Structural principles of transcriptional collisions

John W. Watters[1,2], Andreas U. Mueller[3], Xiangwu Ju[1], Stephany J. Chuquimarca[1], Hannah J. Ye[1], Seth A. Darst[3], Gregory M. Alushin[2,*], Shixin Liu[1,*,#]

[1]Laboratory of Nanoscale Biophysics and Biochemistry, The Rockefeller University, New York, NY, USA
[2]Laboratory of Structural Biophysics and Mechanobiology, The Rockefeller University, New York, NY, USA
[3]Laboratory of Molecular Biophysics, The Rockefeller University, New York, NY, USA
*Correspondence: galushin@rockefeller.edu & shixinliu@rockefeller.edu
[#]Lead contact

Abstract

RNA polymerase (RNAP) must navigate crowded genomic tracks to robustly produce transcripts. DNA-bound proteins act as transcription-impeding roadblocks, which are eventually overcome by RNAPs through unclear mechanisms. Here, we use cryo-electron microscopy to visualize actively transcribing *E. coli* RNAP upon collision with an inactivated restriction enzyme (EcoRI*) or with another converging RNAP. Both collisions induce RNAP backtracking into an inactive swiveled state. Swiveling is coupled to DNA deformation through a characteristic structural landscape, mediating RNAP pause stabilization across distinct collision geometries. EcoRI* roadblock bypass efficiency is impacted by factors that modulate RNAP swiveling, backtracking rescue, and roadblock stability. In comparison, head-on RNAP-RNAP collisions feature substantial heterogeneity suggestive of sustained dynamics, with variable inter-RNAP distances that are modulated by nascent transcript hairpins which position bidirectional termination sites. By resolving, to our knowledge, novel structures of actively transcribing RNAP undergoing collisions, we provide a mechanistic framework for interpreting mechanical conflicts during transcription.

Introduction

The genome has been described as a molecular highway where the macromolecular machines that carry out fundamental central dogma processes (e.g. replication and transcription) function in a densely crowded environment[1]. This milieu includes a variety of static roadblocks (e.g. DNA-binding proteins and DNA lesions), signals (regulatory sequences), and other dynamic machines operating on the same substrate, such that collisions *in vivo* are inevitable[1,2]. In addition to interfering with machine function[2], collisions also provide potential for regulation, for example, by defining genomic boundaries or coupling the activities of multiple machines in tandem[3,4], thereby enhancing biological robustness.

Recent studies have emphasized that multiple categories of collisions between DNA machines can have both positive and negative biological consequences. For example, head-on collisions between replication and transcription machinery lead to increased R-loop formation, which is detrimental to genome stability[5–7]. Conversely, local chromatin looping into topologically associated domains is driven by cohesin-CTCF collisions, mediating functional genome organization[8]. Furthermore, co-directional collisions between machines can enhance their processivity, such as translation-transcription coupling in bacteria, which prevents transcriptional arrest[9–15]. Additionally, we recently showed that sequence-specific pausing of bacterial RNA polymerases (RNAPs) positions RNAP-RNAP head-on collisions, thereby programing transcription termination sites to enhance the production of precise transcripts from convergent genes[16,17]. How collisions both disrupt and enhance machine functions at the protein structural level to produce these diverse outcomes remains unknown.

Because of its central role in many genomic collisions, here we focus on the RNAP as a model machine, which operates as a Brownian rachet and can generate strong forces and torques[18–23]. Previous biophysical studies of RNAP collisions have suggested that upon being halted by roadblocks, the complex can bypass them by two mechanisms: (i) passively waiting for roadblocks to dissociate or (ii) actively consuming chemical energy to dislodge them[24], with the relative contributions of these mechanisms dependent on roadblock identity[15,16,24,25]. Additional factors including transcription elongation factors, co-directional transcribing RNAPs, and applied force have also been shown to help RNAP overcome roadblocks[15,16,24]. While RNAP has been extensively characterized by single-molecule biophysics, structural biology, and biochemistry, it remains unclear how roadblocks effectively stall actively transcribing RNAPs and how RNAPs in turn bypass roadblocks. Collisions occur under non-equilibrium, actively energy-dissipating conditions are technically challenging to recapitulate for structural studies. Conceptually, it is interesting to examine how collisions involving a nanoscale machine like RNAP—which operates under the low Reynolds number regime—compare to macroscopic collisions such as a car crash, which produces heterogeneous outcomes in the final motor configuration. However, such microscopic transcriptional collisions have rarely been directly visualized at the atomic level.

In this study, we use cryo-electron microscopy (cryo-EM) to examine actively transcribing *E. coli* RNAPs collided with (i) a static roadblock, namely the catalytically deficient EcoRI E112Q restriction enzyme mutant (hereafter referred to as EcoRI*), and (ii) another RNAP transcribing in the opposite direction, resulting in a head-on collision. By comparing the structural landscapes of these two physiological collision

scenarios[24,26,27], we uncover principles for mechanical control of RNAP that can be harnessed to bypass deleterious roadblocks and mediate functional positioning of transcription termination sites.

Results

Reconstituting RNAP collision with a stationary roadblock

To capture a collided roadblock complex, we generated a nucleic acid scaffold containing a pre-formed transcription bubble and loaded the scaffold with an RNAP core enzyme and a downstream EcoRI* dimer bound to the cognate binding site (5'-GAATTC-3', K_D ~2.5 fM)[26,27]. Upon ribonucleotide (NTP) addition, RNAP is anticipated to transcribe for ~20 nucleotides (nt) before encountering the EcoRI* roadblock (**Fig. 1A**). Proper loading of RNAP and EcoRI* was confirmed via a native gel-shift assay, and the complex was enriched using gel filtration chromatography (**Fig. 1B, S1**). We used a Cy3-labeled RNA in the scaffold to evaluate the length of transcription products. In the absence of EcoRI*, most RNAPs transcribed until the end of the DNA template, generating run-off transcripts ~70 nt in length (**Fig. 1C**). When EcoRI* was added, shorter RNAs (narrowly distributed around ~40 nt in length) were observed with minimal run-off products (**Fig. 1C**), suggesting high occupancy of EcoRI* and efficient blockage of RNAP readthrough.

We next used the RNAP-EcoRI* assembly supplemented with 1.5 mM Fos-Choline-8 detergent for cryo-EM sample preparation. To induce collision, 0.5 mM NTPs were added and incubated for 60 seconds prior to plunge-freezing (**Fig. 1B**). After data collection, we performed 3D classification to enrich for particles that contained both RNAP and EcoRI*. Unexpectedly, we obtained a single three-dimensional (3D) reconstruction of the RNAP-EcoRI* collided complex with a global Fourier shell correlation (FSC) resolution of 2.68 Å, indicating that we captured a single metastable state instead of multiple states along the collision pathway (**Fig. S2, S3A-D**). The resolution of the downstream EcoRI* roadblock was significantly lower than the RNAP core, suggesting higher flexibility[26,28,29]. We used 3DFlex in cryoSPARC to improve the resolution of the flexible downstream DNA and EcoRI*, producing a reconstruction with a global resolution of 2.83 Å that allowed us to build an atomic model of the entire collided complex (**Fig. 1D, S3E-H, Table S2**)[30].

EcoRI-collided RNAP is backtracked*

We envisioned three possible scenarios for the behavior of RNAP upon colliding with a roadblock: (1) it could advance to the final position along the DNA template free from steric hindrance and remain there; (2) it could halt before this final position due to topological constraints imposed by the DNA; or (3) it could recoil away from the final position to an energetically more favorable site, due to the ability of RNAP to backtrack along DNA[31,32], as suggested by previous studies of RNAP transcription in the presence of an EcoRI* roadblock[15,24]. To assess RNAP backtracking in our system, we took advantage of the fact that backtracked RNA is sensitive to GreA/B-stimulated endonucleolytic cleavage[33]. In the absence of GreA/B, the majority of RNA products were 43 nt long (**Fig. 2A),** while in the presence of GreB and, to a lesser extent, GreA, the distribution of bands was shifted towards shorter lengths. In the presence of both factors, most RNA products were 39-41 nt in length (**Fig. 2A, S4A, B**), suggesting that the collided

RNAP is backtracked 2-4 nt in our reconstitution system, in agreement with previous results[15].

During each nucleotide addition cycle, conformational changes in the bridge helix and trigger loop—structural elements in the RNAP active site—are coordinated to translocate the DNA register, facilitate nucleotide binding, and accelerate the catalysis of nucleotide incorporation[34]. We next examined the RNAP active site in our cryo-EM map and found that it resembles previously reported structures of backtracked RNAP (PDB: 6RI9, 6RIP), featuring an unfolded trigger loop and a straight bridge helix (**Fig. 2B**)[35,36]. Moreover, we observed densities in the secondary channel consistent with extruded, backtracked RNA (**Fig. 2B**). Two backtracked nucleotides are resolved; the lack of density from additional backtracked nucleotides is likely attributable to higher levels of conformational flexibility, as has been observed in previous cryo-EM structures of backtracked RNAP[35].

These biochemical and structural data suggest that the RNAP transcribes 19 nt of RNA from its initial position, but it cannot stably reside at that position, potentially due to steric clashes with the roadblock. After collision, RNAP backtracks by as many as 4 nt, allowing it to re-equilibrate at a more favorable position. To further evaluate this hypothesis, we computationally modeled the collision pathway by moving RNAP structures to each nucleotide position along the collision axis (up to 44 nt, one nt past the final position transcribed by RNAP in the EcoRI* collision scenario). We then measured the overlapping volume enclosed by the solvent-excluded surfaces (SES) of RNAP and EcoRI* as a proxy of steric overlap (**Fig. S4C, D**). The RNAP conformation was modeled based on a previously reported elongation complex (EC) structure (PDB: 6RH3)[35]. When the polymerase was placed less than two bases away from the collision interface, the overlapping SES volume[37] increases substantially, consistent with clashes occurring. While this analysis is limited by the assumption that the RNAP retains the same conformation as it approaches the roadblock, it nonetheless indicates that direct steric hindrance from EcoRI* is a plausible explanation for why RNAP backtracks away from the collision site.

EcoRI-collided RNAP is swiveled*

We next examined the RNAP structural modules that influence its catalytic activity. We find that in the collided RNAP-EcoRI* complex, RNAP is in a 'swiveled' conformation, in which a swivel module (β subunit: AA 1241-1341; β' subunit: AA 1-342, 369-420, 787-930, 1135-1375) is rotated relative to the RNAP structural core. When RNAP is sufficiently swiveled, the bridge helix kinks and the trigger loop is unable to fold into the correct conformation to mediate catalysis, stabilizing long-lived transcriptional pausing[38–40]. In the active elongation complex (EC; PDB: 6rh3)[35], RNAP is in a non-swiveled state, whereas pronounced swiveling (~4.5-6° swivel module rotation relative to the active EC) is observed in various paused and backtracked ECs[35,36,39,41].

In the collided RNAP-EcoRI* complex, the RNAP adopts a moderately swiveled state with a swivel module rotation of 2.8° relative to the active EC, a lower magnitude rotation than observed in a previously reported hairpin-stabilized paused EC (4.4°, PDB: 6ASX)[39] (**Fig. 2C, D**). We then analyzed 3DFlex variability trajectories to assess swiveling heterogeneity in our data. We found that the swivel angle ranged between 2.6° and 3.8° across frames (**Fig. 2D**)[30]. In contrast, previously resolved backtracked RNAPs populate

168 either swiveled (3.8°, 63% of population, PDB: 6RI9) or non-swiveled states (0.7°, 37%
169 of population, PDB: 6RIP), but not in between (**Fig. 2D**)[35]. This result indicates that the
170 collided RNAP is constrained within a moderately swiveled state, suggesting a novel
171 configuration imposed by the physical collision.
172 EcoRI binding to DNA by itself is known to distort DNA from its canonical B-form
173 at the cognate sequence, introducing multiple kinks that result in a net 25° bending of the
174 DNA and concomitant increased accessibility of bases in the major groove[28,42]. Given the
175 ability of DNA to serve as a mechanical transducer, we hypothesized that DNA distortions
176 could be transmitted to RNAP as a mechanism to indicate the presence of a roadblock.
177 To investigate coupling between DNA deformation and RNAP structural remodeling, we
178 once again analyzed 3DFlex trajectories. To assess changes in DNA positioning relative
179 to the active EC (PDB: 6RH3), a reasonable proxy for the spatial orientation of DNA
180 relative to the RNAP structural core during catalysis, we established a frame of reference
181 by subjecting the downstream nucleotides to principal component analysis (PCA). The
182 displacement of downstream DNA in the RNAP-EcoRI* complex was then measured from
183 projections onto the PC1 and PC2 orthoplanes after superimposing the structural cores
184 of the RNAPs. Interestingly, we observed that specific DNA distortions downstream from
185 the RNAP active site are correlated with changes in the RNAP swiveling angle (**Fig. 2E,**
186 **S5, Movie S1**). As the downstream DNA is displaced more along PC2, the magnitude of
187 RNAP swiveling increases (**Fig. 2F, S5**). This observation suggests that modulation of
188 DNA substrate geometry can couple inhibitory RNAP rearrangements to the presence of
189 roadblocks at collision sites.
190
191 *Factors that modulate RNAP backtracking / swiveling impact bypass efficiency*
192 Because RNAP is backtracked and swiveled in the RNAP-EcoRI* complex, we
193 next examined whether the outcome of this collision (persistent RNAP halting versus
194 roadblock bypass) can be biased by factors that modulate these aspects of the RNAP
195 structural landscape. We implemented a bulk biochemical assay to quantitatively evaluate
196 roadblocking efficiency (E_{RB}), conducting transcription reactions at 37 °C for 30 minutes
197 with increasing salt concentrations (50 to 500 mM KOAc) to permit measurable levels of
198 bypass. As anticipated, higher levels of roadblock bypass (lower E_{RB}) were observed at
199 higher salt concentrations, presumably due to destabilization of the EcoRI*-DNA
200 interaction (**Fig. 3A, B**). Addition of NusA (a pro-swiveling factor) and NusG (an anti-
201 swiveling, anti-backtracking factor)[36,43,44] led to decreased and increased bypass,
202 respectively (normalized E_{RB} = 1.2 ± 0.1 for NusA and 0.8 ± 0.1 for NusG in the presence
203 of EcoRI*, compared to 1.0 ± 0.1 for EcoRI* alone) (**Fig. 3C**). These results suggest that
204 biochemically modulating RNAP's swiveling propensity (e.g. through Nus factor
205 engagement) changes its capacity to overcome roadblocks. Strikingly, the presence of
206 the transcript cleavage factor GreB, which rescues backtracked RNAP[33], dramatically
207 reduced roadblocking efficiency (normalized E_{RB} = 0.06 ± 0.03) (**Fig. 3B, C**). This near-
208 complete roadblock bypass can be explained by repeated collisions between RNAP and
209 EcoRI* made possible by GreB-induced transcription restart after RNAP backtracking,
210 consistent with a recently proposed 'battering' model[24].
211
212 *Roadblock structural remodeling facilitates bypass*

