ligated ECs. Average transcription speeds were determined as previously described (41).
Fig. S2. Intrinsic RNA cleavage by the TL mutants in ECs backtracked by multiple nucleotides. (A) Nucleic-acid scaffold 3 used in cleavage experiments. The starting 16mer RNA was radiolabeled at 3′-end by incorporation of [α-32P]CMP and extended to the halt position (C21), previously demonstrated to backtrack by multiple nucleotides (30). (B) Cleavage of the C21 RNA by TL His mutants. The backtrack positions giving rise to the cleavage products observed on the gel (5–7 nt in length) are numbered on the cartoon of the EC. (Inset) Distribution of major cleavage products in end-point hydrolysis reactions of the wild-type and TL His mutant polymerases. (C) Cleavage of the C21 RNA by TL Arg mutants. (Inset) Populations of hydrolysis products in end-point reactions of the wild-type and TL Arg variants.
Fig. S3. Endo- and exonuclease activities of the TL variants. (A) Nucleic-acid scaffold 4 used in cleavage experiments. The scaffold is based on the published consensus pause sequence (31), on which a portion of ECs was shown to backtrack by one nucleotide (32). The 16mer RNA was extended to the pause position by incorporation of [\(\alpha^{-32}\)P]CMP, and the resulting C17 ECs allowed to equilibrate between the pretranslocated and 1-nt backtracked registers. RNA hydrolysis was commenced with high pH and Mg\(^{2+}\) concentration. (B) Cleavage of the C17 RNA by TL His mutants. The 1- and 2-nt cleavage products are marked on the gel. (C) Cleavage of the C17 RNA by TL Arg mutants. The translocation state of an EC that gives rise to each cleavage product is shown as a cartoon.

Table S1. Calculated volume of relevant amino acid side chains

<table>
<thead>
<tr>
<th>Side chain identity</th>
<th>Molar volume (cm(^3)/mol)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Histidine</td>
<td>60.6</td>
</tr>
<tr>
<td>Arginine</td>
<td>73.3</td>
</tr>
<tr>
<td>Glutamine</td>
<td>69.8</td>
</tr>
<tr>
<td>Tyrosine</td>
<td>90.2</td>
</tr>
<tr>
<td>Leucine</td>
<td>63.9</td>
</tr>
<tr>
<td>Alanine</td>
<td>31.9</td>
</tr>
</tbody>
</table>
Table S2. Strains, plasmids, and oligonucleotides

<table>
<thead>
<tr>
<th>Stock no.</th>
<th>Name</th>
<th>Description or sequence</th>
<th>Source or Note</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Strains</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>RL2245</td>
<td>DH10B (DE3)</td>
<td><em>Escherichia coli</em> F- endA1 recA1 galE15 galK16 nupG rpsL(StrR) ΔlacX74 araD139 Δ(ara,leu)7697 mcrA Δ(mrr-hsdRMS-mcrBC) ΔompT774::frt Δ(DE3 [lacI, PlacUV5-T7, ind1, sam7, nin5])</td>
<td>Present work</td>
</tr>
<tr>
<td>2956</td>
<td>pRM756</td>
<td>Expresses wild-type <em>E. coli</em> RNAP (α2ββ′ω) with His10 tag on the β′ C terminus</td>
<td>(47)</td>
</tr>
<tr>
<td>1663</td>
<td>pRL663</td>
<td>Encodes rpoC fragment of <em>E. coli</em> RNAP. Used as PCR template for site-directed mutagenesis.</td>
<td>(48)</td>
</tr>
<tr>
<td>1785</td>
<td>pRL785</td>
<td>Encodes rpoB fragment of <em>E. coli</em> RNAP. Used as PCR template for preparation of dsDNA for ligated-scaffold transcription experiment. Derivative of pRL702 (49), with Ncol site directly fused to N terminus of rpoB</td>
<td></td>
</tr>
<tr>
<td><strong>Plasmids</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10160</td>
<td>T-DNA, scaffold 1</td>
<td>p-CAAGGGACTGGTCTGAATCGGGAATACGGAACGGAAGATCTGAGTTCTCTTCCCCTCTAGCTCAGGACGTACTGACC</td>
<td></td>
</tr>
<tr>
<td>10161</td>
<td>NT-DNA, scaffold 1</td>
<td>GGTCAGTACGTCTGAGCTAGAGGGAAGAGAACTCAAGATCTTCCCTGATCGAGCCAGACCTGACC</td>
<td></td>
</tr>
<tr>
<td>6354</td>
<td>NT-DNA, scaffold 2</td>
<td>GGTCAGTACGTCTAAGACCAGCTCATCGATTCAAGACGG</td>
<td></td>
</tr>
<tr>
<td>6355</td>
<td>T-DNA, scaffold 2</td>
<td>CCTGGTCTGAATCGGGAACCGGCGCTTATAGACGTACGTACC</td>
<td></td>
</tr>
<tr>
<td>9580</td>
<td>NT-DNA, scaffold 3</td>
<td>GGTCAGTACGTCCCTAGTGGAACCTTCTCGGTAAGACGCCAGCTGAGTATCCGATG</td>
<td></td>
</tr>
<tr>
<td>9581</td>
<td>T-DNA, scaffold 3</td>
<td>CCACCTGGGAATCTCAAGGGCTTTTACCAGGAAAGGTTCAGATTGACGTACGTGACC</td>
<td></td>
</tr>
<tr>
<td>8334</td>
<td>T-DNA, scaffold 4</td>
<td>CCACCTGGGAATCTGATTTACGGGCAGCACATCGCTGACGTAATCGACC</td>
<td></td>
</tr>
<tr>
<td>9563</td>
<td>NT-DNA, scaffold 4</td>
<td>GGTCAGTACGTCCCGGATATCCCACTGCAATGGACGTACGTGACC</td>
<td></td>
</tr>
<tr>
<td>8855</td>
<td>RNA, scaffold 1</td>
<td>UUUUUUUGACCGGAG</td>
<td></td>
</tr>
<tr>
<td>11001</td>
<td>RNA, scaffold 2</td>
<td>UUUUUUGCCGGGG</td>
<td></td>
</tr>
<tr>
<td>9594</td>
<td>RNA, scaffold 3</td>
<td>UUUUUUGCGGGAAG</td>
<td></td>
</tr>
<tr>
<td>8401</td>
<td>RNA, scaffold 4</td>
<td>UUUUUUGCCGAAG</td>
<td></td>
</tr>
<tr>
<td><strong>Oligonucleotides (5′→3′)</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>10242</td>
<td>Rev (PCR)</td>
<td>TCGGATACGAGGATGGAATCG</td>
<td></td>
</tr>
<tr>
<td>10551</td>
<td>For (PCR)</td>
<td>CTACCGCCTTGGCGAACCGG</td>
<td></td>
</tr>
<tr>
<td>10746</td>
<td>R933Q (SDM)</td>
<td>GGAGTACGAGGATGGAATCG</td>
<td></td>
</tr>
<tr>
<td>10747</td>
<td>H936Q (SDM)</td>
<td>GGAGTACGAGGATGGAATCG</td>
<td></td>
</tr>
</tbody>
</table>

Mishanina et al. www.pnas.org/cgi/content/short/1702383114