213 Having observed substantial structural alterations in RNAP upon collision, we next
214 examined whether the structure of the EcoRI* roadblock is also impacted. EcoRI*
215 engages DNA as a homodimer[28,42,45]. The protein's DNA-bound conformation is overall
216 highly similar between the RNAP-EcoRI* collided complex and a previous crystal
217 structure in the absence of RNAP (PDB: 1ERI)[42], with the notable exception for the N-
218 terminal helix (AA 20-31) of the protomer proximal to the collision interface, which flips
219 ~150°, orienting it away from RNAP (**Fig. 3D, Movie S2**). This rearrangement was not
220 observed in the distal EcoRI* protomer, suggesting it is specifically linked to contact with
221 RNAP (**Fig. 3D**). We speculate that it occurs at the collision site to reduce steric clashes
222 with RNAP, becoming kinetically trapped as the RNAP backtracks away. Additionally, this
223 rearrangement repositions EcoRI*'s N-terminal extension (NTE) residues (5K, 9R, 15K)
224 in an orientation where they can access DNA (**Fig. 3D**). These amino acids have been
225 previously shown to enhance DNA binding in biochemical assays, presumably by
226 mediating electrostatic interactions[26,28,42,45].
227 In the EcoRI* crystal structure, the N-terminal helix forms hydrophobic contacts
228 with the protein's structural core[42]. As a protein's potency as a roadblock intuitively
229 depends on its mechanical resilience, we posited that destabilizing the overall folded
230 structure of EcoRI* would compromise its roadblocking efficacy. To test this hypothesis,
231 we introduced alanine substitutions at the positions of hydrophobic residues within the
232 EcoRI* core (**Fig. 3E**), distal from the DNA binding interface, and assessed their impact
233 on E_{RB} using our roadblock assay. We found that all point mutations (with the exception
234 of V166A) led to significantly increased RNAP bypass (decreased E_{RB}) (**Fig. 3F**). Notably,
235 the two mutations that caused the largest magnitude changes, I21A and F252A, are
236 internally facing residues adjacent to the predicted collision interface that are likely to
237 mediate a stabilizing contact between the N-terminal helix and the EcoRI* core (**Fig. 3E,**
238 **F**). These results suggest that the structural stability of a roadblock protein—independent
239 from its affinity to DNA—is linked to its capacity to resist physical impacts from motors
240 like RNAP; correspondingly, mechanical destabilization of roadblocks can enhance
241 bypass.
242
243 *Reconstituting head-on collisions between two RNAPs*
244 To explore the generality of RNAP's collision response, we next investigated
245 another physiological scenario, when an RNAP runs into another active motor complex
246 rather than a stationary roadblock. In previous studies, we demonstrated that head-on
247 collisions between two RNAPs produce long-lived complexes on DNA[16], constituting a
248 prevalent mechanism for transcription termination of convergent genes in *E. coli*[17]. To
249 reconstitute a head-on RNAP-RNAP collision in a format suitable for cryo-EM structural
250 interrogation, we constructed a DNA template with two convergent promoters (**Fig. 4A**).
251 Each promoter was bound by a σ^{70}-RNAP holoenzyme that was walked to the +18/+20
252 position by omitting one of the four nucleotides (CTP), followed by purification via size
253 exclusion chromatography. Transcription was then restarted with addition of 1 mM NTPs
254 to induce a collision (**Fig. 4B**). We also introduced a native sequence from the *E. coli*
255 genomic region between the *YnaJ-UspE* gene pair into the template, which encodes an
256 RNA hairpin containing an 8-bp stem. We have previously shown that this sequence is
257 important for positioning the collision site between a pair of RNAPs, mediating precise
258 termination of convergent transcription[16,17].

After biochemically confirming the formation of a head-on RNAP-RNAP collided complex (**Fig. 4C**), we employed this DNA template (once again supplemented with 1.5 mM Fos-Choline-8 detergent) for cryo-EM sample preparation. Initial 2D class averages displayed features consistent with two RNAPs in close proximity (**Fig. S6**). Unlike the RNAP-EcoRI* collided complex, 3D classification of the RNAP-RNAP collided complex revealed substantial heterogeneity. Notably, the distance between the two RNAP molecules varied, with the best resolved classes featuring 26, 29, and 31 bp of DNA between the active sites of the polymerases (**Fig. 4D-F)**. Because of the ~10.5-bp helical pitch of B-form DNA, these three configurations all feature approximate C2 symmetry, a plausible explanation for why both RNAPs could be effectively aligned and averaged when separated by these particular distances. We exploited this C2 pseudosymmetry to perform symmetry expansion and analyzed all RNAP protomers from a given inter-RNAP distance class within a single frame of reference, facilitating high-resolution refinement (26 bp, 2.65 Å; 29 bp, 2.87 Å; 31 bp, 2.62 Å global resolution) (**Fig. S7**).

We could not interpret nucleic acid sequences in the density maps (which were modeled as homopolymers), presumably due to heterogeneity in where collisions occur along the DNA template. An additional potential source of heterogeneity in our preparation is the presence of σ^{70} to promote initiation, but the 31-bp class was the only one that featured weak density (only visible at low map contour levels) for the σ^{70} domain 2 bound to RNAP. Therefore, the variable inter-RNAP spacing we recovered through classification likely reflects inherent heterogeneity in RNAP-RNAP collisions. We once again employed 3DFlex to further improve the resolution of these classes and analyzed their conformational landscapes (**Fig. S7**).

The RNAP-RNAP interface varies considerably among the three classes due to the differential translational and rotational relative positions of the RNAP pair on the helical DNA template. Nevertheless, in all three classes, the clamp domain of one RNAP was engaged with different domains of the other collided RNAP, notably the β lobe/SI2 domain (26 bp class), SI3 domain (29 bp class), and jaw domain (31 bp class) (**Fig. 4G-I**). We reasoned this could influence RNAP swiveling, as the clamp domain is a component of the swivel module. Indeed, all of the RNAPs in head-on collided complexes adopted a swiveled conformation (26 bp, 3.9°; 29 bp, 2.3°; 31 bp, 4.0° relative to PDB: 6RH3) (**Fig. 4J-L**). Additionally, despite sequence ambiguity, all the RNAPs showed clear density for backtracked RNA in their secondary channel regardless of inter-RNAP distance (**Fig. 4M-O**). Thus, despite substantial differences at the collision interface, RNAP-RNAP collisions populate a similar conformational landscape as the RNAP-EcoRI* collision, suggesting a fundamental RNAP response to roadblocks.

RNA hairpin impacts collision positioning and stability

To evaluate the bulk transcriptional output of the paired RNAPs in this system, which also serves as a proxy for their positioning along the DNA, we performed in vitro SEnd-seq (simultaneous 5' and 3' end RNA sequencing) to systematically profile end-to-end RNA products[17,46]. End-biotinylated DNA templates were used in these experiments to pull down collided complexes, enriching for RNAs that remained bound to RNAP. We also constructed a template without the hairpin-encoding sequence for comparison. The sequenced RNAs displayed homogeneous 5’ ends (corresponding to initiation from either promoter) and more heterogeneous 3’ ends (**Fig. 5A, B, S8**), as anticipated given the

305 stochastic nature of RNAP restart[47], which will impact the occurrence and location of
306 collisions. Samples including the hairpin-encoding sequence in the template featured a
307 pronounced drop in the RNA length profile at the predicted hairpin sequence boundaries
308 (**Fig. 5A**), resulting in significantly shorter transcripts for both the forward and reverse
309 directions compared to the no-hairpin template (**Fig. 5C, S8**). These results indicate that
310 the hairpin fosters consistent positioning of RNAP-RNAP collisions, mitigating
311 heterogeneity attributable to the stochastic mechanochemistry of these nanoscale
312 machines.
313 We observed that the complex disassembles when incubated for longer time
314 periods, which could occur when one RNAP is able to overcome the other and push it off
315 the DNA template. Given the observation that all RNAPs in the collided complex were
316 backtracked, we reasoned that the hairpin could serve as a barrier against long-range
317 backtracking, thereby stabilizing the complex. To test this hypothesis, we performed a
318 time course assay to measure the half-life of the collided complex by halting the
319 transcription reaction with an EDTA quench at different timepoints post NTP addition.
320 While the collided complex on the template with the hairpin remained stable for hours ($T_{1/2}$
321 = 150 ± 40 min), the collided complex on the template without the hairpin disassembled
322 within minutes ($T_{1/2}$ = 6 ± 3 min) (**Fig. 5D, E**), confirming the hairpin's contribution to
323 complex stability.
324 We next sought to examine how the presence of a hairpin in the nascent transcript,
325 which would form if one of the RNAPs has transcribed past the hairpin-encoding
326 sequence before collision, influences the structure of the collided complex. No clear RNA
327 hairpin density was observed in the exit RNA channel for any of the three RNAP-RNAP
328 complex classes, indicative of flexibility or variable occupancy. We then reprocessed the
329 data, re-extracting particles using a smaller box size encompassing one RNAP followed
330 by focused classification at the RNA exit channel (**Fig. S9**). In a single class obtained
331 using this approach (representing ~7% of the data), we observed hairpin density (**Fig.
332 5F**). RNAP was still backtracked and swiveled (~3.4°) in this class, suggesting that the
333 presence of a well-ordered hairpin does not grossly alter RNAP's conformational
334 response to collision (**Fig. 5G**). To assess the hairpin's impact on collision geometry, we
335 recentered our particle picks to capture the other collided polymerase, allowing us to
336 analyze the distribution of inter-RNAP spacings for particles contributing to the hairpin-
337 containing class versus the distribution for the entire collided population. In the hairpin-
338 containing particles, we found the class with the largest separation between RNAPs (31
339 bp) was mildly enriched (~1.3-fold increase), while the classes with shorter inter-RNAP
340 spacing were less populated, with a lower 26-bp class abundance (~1.5-fold decrease)
341 and a complete absence of the 29-bp class (**Fig. 5H**). These data are consistent with the
342 notion that the hairpin reduces the heterogeneity of RNAP-RNAP collisions (**Fig. 5H**,
343 $p<0.001$), potentially by restricting the dynamics of the hairpin-associated RNAP and
344 maintaining its position on DNA.
345
346 *DNA deformation-RNAP swivel coupling depends on inter-RNAP distance*
347 As a physical connection between RNAP and roadblock through the template DNA
348 is a common element of transcriptional collisions with distinct composition and
349 geometries, we reasoned this linkage could mediate characteristic RNAP remodeling. To
350 test this hypothesis, we examined the coupling between DNA deformation and RNAP

351 swiveling across RNAP-RNAP collision configurations. Swiveling was present in all three
352 head-on collision orientations. However, interestingly, the degree of RNAP swiveling
353 varied between the different classes in a manner that was non-monotonic with inter-RNAP
354 distance, as the 29-bp spaced class featuring a substantially lower swivel angle in the
355 consensus class (~2.3°) than the 26-bp (~3.9°) and 31-bp (~4.0°) classes (**Fig. 4J-L**). We
356 speculate this variation is linked to the differential interfaces between the RNAPs, which
357 likely impact the local mechanical environment. We also found that the degree of DNA
358 deformation downstream of the collided RNAP varied among classes (**Fig. 6A**). Finally,
359 we examined whether the RNAP swivel angle and the DNA deformation remain correlated
360 across different collision scenarios. When compared alongside previous RNAP structures
361 representative of active EC, backtracked, and swiveled complexes (PDB: 6RH3, 6ASX,
362 6RI9, 6RIP)[35,39], changes in the DNA PC1-axis did not correlate with changes in swiveling
363 while changes in the DNA PC2-axis displayed a strong correlation, with the displacement
364 increasing linearly with the magnitude of the RNAP swivel angle (**Fig. 6B, C**). Additionally,
365 the extent of PC2 displacement within individual 3DFlex variability trajectories also
366 correlates with the magnitude of swivel angle change (**Fig. 6D, S10, Movie S1, 3**)[30].
367 These analyses highlight that while RNAP's conformational landscape varies across
368 collisions, likely due to variations in the details of the RNAP-roadblock interface restricting
369 the accessible DNA bending landscape, DNA deformation serves as a common
370 mechanism to transduce collisions into RNAP swiveling and inactivation.
371
372
373 **Discussion**
374 In this study, we visualized the structural dynamics of RNAP upon collisions with
375 both a passive (EcoRI*) and an active (another RNAP) roadblock under non-equilibrium,
376 actively transcribing conditions. Despite variability in the composition and geometry of
377 these collisions, our results reveal a characteristic RNAP response consisting of
378 backtracking and swiveling to enter a transcriptionally inactive state (**Fig. 7A**). Our data
379 highlight how shared mechanical elements across collisions, including roadblock steric
380 hindrance and DNA deformation, can coordinate this common RNAP response
381 irrespective of the detailed stereochemistry of a collision.
382 As we did not directly visualize RNAP at its site of closest contact with roadblocks
383 (indicated from transcript lengths), we cannot comment on how backtracking is triggered.
384 We nevertheless speculate that roadblock steric hindrance ultimately limits forward
385 progress of RNAP, triggering backtracking to a metastable position where clashes are
386 moderated. At this position, RNAP dynamics can be modulated by coupling between DNA
387 template deformation and RNAP swiveling, facilitating the maintenance of an inactivated
388 state. Consistent with a previously proposed battering model[15,24], we find that factors
389 which reduce RNAP swiveling and enhance transcription restart after backtracking
390 enhance roadblock bypass (**Fig. 3B-C, 7B**). Our structures also rationalize the previous
391 observation that applying small opposing forces (~0.2 pN) counterintuitively increases the
392 rate of EcoRI* bypass by RNAP[24], as this likely promotes RNAP un-swiveling by altering
393 the downstream DNA deformation profile to more efficiently restart elongation after
394 backtracked RNA cleavage, leading to eventual bypass. We furthermore uncovered an
395 additional layer of the battering mechanism, wherein RNAP can physically deform
396 roadblocks to eventually displace them (**Fig. 3D-F, 7B**). This suggests mechanical

397 destabilization of deleterious roadblocks plays an instrumental role in their clearance from
398 the genome.

399 Our observation that RNAPs in head-on collisions are also backtracked and
400 swiveled, but feature variable inter-RNAP distances, is compatible with a scenario
401 featuring intermittent, ping-ponging restart dynamics between the two RNAPs. We
402 propose that when one RNAP stochastically restarts, it backtracks upon colliding with the
403 other paused RNAP. This clears the path for either RNAP to once again restart,
404 generating cycles of randomly alternating transcription and backtracking (**Fig. 7C**). This
405 model can explain our previous observation that both RNAPs can transcribe through
406 bidirectional convergent terminators like YnaJ-UspE before termination[17], which is
407 otherwise incompatible with a static collision. Our analysis also provides a mechanistic
408 explanation for the positioning effect of the RNA hairpin when bound to RNAP, as we find
409 its presence is correlated with increased complex stability (**Fig. 5D-E**) and reduced
410 heterogeneity in inter-RNAP spacing (**Fig. 5H**). This suggests that the hairpin likely
411 suppresses the motions of the engaged RNAP, thereby stabilizing it in a particular region
412 along the DNA template and attenuating the variability in transcript length. We did still
413 observe substantial variability in transcripts even in the presence of the hairpin in our
414 minimal reconstitution system (**Fig. 5A-C**), suggesting additional cellular factors may
415 further enhance the robustness of this mechanism, an important topic for future studies.

416 This work establishes the feasibility of scrutinizing genomic collisions using cryo-
417 EM, providing a path towards visualizing physiological scenarios of increasing complexity.
418 An important consideration for future studies is the topology of the template DNA, which
419 may be addressable by employing closed templates with varying degrees of
420 supercoiling[48,50]. Additionally, collisions featuring additional machines, such as co-
421 directionally transcribing RNAP and translating ribosomes (both of which can promote
422 roadblock bypass[3,10,15]) or head-on/co-directional DNA replication forks (replication-
423 transcription conflict as a primary source of genome instability[4–6]) are likely accessible by
424 similar approaches. Such studies will be critical for a mechanistic understanding of how
425 mechanical conflicts are managed across the crowded genome and functionally
426 harnessed to enhance the robustness of gene expression.

427 **Methods**

428 **Protein purification**
429 *E. coli RNA Polymerase core*
430 RNAP was expressed and purified as described previously[49], with an additional step of
431 3C cleavage of the β'-3C-His10 tag and a reverse Nickel-IMAC (immobilized metal affinity
432 chromatography) column to isolate RNAP core without the His10 tag.
433
434 *Maltose Binding Protein (MBP)-EcoRI* (E112Q) and additional mutants*
435 EcoRI* E112Q (a gift from Dr. Ruben Gonzalez's lab) was cloned into a pET21 vector
436 with an N-terminal His10-MBP-3C sequence using Gibson assembly. The plasmid was
437 transformed into competent Eco BL21(DE3) by heat shock. The cells were grown in the
438 presence of 50 µg/mL kanamycin to an optical density (OD) at 600 nm of ~0.5-0.6 in a 37
439 °C shaker. Expression was induced with addition of 0.5 mM Isopropyl *β*-D-1-
440 thiogalactopyranoside (IPTG) at 16 °C overnight. Cells were harvested by centrifugation,
441 resuspended in lysis buffer [20 mM potassium phosphate, 300 mM sodium chloride
442 (NaCl), 1 mM dithiothreitol (DTT), 0.1 mM ethylenediaminetetraacetic acid (EDTA), 5%
443 glycerol (v/v), 0.15% Triton X-100, 1x protease inhibitor cocktail (Sigma-Aldrich), pH 7.5
444 at room temperature], and flash frozen in liquid nitrogen prior to storage at -80 °C.
445 Cell pellets were thawed, then lysed with an Avestin Emulsiflex C5 homogenizer.
446 The lysate was then clarified (2x centrifugation steps, 30 minutes each, 48,000 x *g*) and
447 applied to a 5 mL MBPTrap HP column (Cytiva). The column was washed with MBP
448 Buffer A [20 mM potassium phosphate, 300 mM NaCl, 1 mM DTT, 0.1 mM EDTA, 5%
449 glycerol (v/v), 0.15% Triton X-100, pH 7.5 at room temperature] and eluted with MBP
450 Buffer A supplemented with 10 mM maltose. Peak fraction were pooled, diluted to in
451 Heparin Buffer A [20 mM potassium phosphate, 1 mM DTT, 0.1 mM EDTA, 5% glycerol
452 (v/v), 0.15% Triton X-100, pH 7.5 at room temperature] to 50 mM NaCl final concentration
453 and loaded onto a 10 mL Heparin HP (Cytvia) column equilibrated in 10% Heparin Buffer
454 A/90% Heparin Buffer B [20 mM potassium phosphate, 1 M NaCl,1 mM DTT, 0.1 mM
455 EDTA, 5% glycerol (v/v), 0.15% Triton X-100, pH 7.5 at room temperature]. The bound
456 protein was eluted with a gradient of 10-100% Heparin Buffer B over 10 CVs. Peak
457 fractions of protein without nucleic acid contamination were eluted in a peak centered at
458 ~45 mS/cm conductivity and concentrated with a 30 kDa molecular weight cut-off
459 (MWCO) concentrator (Amicon). Concentrated protein was clarified (5 minutes, 21,000 x
460 *g*) before application onto a HiLoad 16/600 Superdex (Cytiva) gel filtration column
461 equilibrated with gel filtration buffer [20 mM potassium phosphate, 300 mM NaCl, 1 mM
462 tris(2-carboxyethyl)phosphine (TCEP), 0.1 mM EDTA, 5% glycerol (v/v), 0.15% Triton X-
463 100, pH 7.5 at room temperature]. Peak fractions corresponding to the EcoRI* dimer were
464 pooled, concentrated in a 30 kDa MWCO concentrator (Amicon), diluted with gel filtration
465 buffer supplemented with 50% glycerol to a final concentration of 20% glycerol, flash
466 frozen in liquid nitrogen, and stored at −80°C until use. Protein concentration was
467 estimated using a Bradford assay[51] (Bio-Rad Protein Assay Dye Reagent Concentrate)
468 with bovine serum albumin standards.
469 For single point mutants (I21A, L108A, V166A, F252A, I258A), the same protocol
470 was used with the following substitutions: 1 mM DTT was replaced in all buffers with 1
471 mM TCEP, and a Superdex 200 Increase (Cytiva) column was used for gel filtration
472 instead of a HiLoad 16/600 Superdex (Cytiva) column.

473
474 *GreA, GreB, NusA, NusG, σ^{70}*
475 GreA, GreB, NusA, NusG, and σ^{70} were expressed, purified was performed as described
476 previously[16,49,52,53].
477 Endonucleolytic cleavage activity of GreA/B was verified as described previously[16].
478 In brief, a backtracked RNAP EC scaffold (nucleotide sequences in **Table S1**) was formed
479 with a short (2 nt RNA backtracked) and long backtracked EC (6 nt RNA backtracked).
480 Oligonucleotides were high-performance liquid chromatography (HPLC) purified (IDT).
481 Activity assays were performed at 37 °C for 10 minutes by adding GreB or GreA at a final
482 concentration of 600 nM in the absence or presence of NTPs. The reactions were stopped
483 by addition of an equal volume of 2x RNA loading buffer and incubation at 90 °C for 2
484 minutes. Products were separated on a 25% Urea-PAGE gel, followed by fluorescence
485 visualization on a Typhoon FLA 7000 scanner. Full-length RNA products indicative of
486 endonucleolytic cleavage and transcriptional restart were enhanced in the presence of
487 GreA (short backtracked EC) and GreB (long backtracked EC) as expected[33].
488 As reported previously, NusA and NusG were verified to be active by bulk in vitro
489 transcription assays in transcription buffer (25 mM Tris-hydrochloride pH 7.5, 150 mM
490 potassium chloride, 10 mM magnesium chloride, 1mM DTT). In brief, Rho and NusG in
491 combination were shown to enhance Rho mediated termination efficiency on the Rho-
492 dependent λtR1terminator[53]. NusA reduced transcriptional readthrough across the ynaJ-
493 uspE bidirectional terminator[54]. σ^{70} was verified to be active in promoting T7A1 and
494 LambdaPR dependent promoter complex formation and transcription in in vitro
495 transcription assays in a transcription buffer (25 mM Tris-HCl pH 7.5, 150 mM KCl, 10
496 mM $MgCl_2$, 1mM DTT)[55].
497
498 **RNAP-EcoRI* Scaffold Complex Assembly**
499 DNA oligonucleotide sequences (Polyacrylamide gel electrophoresis (PAGE)-purified
500 from IDT) and a 5' Cy3-labeled RNA oligonucleotide (Dharmacon) were designed using
501 a previously used scaffold as a starting point (sequences of template, non-template, and
502 RNA oligos in **Table S1**)[56]. Template DNA and RNA were annealed in a 1:1 ratio in 1x
503 IDT annealing buffer (100 mM potassium acetate (KOAc); 30 mM HEPES, pH 7.5) for 5
504 minutes at 95º C, 5 minutes at 75º C, 5 minutes at 45º C, and 40 cycles of 1º C steps (1
505 minute hold each step, final temperature 4º C) in a thermocycler. RNAP core desalted
506 into scaffolding buffer (20 mM Tris pH 7.5 at RT, 150 mM potassium glutamate (KGlu), 5
507 mM magnesium acetate ($Mg(OAc)_2$), 2.5 mM DTT) was then incorporated and incubated
508 for 15 minutes at RT. 1.5-fold excess of non-template DNA was then added to the
509 complex mixture and incubated for 10 minutes at RT. 5-fold excess of EcoRI* was then
510 added to the mixture (this step was excluded for the -EcoRI* controls) and incubated for
511 10 minutes at RT before addition of 0.1 mg/ml final concentration of heparin sodium salt
512 to remove non-specifically bound RNAP and EcoRI* proteins. The fully assembled
513 complex was purified away from incompletely assembled complexes via gel filtration
514 chromatography using a Superdex 200 Increase column equilibrated with low salt buffer
515 (20 mM Tris-acetate (Tris-OAc) pH 7.8 at RT, 50 mM KOAc, 5 mM $Mg(OAc)_2$, 2.5 mM
516 DTT) containing 0.1 mg/ml heparin (Sigma-Aldrich) to prevent re-binding of non-
517 specifically loaded RNAP/EcoRI* to the DNA template during the purification. Complex

assembly was monitored via native gel electrophoresis on a 5% polyacrylamide 0.5X TBE gel.

RNAP Head-On Complex Assembly

DNA constructs featuring either two RNAP promoters (T7A1 and λPR) or only one promoter were generated by PCR, purified via gel filtration chromatography on a Superdex 200 Increase column equilibrated with DNA gel filtration buffer (20 mM Tris pH 8.0 at RT, 200 mM NaCl, 0.1 mM EDTA), and concentrated into Tris Low EDTA buffer (10 mM Tris pH 8.0 at RT, 0.1 mM EDTA) for long term storage. PCR oligos and full DNA sequences of the templates are in **Table S1**.

RNAP holoenzyme was formed by incubating σ^{70} with RNAP core for 20 minutes at 37º C in their purification buffers, using a 3:1 ratio of σ^{70}:RNAP core. The complex was then desalted into transcription buffer (25 mM Tris-HCl pH 7.5 at RT, 150 mM KCl, 10 mM $MgCl_2$, 1mM DTT), and its concentration measured using A280 and its calculated extinction coefficient. To form open promoter complexes, 4-fold excess of holoenzyme was added to the DNA template (final concentration ~1-2 μM DNA) and incubated for 15 minutes at 37º C. Non-specifically bound RNAPs were removed by incubation with 0.1 mg/mL heparin for 10 minutes at 37º C. 0.5 mM ATP, UTP, and GTP were then added to the reaction mixture and incubated for 15 minutes at 37º C to walk the promoter bound RNAPs out to the dual CTP nucleotide deprivation stall site. The fully assembled complex was purified away from incompletely assembled complexes via gel filtration chromatography using a Superdex 200 Increase (Cytiva) column equilibrated with low salt buffer (20 mM Tris-OAc pH 7.8 at RT, 50 mM KOAc, 5 mM $Mg(OAc)_2$, 2.5 mM DTT) containing 0.1 mg/ml heparin to prevent re-binding of non-specifically loaded RNAP to the DNA template during the purification. Complex assembly was monitored via native gel electrophoresis on a 2% agarose 0.5X TBE gel.

Cryo-EM

Sample Preparation

After gel filtration, the complex was concentrated using Amicon 100 kDa concentrators and washed 7x (450 μL added per wash) into low salt buffer without heparin to dilute any remaining heparin still in solution to a negligible amount. The complex was then diluted with low salt transcription buffer supplemented with Fos-Choline-8 (Anatrace, final concentration 1.5 mM). 3 μL of sample was incubated with 0.5 μl of 7x NTP stock solution (3.5 mM NTPs in low salt transcription buffer plus 1.5 mM Fos-Choline 8) for one minute (EcoRI* complex) or 30 seconds (RNAP-RNAP) at room temperature before vitrification. Samples were prepared on glow-discharged (Gatan Solarus, O_2/H_2, 10 seconds) C-Flat Holey Carbon Film CF-1.2/1.3-4, 400 mesh gold grids at concentrations ranging from 10-14 μM for RNAP-EcoRI* and ~4 μM for RNAP-RNAP. Grids were loaded into Leica EM GP apparatus, blotted for 4-5 seconds at 90% humidity at room temperature, and plunge frozen in liquid ethane.

For the RNAP-EcoRI* complex, the sample concentration was estimated by the A550 signal (Cy3 extinction coefficient of 150000 M^{-1} cm^{-1}). For the RNAP-RNAP complex, sample concentration was estimated using A260 and A280 absorbance measurements using combined extinction coefficients from the individual components.

Cryo-EM data acquisition and processing

For RNAP-EcoRI* complex, three different grids were imaged from two different sample preparations of the complex, frozen on different days. Datasets were collected on the same FEI Titan Krios G3i system operating at 300 kV and equipped with a Gatan K3 direct electron detector, BioQuantum energy filter, and a spherical aberration corrector using super-resolution mode. Frame sequences ("movies") were recorded using the SerialEM software[57] suite at a nominal magnification of ×81,000, corresponding to a calibrated pixel size of 0.86 Å at the specimen level (super-resolution pixel size of 0.43 Å per pixel)[58]. Each 1.4 s exposure was dose-fractionated across 35 frames, with a total electron dose of 47.3 e^{-} $Å^{-2}$ (1.35 e^{-} $Å^{-2}$ per frame), with defocus values ranging from −0.8 to −2.0 µm underfocus. Beam-image shift was used to collect 43,378 single exposures (4.6k, 17.8k, 23.8k exposures from the different grids) from nine holes in a 3-by-3 grid per each stage translation.

The RNAP-RNAP complex dataset was collected similarly, with the following alterations. Data were collected on a single grid on a FEI Titan Krios G2 operating at 300 kV and equipped with a Gatan K3 direct electron detector and BioQuantum energy filter. Movies were recorded at a nominal magnification of x105,000, corresponding to a calibrated pixel size of 0.847 Å at the specimen level (super-resolution pixel size of 0.4235 Å per pixel)[58]. The total electron dose was 48.79 e^{-} $Å^{-2}$ (1.39 e^{-} $Å^{-2}$ per frame). Beam-image shift was used to collect 20,080 single exposures.

Cryo-EM data processing

All processing was performed using cryoSPARC[59]. Movies were aligned with Patch Motion correction (B-factor: 150), and contrast transfer function (CTF) parameters were estimated with Patch CTF using default settings. For each dataset, blob picker was initially used to pick particles from a 256-pixel box, followed by 2D and 3D classification on extracted particles to remove junk particles and RNAPs that did not restart. This initial particle stack was used to generate a preliminary reconstruction that was used to re-pick particles using a template picker approach to yield more particles. All of the picked particles were then re-extracted with a 400-pixel box and separated with multiple rounds of 3D heterogeneous refinement job with candidate 3D volumes as well as junk particle volumes (**Fig. S2, S6**). All models were then further refined via a non-uniform refinement job[60], local CTF refinement, and global CTF refinement. Then low per-particle scale particles were removed, followed by per-particle reference-based motion correction and subsequent local refinement. 3DFlex[30] was then performed (N=5 modes) to improve the density of high flexible regions in the particles and generate variability trajectories for DNA deformation-RNAP swiveling analysis.

For the RNAP-RNAP dataset, the following additional steps were performed. During test dataset processing a 400-pixel box size was used and binned by 4 to enhance processing speed during initial processing steps. Particle box size was increased to 700 pixels (no binning) to accommodate the additional RNAP. After heterogeneous refinement, an additional round of 3D classification was performed using low-resolution volumes as bait volumes (generated by manual docking to mimic RNAPs with 26-36 bp of separation followed by molmap (resolution 20) volume generation from atomic models) using ChimeraX to see if any additional states could be extracted. This resulted in the detection of the 29 bp inter-RNAP class. After non-uniform refinement, models were re-

oriented along their C2 symmetry axis and symmetry expanded before local refinement and further polishing with a one RNAP sized mask.

To isolate particles that had clear hairpin density in the exit channel, a new set of particles were extracted from the same micrographs (to remove any potential bias from the C2 symmetric particles that were able to be resolved to high resolution) and subjected to a similar processing pipeline to a resolve a single RNAP class (box size 400 pixels) using a focus mask corresponding to a single RNAP (**Fig. S9**). This class contained downstream DNA and low-resolution signal further downstream that could correspond to collided RNAP density being averaged between different states (since the second RNAP was not included in the focus mask). Taken together, this indicates that these single RNAPs are still in a post-collision state instead of the alternate hypothesis such as run-off RNAPs or non-bound RNAPs. These particles were subjected to polishing and refinement prior to an additional round of focused classification performed with a mask around the RNAP exit channel that showed one class with clear RNA hairpin density. Particles contributing to the hairpin-containing class were subjected to final polishing and refinement. Additionally, to compare the collided states of the hairpin-containing RNAPs, hairpin-containing particles were re-extracted with a larger box size (700 pixels) and subjected to 3D classification with a focus mask excluding the high-resolution hairpin-containing RNAP (N=8 classes). The relative distribution of classes representing either the 26, 29, 31 bp inter-RNAP distances, or other classes consisting of only a single well-refined RNAP was then determined via manual inspection of the density maps and comparison to the 26, 29, and 31 bp maps constructed prior.

Model building and refinement

The RNAP-EcoRI* initial model was derived from PDB 6RIP[35]. The model was manually fit into the cryo-EM density maps using ChimeraX[37], modified in Coot/PyMOL/Isolde[61–63], and followed by real-space refinement using PHENIX[64]. The nucleic acid sequence was fixed in the same reference frame as the EcoRI* cognate binding sequence, which matched the cryo-EM map along the rest of the nucleic acid sequence. For real-space refinement, rigid body refinement was followed by all-atom refinement with Ramachandran and secondary structure restraints. For the RNAP-RNAP complexes, the same refinement methodology was applied with the following modifications. The RNAP coordinates from the RNAP-EcoRI* complex were used as the starting model for the RNAP-RNAP complexes. Additionally, since we were unable to determine the nucleic acid sequence by visual inspection in any of the RNAP-RNAP states (likely due to heterogeneity in the positioning of collision sites along the template), poly-G and poly-C sequences were used during modeling building.

3DFlex model fitting individual frames

For each frame of the 3DFlex variability trajectories, docking of individual rigid bodies (as defined in **Table S3**) from the consensus model was performed in ChimeraX[37] followed by subsequent unsupervised all-atom refinement in Isolde (5 minutes, 20K, nucleotides restrained by base-pair hydrogen bonding)[63].

Swivel domain motion calculation

655 The core modules of the different RNAP structures were superimposed in PyMOL by
656 aligning the following the α-carbons: all modeled residues from the α subunit, residues
657 3:27, 142:152, 445:455, 520:713, 786:828 from the β subunit, residues 343:368, 412:524,
658 530:552, 569:701, 720:786 from the β' subunit, and all modeled residues from the ω
659 subunit.
660 RNAP EC (PDB: 6RH3)[35] was used as the reference structure (swivel angle = 0°)
661 for swivel angle measurements. To determine the swivel angle for a given RNAP
662 structure, residues comprising the swivel module from the RNAP structure and the
663 reference were selected (β-subunit AA 1241:1341; β'-subunit AA 1:342, 369:420,
664 787:930, 1135:1375). The rotation axis and rotation angle transforming one swivel
665 module into the other were then computed using the PyMOL script draw_rotation_axis.py
666 (https://pymolwiki.org/index.php/RotationAxis). For DNA deformation-RNAP swiveling
667 coupling analysis, the same method was applied. For symmetry expanded particles in the
668 head-on collision state, the swivel angle of only the RNAP contained in the refinement
669 mask was analyzed.
670
671 *DNA displacement calculation*
672 Generation of a frame of reference was necessary to orient downstream DNA
673 displacement on different axes. To accomplish this, the center of masses of twelve
674 downstream nucleotides of the active RNAP EC (PDB: 6rh3)[35] were extracted using
675 PyMOL[62]. PCA was performed on their 3D coordinates to generate the PC1 and PC2
676 orthoplanes along the long axis of the downstream DNA. Since this definition of PC1 and
677 PC2 is due to inherent variation of the center of mass of the nucleotides downstream of
678 the active site in the reference frame RNAP EC (PDB: 6rh3)[35], the plane definition is
679 arbitrary, meaning that the definition of which axis is most responsible for variance (PC1)
680 is descriptive of the RNAP reference structure variability, not descriptive of the variance
681 between the different structures being compared in this study. For each query structure
682 (consensus structures and 3DFlex frames), the structural cores of the RNAPs were
683 aligned to the structural core of the active RNAP EC (PDB: 6rh3)[35], and displacement of
684 the center of mass of each nucleotide was then projected onto PC1 and PC2, summed
685 across the sequence of twelve nucleotides, and averaged.
686
687 **Biochemical Assays**
688 *RNAP-EcoRI* Collided Complex*
689 *Backtracking assays*
690 5' Cy3-labeled RNA oligonucleotides corresponding to zero to four nucleotides of
691 backtracking (based on the RNAP starting position along the nucleic acid sequence) were
692 synthesized (Dharmacon) and used as a ladder to verify the RNA length. For the
693 backtracking assay, the scaffold was assembled as described previously except for
694 omitting the EcoRI* binding step before gel filtration purification. Twenty-fold excess of
695 EcoRI* (if applicable) and 1 μM Gre factors (if applicable) were added to the assembled
696 complex and incubated at RT for 20 minutes before addition of 0.5 mM NTPs for 1 minute
697 in a final volume of 15 μL. To process the RNA samples, 1.5 μL of Turbo DNase
698 (ThermoScientific) was added and incubated for 15 minutes at 37º C, 1.5 μL of proteinase
699 K (New England Biolabs) and 50 mM EDTA for 15 minutes at RT, and finally an equal
700 volume of 2x RNA loading dye (1x components - 47.5% formamide, 0.01% SDS, 0.01%

701 bromophenol blue, 0.5 mM EDTA) was added to the samples. Samples were boiled at
702 95º C for 5 minutes and separated on an 18% TBE-Urea PAGE gel and visualized using
703 Cy3-fluorescence on a Typhoon gel imager. Three biological replicates of this assay were
704 performed on different days.
705
706 *EcoRI* bypass assays*
707 RNAP-EcoRI* containing scaffolds were purified as described previously. Samples were
708 then diluted into heparin-free buffer to a final concentration of ~200 nM as quantified by
709 Cy3 RNA fluorescence. Samples were then diluted two-fold into a mixture of low salt
710 transcription buffer supplemented with varying KOAc concentrations and 0.5 mM NTPs.
711 These reactions were incubated at 37º C for 30 minutes and then quenched by addition
712 of 2X RNA loading dye (95% formamide, 0.02% SDS, 0.01% bromophenol blue, 1 mM
713 EDTA) supplemented with KOAc to equalize the final salt concentration between
714 conditions. For analysis, the intensity of the roadblock and scaffold band were
715 background subtracted, measured in FIJI[65], and normalized relative to each other
716 (roadblock intensity divided by scaffold intensity, and additionally rescaled for each salt
717 titration gel so that the 50 mM normalized roadblock intensity was set to E_{RB} = 1, and the
718 background intensity was set to E_{RB} = 0). E_{RB} calculated at 300 mM KOAc (because that
719 salt concentration had the largest dynamic range when all the factor addition experiments
720 were compared) was rescaled so that the average value of the EcoRI* ("normalized E_{RB}")
721 was equivalent to 1 for each sample condition (factor addition experiments in **Fig. 3C** and
722 stability mutants experiments in **Fig. 3F**) to adjust for observed batch effects when
723 different NTP stocks were used.
724
725 *RNAP Head-On Collided Complex*
726 *Complex stability assay*
727 RNAP-RNAP head-on complexes were assembled and purified for DNA templates with
728 and without the YnaJ-UspE hairpin as previously described in low salt transcription buffer
729 supplemented with 0.1 mg/ml heparin. Samples were concentrated using a 100 kDa
730 MWCO concentrator (Amicon). Control samples to measure total DNA amount loaded
731 per well were generated by treating samples with 1 µL proteinase K (New England
732 Biolabs) per 10 µL reaction volume supplemented with 25 mM EDTA (final concentration)
733 for over one hour at room temperature. 1 mM NTPs were added to initiate transcription
734 restart, and aliquots of the reaction were quenched with addition of 25 mM EDTA (final
735 concentration) at the specified time interval. Samples were run on a 2% 0.5X TBE
736 agarose gel after being diluted with a 6x stock of 2M sucrose, and the resulting gel was
737 stained with SYBR safe to visualize the DNA. Lane profiles were extracted using a 0.5x
738 lane width window. Signal was summed in the 400-1000 bp range, then normalized
739 between the -NTP control (corresponding to 100% RNAP-RNAP complexes) and the
740 proteinase K control (corresponding to background signal) to calculate fraction intact
741 complex. Four different experimental replicates were performed. Bi-exponentials were
742 fitted through the data using the scipy python module[66] to approximate the half-life of the
743 collided complex, and 95% confidence intervals are displayed in the plot.
744
745 *SEnd-seq*

RNAP-RNAP head-on complexes were assembled and purified as previously described in low salt transcription buffer supplemented with 0.1 mg/ml heparin. For single promoter templates, the same general protocol was followed with a polymerase:DNA ratio of 2:1 instead of 4:1 as used for the double promoter templates. Samples were concentrated using a 100 kDa MWCO concentrator (Amicon). +PK labeled samples were treated with proteinase K (New England Biolabs) and 25 mM EDTA for over one hour at room temperature. 1 mM NTPs were added and incubated at RT for 1 minute before quenching with 25 mM EDTA. For complex pulldown, samples were washed and incubated with streptavidin M-280 Dynabeads (Thermo Fisher Scientific) before being pulled down via a magnetic separation rack (Permagen), then washed twice with 200 µL low salt buffer supplemented with 25 mM EDTA and 0.1 mg/mL heparin. The complexes were then treated with proteinase K for 1 hour at 37 °C to elute RNAs off the beads followed by overnight incubation at 4°C. The reaction was purified using 2.2x vol of Agencourt RNAClean XP beads (Beckman Coulter, A63987). After elution, nucleic acids were treated with 0.5 µl TURBO DNase (Life Technologies, AM2238) to remove template DNA. The remaining RNA was further purified via phenol-chloroform-isoamyl alcohol (Invitrogen, 15593031) extraction (25:24:1, v/v/v) and recovered by ethanol precipitation. The RNA pellet was resuspended, mixed with 1 µL of spike-in RNA, and subsequently ligated to a 3' RNA adapter for SEnd-seq library preparation, as previously described[17,46]. Sequencing was performed on an Illumina MiSeq instrument in a paired-end mode (150 nt x2).

Reads were barcode sorted to the correct template and experimental condition. The separate 5' and 3' end reads of each RNA sequence were aligned to both the forward and reverse candidate RNA sequence using a local pairwise aligner in the Biopython module[67] (match_score=2, mismatch_score=-1, open_gap_score=-5, extend_gap_score=-0.5). Alignment scores were calculated by the summation of the two best candidate positions. The higher scoring alignment was taken, as long as the distance between the 5' and 3' ends was greater than 15 nucleotides, as all the RNAs should be at least 17 nucleotides from the A/G/UTP only transcription initiation walkout before subsequent purification. Unmapped reads either had the wrong direction of 5' and 3' alignment or were too short. Furthermore, RNAs that did not initiate from within a 10-nucleotide range near the transcription start site were filtered out. These lost reads upon filtering could come from RNAP initiation from the end of the DNA template, RNA degradation of the 5' end, or non-templated addition at the 3' end[68].

Molecular graphics and data analysis

Molecular graphics were prepared with UCSF ChimeraX[37]. Plotting and statistical analysis were performed with Python using the matplotlib module[69]. Python codes were prepared with the assistance of ChatGPT 4.0.

Statistical analysis

Conditions were compared via two-tailed, unpaired t-test including Welch's correction. If applicable, a Benjamini-Hochberg correction for multiple hypothesis testing was added. S.D. stands for standard deviation and S.E. stands for standard error when used in the text. P-value asterisk reported as *$p < 0.05$, **$p < 0.01$, ***$p < 0.001$.

Acknowledgements

We thank Johanna Sotiris, Honkit Ng, and Mark Ebrahim at the Rockefeller University Evelyn Gruss Lipper Cryo-EM Resource Center for their support in data collection. The expression plasmids for Rho, NusA, NusG, GreA and GreB were provided by Rachel Anne Mooney (Landick Laboratory, University of Wisconsin-Madison). NusA, NusG, GreA, and GreB proteins were purified by Rui Gong (Alushin Laboratory, Rockefeller University). S.L. acknowledges support from the National Institute of Health (R01GM149862). S.L. and G.M.A acknowledge support from the Stavros Niarchos Foundation Institute for Global Infectious Disease Research at the Rockefeller University and a joint award from the Alfred P. Sloan Foundation Matter-to-Life Program (G-2024-22709).

Author contributions

J.W.W., S.L., and G.M.A. designed research. J.W.W. purified proteins, collected cryo-EM datasets, and performed biochemical studies. A.U.M. purified proteins and helped with biochemical experimental design. A.U.M. and S.A.D. helped with cryo-EM processing and data interpretation. X.J. performed SEnd-Seq experiments. S.J.C. performed biochemical assays for the RNAP-EcoRI* complex. H.J.Y. performed biochemical assays for the RNAP-RNAP collided complex and collected the RNAP-RNAP collided cryo-EM dataset. J.W.W., S.L., and G.M.A. wrote the paper with input from all other authors.

Data availability

Cryo-EM density maps and corresponding atomic models have been deposited in the PDB and EMDB with the following accession codes: RNAP-EcoRI* Collision (PDB: XXXX, EMDB: YYYYY); Head-on RNAP-RNAP Collision: 26 bp distance (PDB: XXXX, EMDB: YYYYY); Head-on RNAP-RNAP Collision: 29 bp distance (PDB: XXXX, EMDB: YYYYY); Head-on RNAP-RNAP Collision: 31 bp distance (PDB: XXXX, EMDB: YYYYY); Head-on RNAP-RNAP Collision: RNA hairpin (PDB: XXXX, EMDB: YYYYY). 3DFlex volumes and fit models for both consensus and variability trajectory movies have been deposited in zenodo (doi: 10.5281/zenodo.19317100).

Competing interests

The authors declare no competing interests.

References

1. Le, T.T., and Wang, M.D. (2018). Molecular Highways—Navigating Collisions of DNA Motor Proteins. J. Mol. Biol. *430*, 4513–4524. https://doi.org/10.1016/j.jmb.2018.08.006.
2. Sneppen, K., Dodd, I.B., Shearwin, K.E., Palmer, A.C., Schubert, R.A., Callen, B.P., and Egan, J.B. (2005). A Mathematical Model for Transcriptional Interference by RNA Polymerase Traffic in Escherichia coli. J. Mol. Biol. *346*, 399–409. https://doi.org/10.1016/j.jmb.2004.11.075.
3. Klumpp, S., and Hwa, T. (2008). Stochasticity and traffic jams in the transcription of ribosomal RNA: Intriguing role of termination and antitermination. Proc. Natl. Acad. Sci. U. S. A. *105*, 18159–18164. https://doi.org/10.1073/pnas.0806084105.
4. Helmrich, A., Ballarino, M., and Tora, L. (2011). Collisions between Replication and Transcription Complexes Cause Common Fragile Site Instability at the Longest Human Genes. Mol. Cell *44*, 966–977. https://doi.org/10.1016/j.molcel.2011.10.013.
5. Sankar, T.S., Wastuwidyaningtyas, B.D., Dong, Y., Lewis, S.A., and Wang, J.D. (2016). The nature of mutations induced by replication–transcription collisions. Nature 2016 535:7610 *535*, 178–181. https://doi.org/10.1038/nature18316.
6. Prado, F., and Aguilera, A. (2005). Impairment of replication fork progression mediates RNA polII transcription-associated recombination. EMBO Journal *24*, 1267–1276. https://doi.org/10.1038/sj.emboj.7600602.
7. Pomerantz, R.T., and O'Donnell, M. (2008). The replisome uses mRNA as a primer after colliding with RNA polymerase. Nature *456*, 762. https://doi.org/10.1038/nature07527.
8. Zhang, H., Shi, Z., Banigan, E.J., Kim, Y., Yu, H., Bai, X.-C., and Finkelstein Correspondence, I.J. (2023). CTCF and R-loops are boundaries of cohesin-mediated DNA looping. Mol. Cell *83*, 2856-2871.e8. https://doi.org/10.1016/j.molcel.2023.07.006.
9. Epshtein, V., and Nudler, E. (2003). Cooperation between RNA polymerase molecules in transcription elongation. Science (1979). *300*, 801–805. https://doi.org/10.1126/science.1083219.
10. Liang Meng Wee, A., Tong, A.B., Jose Florez Ariza, A., Cañ ari-Chumpitaz, C., Grob, P., Nogales, E., Bustamante, C.J., and Meng Wee, L. (2023). A trailing ribosome speeds up RNA polymerase at the expense of transcript fidelity via force and allostery. Cell *186*, 1244-1262.e34. https://doi.org/10.1016/j.cell.2023.02.008.
11. Wang, C., Molodtsov, V., Firlar, E., Kaelber, J.T., Blaha, G., Su, M., and Ebright, R.H. (2020). Structural basis of transcription-translation coupling. Science (1979). *369*, 1359–1365. https://doi.org/10.1126/SCIENCE.ABB5317.
12. Webster, M.W., Takacs, M., Zhu, C., Vidmar, V., Eduljee, A., Abdelkareem, M., and Weixlbaumer, A. (2020). Structural basis of transcription-translation coupling and collision in bacteria. Science (1979). *369*, 1355–1359. https://doi.org/10.1126/SCIENCE.ABB5036.

13. Zhang, J., Lu, G., Zhou, W., Molodtsov, V., Cheng, Z., Zhang, Y., Li, R., Jiang, Z., Yang, M., Li, L., et al. (2026). Structural basis of translation in transcription-translation coupling. bioRxiv, 2024.03.13.584796. https://doi.org/10.1101/2024.03.13.584796.
14. Proshkin, S., Rachid Rahmouni, A., Mironov, A., and Nudler, E. (2010). Cooperation between translating ribosomes and RNA polymerase in transcription elongation. Science (1979). *328*, 504–508. https://doi.org/10.1126/science.1184939.
15. Epshtein, V., Toulmé, F., Rachid Rahmouni, A., Borukhov, S., and Nudler, E. (2003). Transcription through the roadblocks: the role of RNA polymerase cooperation. EMBO J. *22*, 4719. https://doi.org/10.1093/emboj/cdg452.
16. Wang, L., Watters, J.W., Ju, X., Lu, G., and Liu, S. (2023). Head-on and co-directional RNA polymerase collisions orchestrate bidirectional transcription termination. Mol. Cell *83*, 1153-1164.e4. https://doi.org/10.1016/j.molcel.2023.02.017.
17. Ju, X., Li, D., and Liu, S. (2019). Full-length RNA profiling reveals pervasive bidirectional transcription terminators in bacteria HHS Public Access. Nat Microbiol *4*, 1907–1918. https://doi.org/10.1038/s41564-019-0500-z.
18. 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. https://doi.org/10.1038/nature04268.
19. Ma, J., Bai, L., and Wang, M.D. (2013). Transcription under torsion. Science (1979). *340*, 1580–1583. https://doi.org/10.1126/science.1235441.
20. Yin, H., Wang, M.D., Svoboda, K., Landick, R., Block, S.M., and Gelles, J. (1995). Transcription against an applied force. Science (1979). *270*, 1653–1657. https://doi.org/10.1126/science.270.5242.1653.
21. Wang, M.D., Schnitzer, M.J., Yin, H., Landick, R., Gelles, J., and Block, S.M. (1998). Force and velocity measured for single molecules of RNA polymerase. Science (1979). *282*, 902–907. https://doi.org/10.1126/science.282.5390.902.
22. Galburt, E.A., Grill, S.W., Wiedmann, A., Lubkowska, L., Choy, J., Nogales, E., Kashlev, M., and Bustamante, C. (2007). Backtracking determines the force sensitivity of RNAP II in a factor-dependent manner. Nature *446*, 820–823. https://doi.org/10.1038/nature05701.
23. Ma, J., Tan, C., Gao, X., Fulbright, R.M., Roberts, J.W., and Wang, M.D. (2019). Transcription factor regulation of RNA polymerase's torque generation capacity. Proc. Natl. Acad. Sci. U. S. A. *116*, 2583–2588. https://doi.org/10.1073/pnas.1807031116.
24. Qian, J., Cartee, A., Xu, W., Yan, Y., Wang, B., Artsimovitch, I., Dunlap, D., and Finzi, L. (2024). Reciprocating RNA Polymerase batters through roadblocks. Nature Communications 2024 15:1 *15*, 1–11. https://doi.org/10.1038/s41467-024-47531-x.
25. Crampton, N., Bonass, W.A., Kirkham, J., Rivetti, C., and Thomson, N.H. (2006). Collision events between RNA polymerases in convergent transcription studied by atomic force microscopy. Nucleic Acids Res. *34*, 5416. https://doi.org/10.1093/NAR/GKL668.
26. Wright, D.J., King, K., and Modrich, P. (1989). The Negative Charge of Glu-11 1 Is Required to Activate the Cleavage Center of EcoRI Endonuclease. Journal of Biological Chemistry *264*, 11816–11821. https://doi.org/10.1016/S0021-9258(18)80138-7.

27. Finkelstein, I.J., Visnapuu, M.-L., and Greene, E.C. (2010). Single-molecule imaging reveals mechanisms of protein disruption by a DNA translocase. Nature *468*, 983–987. https://doi.org/10.1038/nature09561.
28. Frederick, C.A., Grable, J., Melia, M., Samudzi, C., Jen-Jacobson, L., Wang, B.C., Greene, P., Boyer, H.W., and Rosenberg, J.M. (1984). Kinked DNA in crystalline complex with EcoRI endonuclease. Nature 1984 309:5966 *309*, 327–331. https://doi.org/10.1038/309327a0.
29. Doughty, B., Kazer, S.W., and Eisenthal, K.B. (2011). Binding and cleavage of DNA with the restriction enzyme EcoR1 using time-resolved second harmonic generation. Proc. Natl. Acad. Sci. U. S. A. *108*, 19979–19984. https://doi.org/10.1073/PNAS.1115498108.
30. Punjani, A., and Fleet, D.J. (2023). 3DFlex: determining structure and motion of flexible proteins from cryo-EM. Nat. Methods *20*, 860. https://doi.org/10.1038/S41592-023-01853-8.
31. Komissarova, N., and Kashlev, M. (1997). RNA polymerase switches between inactivated and activated states by translocating back and forth along the DNA and the RNA. Journal of Biological Chemistry *272*, 15329–15338. https://doi.org/10.1074/jbc.272.24.15329.
32. Nudler, E., Kashlev, M., Nikiforov, V., and Goldfarb, A. (1995). Coupling between transcription termination and RNA polymerase inchworming. Cell *81*, 351–357. https://doi.org/10.1016/0092-8674(95)90388-7.
33. Borukhov, S., Sagitov, V., and Goldfarb, A. (1993). Transcript cleavage factors from E. coli. Cell *72*, 459–466. https://doi.org/10.1016/0092-8674(93)90121-6.
34. Zhang, J., and Landick, R. (2009). Substrate Loading, Nucleotide Addition, and Translocation by RNA Polymerase. RNA Polymerases as Molecular Motors, 206–235. https://doi.org/10.1039/9781847559982-00206.
35. Abdelkareem, M., Saint-André, C., Takacs, M., Papai, G., Crucifix, C., Guo, X., Ortiz, J., and Weixlbaumer, A. (2019). Structural Basis of Transcription: RNA Polymerase Backtracking and Its Reactivation. Mol. Cell *75*, 298-309.e4. https://doi.org/10.1016/j.molcel.2019.04.029.
36. Guo, X., Myasnikov, A.G., Chen, J., Takacs, M., Schultz, P., and Weixlbaumer, A. (2018). Structural Basis for NusA Stabilized Transcriptional Pausing. Mol. Cell *69*, 816–827. https://doi.org/10.1016/j.molcel.2018.02.008.
37. Meng, E.C., Goddard, T.D., Pettersen, E.F., Couch, G.S., Pearson, Z.J., Morris, J.H., and Ferrin, T.E. (2023). UCSF ChimeraX: Tools for structure building and analysis. Protein Science *32*, e4792. https://doi.org/10.1002/pro.4792.
38. Kang, J.Y., Mishanina, T. V., Landick, R., and Darst, S.A. (2019). Mechanisms of transcriptional pausing in bacteria. J. Mol. Biol. *431*, 4007. https://doi.org/10.1016/J.JMB.2019.07.017.
39. Kang, J.Y., Mishanina, T. V., Bellecourt, M.J., Mooney, R.A., Darst, S.A., and Landick, R. (2018). RNA polymerase accommodates a pause RNA hairpin by global conformational rearrangements that prolong pausing. Mol. Cell *69*, 802. https://doi.org/10.1016/j.molcel.2018.01.018.

40. Delbeau, M., Omollo, E.O., Froom, R., Koh, S., Mooney, R.A., Lilic, M., Brewer, J.J., Rock, J., Darst, S.A., Campbell, E.A., et al. (2023). Structural and functional basis of the universal transcription factor NusG pro-pausing activity in Mycobacterium tuberculosis. Mol. Cell *83*, 1474-1488.e8. https://doi.org/10.1016/j.molcel.2023.04.007.
41. Kang, J.Y., Mishanina, T. V., Bao, Y., Chen, J., Llewellyn, E., Liu, J., Darst, S.A., and Landick, R. (2023). An ensemble of interconverting conformations of the elemental paused transcription complex creates regulatory options. Proc. Natl. Acad. Sci. U. S. A. *120*, e2215945120. https://doi.org/10.1073/PNAS.2215945120.
42. Kim, Y., Grable, J.C., Love, R., Greene, P.J., and Rosenberg, J.M. (1990). Refinement of Eco RI endonuclease crystal structure: A revised protein chain tracing. Science (1979). *249*, 1307–1309. https://doi.org/10.1126/science.2399465.
43. Artsimovitch, I., and Landick, R. (2000). Pausing by bacterial RNA polymerase is mediated by mechanistically distinct classes of signals. Proc. Natl. Acad. Sci. U. S. A. *97*, 7090–7095. https://doi.org/10.1073/pnas.97.13.7090.
44. Zhu, C., Guo, X., Dumas, P., Takacs, M., Abdelkareem, M., Vanden Broeck, A., Saint-André, C., Papai, G., Crucifix, C., Ortiz, J., et al. (2022). Transcription factors modulate RNA polymerase conformational equilibrium. Nature Communications 2022 13:1 *13*, 1546-. https://doi.org/10.1038/s41467-022-29148-0.
45. Jen-Jacobson, L., Lesser, D., and Kurpiewski, M. (1986). The enfolding arms of EcoRI endonuclease: Role in DNA binding and cleavage. Cell *45*, 619–629. https://doi.org/10.1016/0092-8674(86)90294-1.
46. Ju, X., Li, S., Froom, R., Wang, L., Lilic, M., Delbeau, M., Campbell, E.A., Rock, J.M., and Liu, S. (2024). Incomplete transcripts dominate the Mycobacterium tuberculosis transcriptome. Nature 2024 627:8003 *627*, 424–430. https://doi.org/10.1038/s41586-024-07105-9.
47. Shundrovsky, A., Santangelo, T.J., Roberts, J.W., and Wang, M.D. (2004). A Single-Molecule Technique to Study Sequence-Dependent Transcription Pausing. Biophys. J. *87*, 3945. https://doi.org/10.1529/biophysj.104.044081.
48. Zhang, M., Cañari-Chumpitaz, C., Liu, J., Onoa, B., Cleir, S. de, Cheng, E., Requejo, K.I., and Bustamante, C. (2026). Apical Localization of RNA Polymerases Modulate Transcription Dynamics and Supercoiling Domains Revealed by Cryo-ET. bioRxiv, 2026.03.25.714350. https://doi.org/10.64898/2026.03.25.714350.
49. Chen, J., Gopalkrishnan, S., Chiu, C., Chen, A.Y., Campbell, E.A., Gourse, R.L., Ross, W., and Darst, S.A. (2019). E. Coli TraR allosterically regulates transcription initiation by altering RNA polymerase conformation. Elife *8*. https://doi.org/10.7554/eLife.49375.
50. Qian, J., Lubkowska, L., Zhang, S., Tan, C., Hong, Y., Jia, X., Fulbright, R.M., Inman, J.T., Kay, T.M., Jeong, J., et al. (2025). Chromatin buffers torsional stress during transcription. Science (1979). https://doi.org/10.1126/science.adv0134.
51. Bradford, M.M. (1976). A rapid and sensitive method for the quantitation of microgram quantities of protein utilizing the principle of protein-dye binding. Anal. Biochem. *72*, 248–254. https://doi.org/10.1016/0003-2697(76)90527-3.

52. Perederina, A.A., Vassylyeva, M.N., Berezin, I.A., Svetlov, V., Artsimovitch, I., and Vassylyev, D.G. (2005). Cloning, expression, purification, crystallization and initial crystallographic analysis of transcription elongation factors GreB from Escherichia coli and Gfh1 from Thermus thermophilus. urn:issn:1744-3091 *62*, 44–46. https://doi.org/10.1107/S1744309105040297.
53. Hao, Z., Epshtein, V., Kim, K.H., Proshkin, S., Svetlov, V., Kamarthapu, V., Bharati, B., Mironov, A., Walz, T., and Nudler, E. (2021). Pre-termination Transcription Complex: Structure and Function. Mol. Cell *81*, 281-292.e8. https://doi.org/10.1016/j.molcel.2020.11.013.
54. Schmidt, M.C., and Chamberlin, M.J. (1987). nusA Protein of Escherichia coli is an efficient transcription termination factor for certain terminator sites. J. Mol. Biol. *195*, 809–818. https://doi.org/10.1016/0022-2836(87)90486-4.
55. Burgess, R.R., Travers, A.A., Dunn, J.J., and Bautz, E.K.F. (1969). Factor Stimulating Transcription by RNA Polymerase. Nature 1969 221:5175 *221*, 43–46. https://doi.org/10.1038/221043a0.
56. Kang, J.Y., Llewellyn, E., Chen, J., Olinares, P.D.B., Brewer, J., Chait, B.T., Campbell, E.A., and Darst, S.A. (2021). Structural basis for transcription complex disruption by the Mfd translocase. Elife *10*, 1–86. https://doi.org/10.7554/eLife.62117.
57. Mastronarde, D.N. (2003). SerialEM: A Program for Automated Tilt Series Acquisition on Tecnai Microscopes Using Prediction of Specimen Position. Microscopy and Microanalysis *9*, 1182–1183. https://doi.org/10.1017/s1431927603445911.
58. Mastronarde, D.N. (2005). Automated electron microscope tomography using robust prediction of specimen movements. J. Struct. Biol. *152*, 36–51. https://doi.org/10.1016/j.jsb.2005.07.007.
59. Punjani, A., Rubinstein, J.L., Fleet, D.J., and Brubaker, M.A. (2017). cryoSPARC: algorithms for rapid unsupervised cryo-EM structure determination. Nature Methods 2017 14:3 *14*, 290–296. https://doi.org/10.1038/nmeth.4169.
60. Punjani, A., Zhang, H., and Fleet, D.J. (2020). Non-uniform refinement: adaptive regularization improves single-particle cryo-EM reconstruction. Nature Methods 2020 17:12 *17*, 1214–1221. https://doi.org/10.1038/s41592-020-00990-8.
61. Emsley, P., Lohkamp, B., Scott, W.G., and Cowtan, K. (2010). Features and development of Coot. Acta Crystallogr. D Struct Biol. *66*, 486–501. https://doi.org/10.1107/S0907444910007493.
62. Schrödinger LLC (2015). The PyMOL Molecular Graphics System, Version~1.8.
63. Croll, T.I. (2018). ISOLDE: A physically realistic environment for model building into low-resolution electron-density maps. Acta Crystallogr. D Struct. Biol. *74*, 519–530. https://doi.org/10.1107/S2059798318002425.
64. Liebschner, D., Afonine, P. V., Baker, M.L., Bunkoczi, G., Chen, V.B., Croll, T.I., Hintze, B., Hung, L.W., Jain, S., McCoy, A.J., et al. (2019). Macromolecular structure determination using X-rays, neutrons and electrons: recent developments in Phenix. Acta Crystallogr. D Struct Biol. *75*, 861–877. https://doi.org/10.1107/S2059798319011471.
65. Schindelin, J., Arganda-Carreras, I., Frise, E., Kaynig, V., Longair, M., Pietzsch, T., Preibisch, S., Rueden, C., Saalfeld, S., Schmid, B., et al. (2012). Fiji: an open-source

platform for biological-image analysis. Nature Methods 2012 9:7 *9*, 676–682. https://doi.org/10.1038/nmeth.2019.

66. Virtanen, P., Gommers, R., Oliphant, T.E., Haberland, M., Reddy, T., Cournapeau, D., Burovski, E., Peterson, P., Weckesser, W., Bright, J., et al. (2020). SciPy 1.0: fundamental algorithms for scientific computing in Python. Nature Methods 2020 17:3 *17*, 261–272. https://doi.org/10.1038/s41592-019-0686-2.
67. Cock, P.J.A., Antao, T., Chang, J.T., Chapman, B.A., Cox, C.J., Dalke, A., Friedberg, I., Hamelryck, T., Kauff, F., Wilczynski, B., et al. (2009). Biopython: freely available Python tools for computational molecular biology and bioinformatics. Bioinformatics *25*, 1422–1423. https://doi.org/10.1093/bioinformatics/btp163.
68. Galls, D., Mueller, A.U., Greenwald, E., and Fire, A.Z. (2025). A general RNA-templated RNA extension activity of E. coli RNA polymerase. RNA *31*, 663–678. https://doi.org/10.1261/rna.080238.124.
69. Hunter, J.D. (2007). Matplotlib: A 2D graphics environment. Comput. Sci. Eng. *9*, 90–95. https://doi.org/10.1109/MCSE.2007.55.

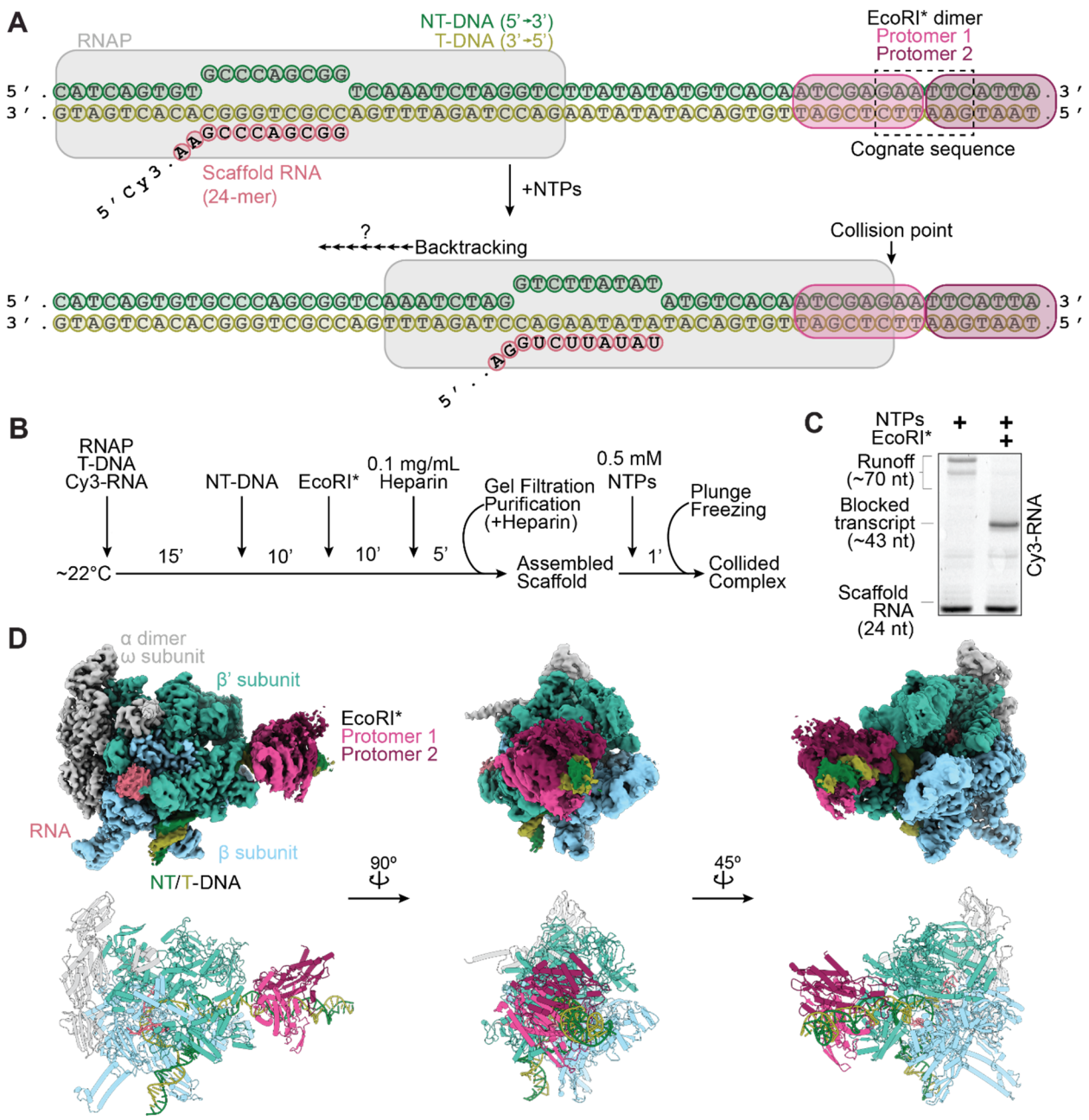

Figure 1. Reconstituting transcriptional collision with a stationary roadblock.
[A] Synthetic nucleic acid scaffold design to assemble E. coli RNAP and EcoRI* onto a DNA template. NT-DNA, non-template DNA strand; T-DNA, template DNA strand.
[B] Scaffold assembly and purification scheme.
[C] Representative denaturing gel showing blocked transcription in the presence of EcoRI*, suggesting collided complex assembly. Cy3-labeled RNA was visualized.
[D] Cryo-EM density map (top) and atomic model (bottom) of the RNAP-EcoRI* collided complex.

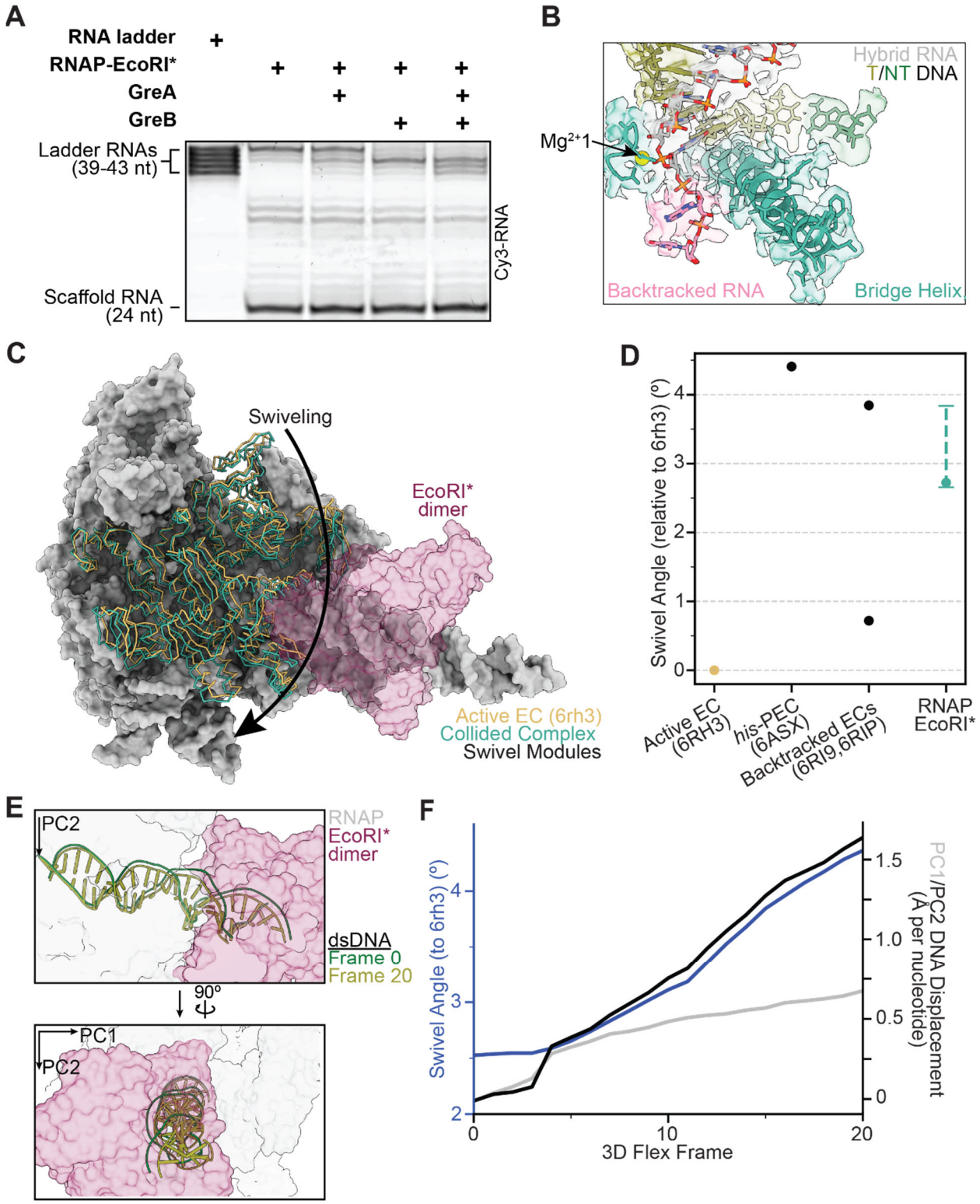

Figure 2. EcoRI* collided complex is backtracked and moderately swiveled. **[A]** Representative gel of in vitro transcription assay in the presence and absence of Gre factors, with an RNA ladder corresponding to potential transcript lengths within the collided complex.

[B] View of the active site of the collided complex (cryo-EM map and atomic model). Resolvable backtracked bases are highlighted in pink.
[C] View of the swivel module of an active elongation complex (PDB: 6RH3) versus that of the EcoRI* collided complex, superimposed on active EC structure after alignment, highlighting differential swiveling.
[D] Swivel angle magnitudes of prior RNAP structures versus that of the RNAP-EcoRI* collided complex. The value for the consensus reconstruction is displayed as a solid dot, while the range measured across all five 3D flex trajectories for the collided complex is displayed as a dashed line.
[E] Initial and final frames of the trajectory displayed in **F**, highlighting DNA displacement.
[F] Swivel angle and DNA displacement measured for each frame of the first series (accounting for the largest conformational movements) of the 3DFlex variability trajectory for the collided complex.

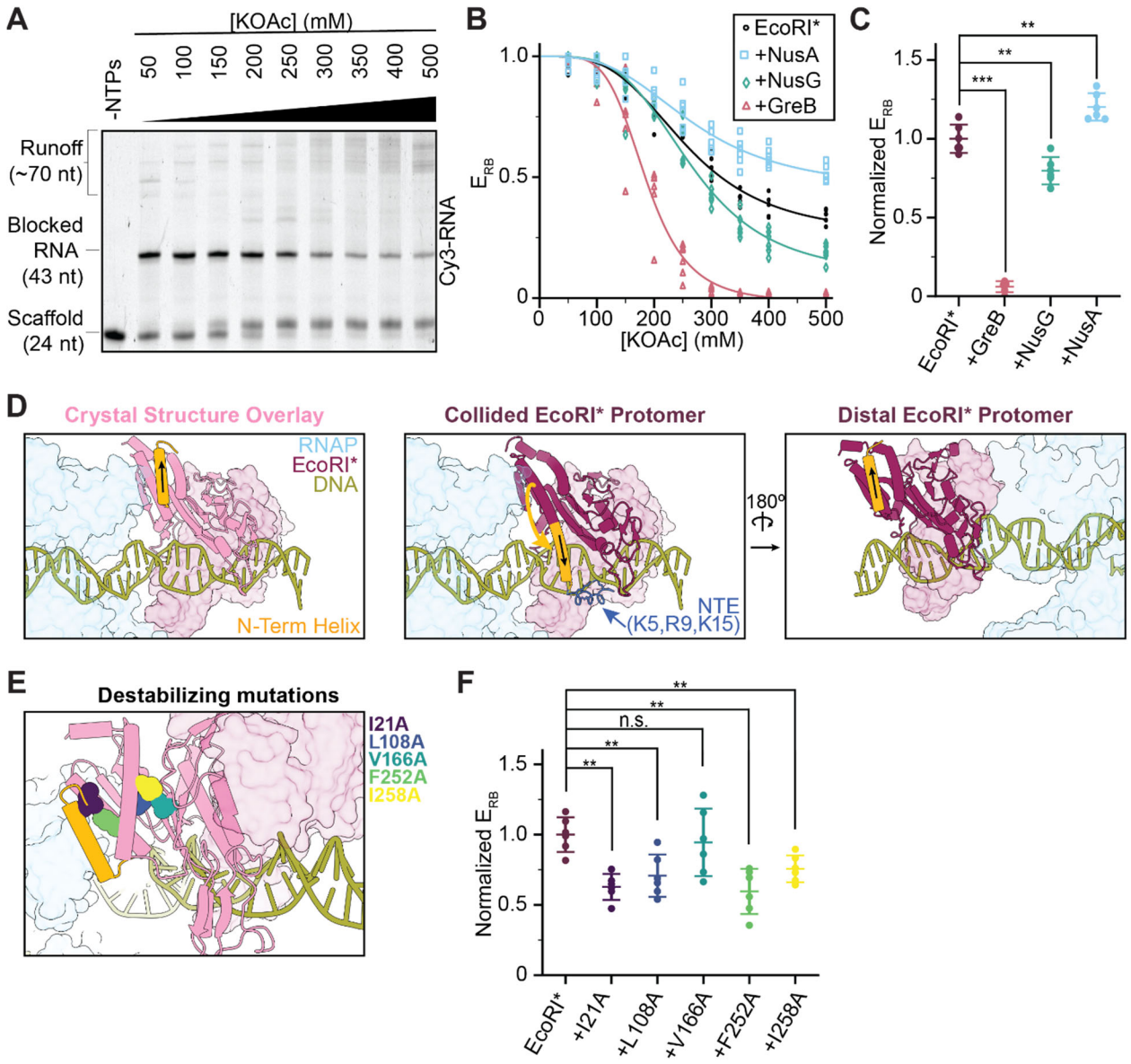

Figure 3. Factors that affect bypass efficiency.
[A] Representative EcoRI* bypass assay. The RNA lengths under different conditions are visualized via Cy3 fluorescence.
[B] Fraction of blocked RNAP in the absence and presence of factors (GreB, NusA, NusG). Raw measurements and their sigmoidal fit for each condition are displayed.
[C] Normalized roadblock efficiencies in the presence of indicated factors at 300 mM KOAc. Data are shown as mean ± S.D. Individual points are also displayed (N = 6).
[D] Views highlighting the orientation of EcoRI* N-terminal helix in the collided complex versus in the EcoRI* crystal structure (PDB 1ERI). The helix of the collided EcoRI* protomer is flipped and re-oriented toward the downstream dsDNA, which may be stabilized by NTE-DNA interactions, while the distal, non-collided protomer (right) adopts a conformation similar to the crystal structure.
[E] Positions of single alanine mutations are highlighted on the crystal structure overlay.

1105 **[F]** Normalized roadblock efficiencies for the destabilizing EcoRI* mutations at 300 mM
1106 KOAc. Data are shown as mean ± S.D. Individual points are also displayed (N = 6).
1107 Conditions were compared via two-tailed, unpaired t test including Welch's correction and
1108 Benjamini-Hochberg correction for multiple testing. n.s., not significant; *$p < 0.05$; **$p <$
1109 0.01; ***$p < 0.001$.

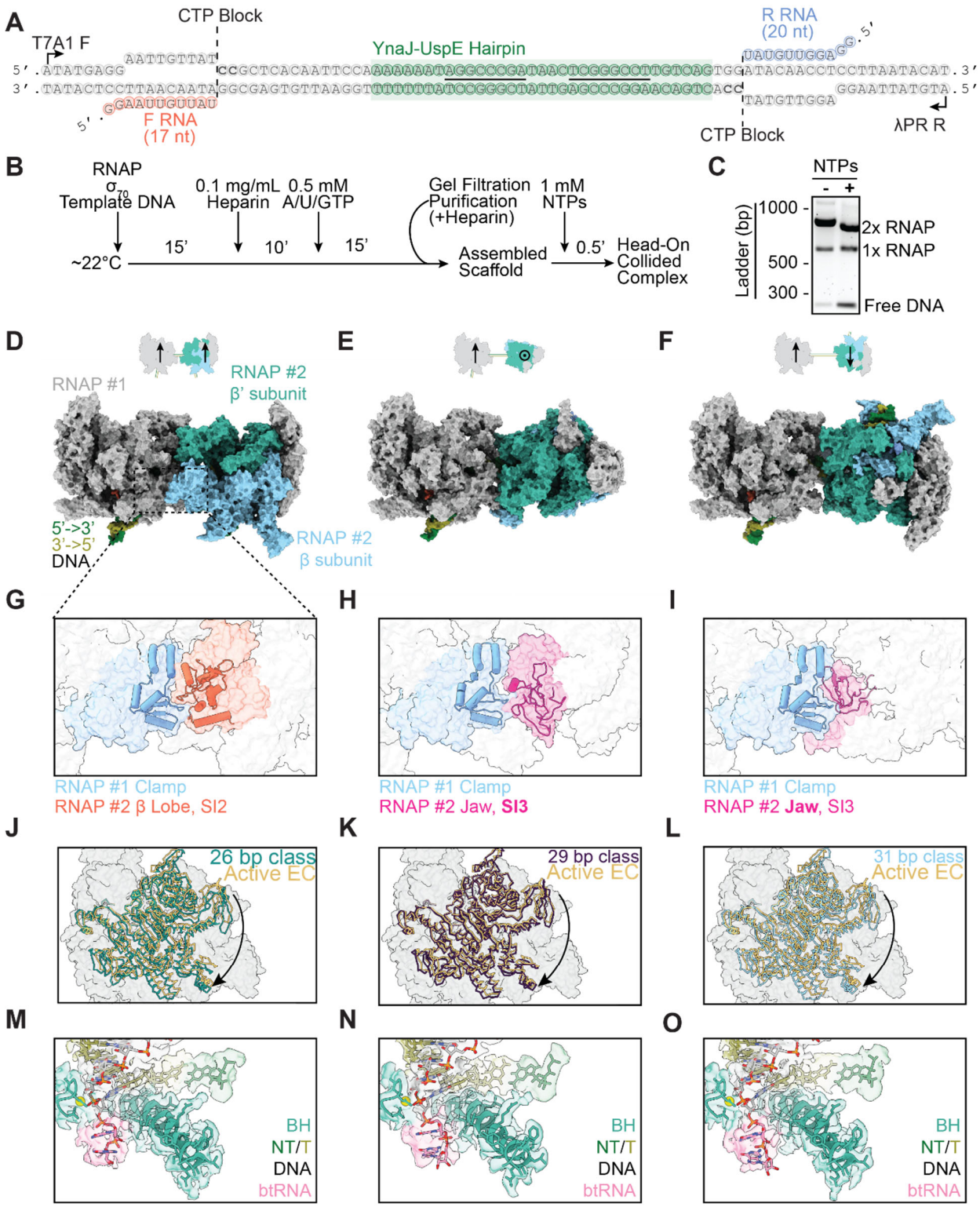

Figure 4. Reconstituting head-on RNAP-RNAP collisions.
[A] DNA template featuring two promoters, designed to stall RNAP at cytosine sites. F RNA, forward RNA; R RNA, reverse RNA. Underlined segments of the YnaJ-UspE hairpin indicate complementary regions.

[B] Experimental scheme for head-on RNAP-RNAP collided complex assembly and purification.
[C] Native gel shift assay for evaluating collided complex formation.
[D-O] Columns correspond to 26-bp (left), 29-bp (middle), and 31-bp (right) DNA inter-RNAP distance complexes. **[D-F]** Top: Cartoons indicating the relative orientation (arrows) and position of polymerases. Bottom: Surface representations of the corresponding atomic models. **[G-I]** Zoom-in views of the inter-RNAP interface for each class. **[J-L]** Views of the swivel module for each class (from symmetry-expanded reconstructions) superimposed on an active EC structure after alignment. **[M-O]** Zoom-in views of the backtracked RNA (btRNA). Cryo-EM maps and atomic models are displayed.

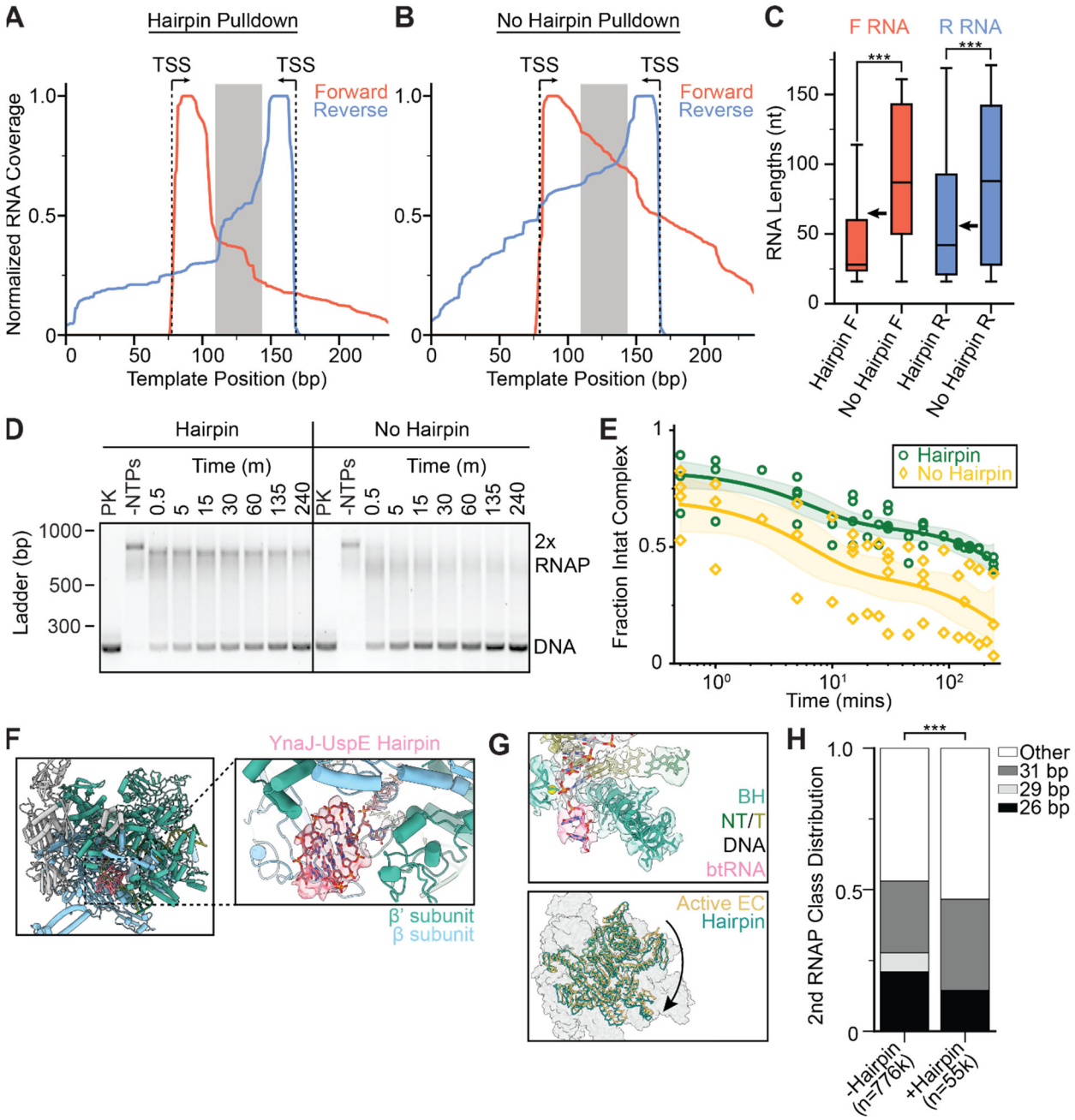

Figure 5. RNA hairpin impacts collision positioning and stability.
[A, B] SEnd-seq profiles of forward and reverse RNAs in head-on RNAP-RNAP collided complexes on a template encoding the YnaJ-UspE hairpin (**A**) or a no-hairpin template (**B**). The grey box defines the hairpin sequence region or the corresponding scrambled sequence region.
[C] Quantification of RNA lengths (mapped back to the originating promoter). N = 1789 (hairpin F), 8075 (no hairpin F), 608 (hairpin R), and 5945 (no hairpin R). Arrows indicate the predicted RNA lengths if transcription terminates right after the hairpin. Conditions were compared with a two-tailed, unpaired t test with Welch's correction. ***$p < 0.001$.

1137 **[D]** Native agarose gels for evaluating collided complex disassembly. PK: Proteinase K
1138 treated. Timepoints indicate time elapsed after NTP addition, followed by EDTA quench.
1139 Left: hairpin-containing template DNA; right, no-hairpin template DNA. DNA was
1140 visualized with SYBR safe.
1141 **[E]** Quantification of the fraction of intact collided complex under indicated conditions.
1142 Nucleic acid intensity values were normalized for each of four experimental replicates to
1143 the -NTPs and +Proteinase K conditions. Shaded areas indicate 95% confidence interval
1144 of the fits.
1145 **[F]** View of the hairpin-containing head-on collided RNAP structure obtained through
1146 symmetry expansion.
1147 **[G]** Views of the hairpin-containing RNAP (cryo-EM map and atomic model) highlighting
1148 the backtracked RNA (top) and rotation of the swivel module relative to the active
1149 elongation complex.
1150 **[H]** Class distributions of no-hairpin and hairpin-containing collided RNAP complexes.
1151 "Other" is defined as particles that show well-defined density for one RNAP but the density
1152 downstream of the well-defined RNAP is too low of a resolution to confidently model an
1153 RNAP. Significance was calculated via a chi-squared test. ***$p < 0.001$.

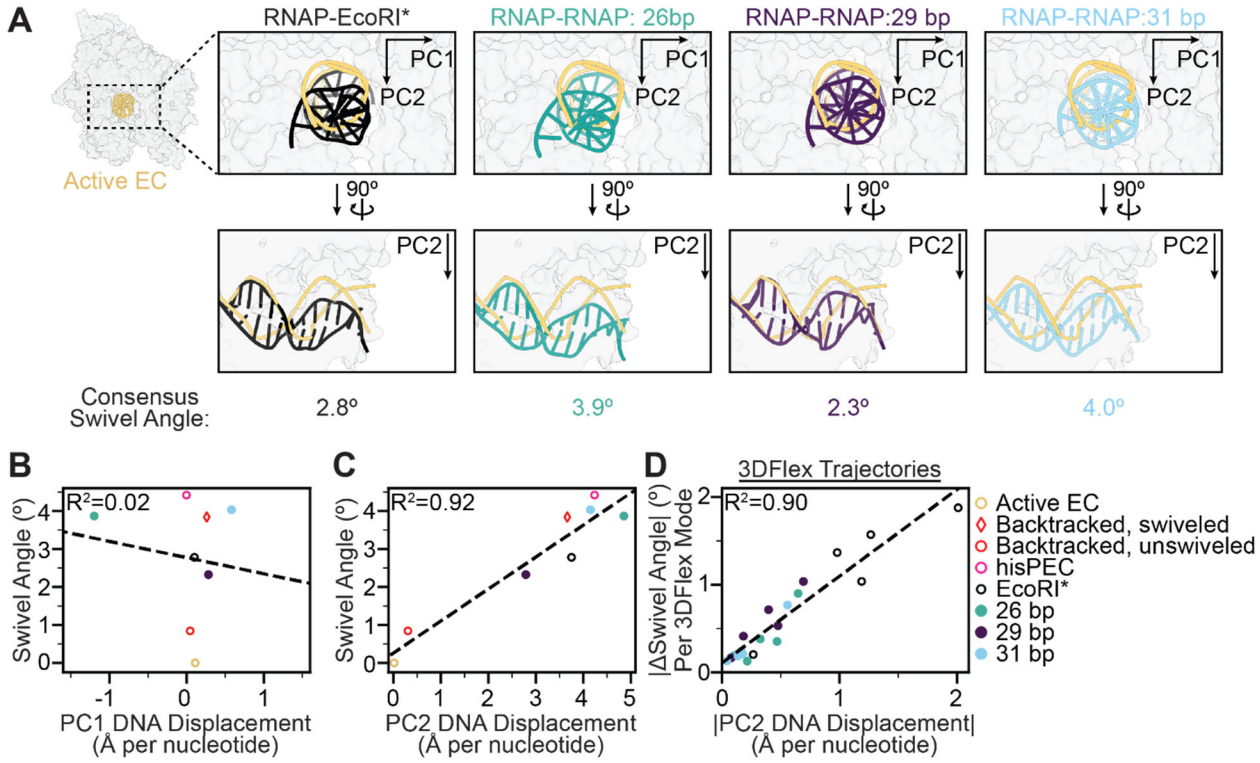

Figure 6. Inter-RNAP distance modulates deformation-swivel coupling.
[A] Downstream DNA deformation of each collision scenario compared against active EC (PDB 6RH3).
[B-C] Plots of swiveling angle versus DNA displacement along **[B]** PC1 and **[C]** PC2 of consensus collision structures (this study) and previously reported structures of an active EC (PDB 6RH3), backtracked ECs (unswiveled, PDB 6RI9; swiveled, PDB 6RIP), and his pause EC (PDB 6ASX). Linear fits are displayed (dashed lines).
[D] Plot of the magnitudes of swiveling and PC2 displacement measured in individual 3DFlex series trajectories for the four collided structures (EcoRI*, 26 bp, 29 bp, 31 bp – 5 series each). A linear fit is displayed (dashed line).

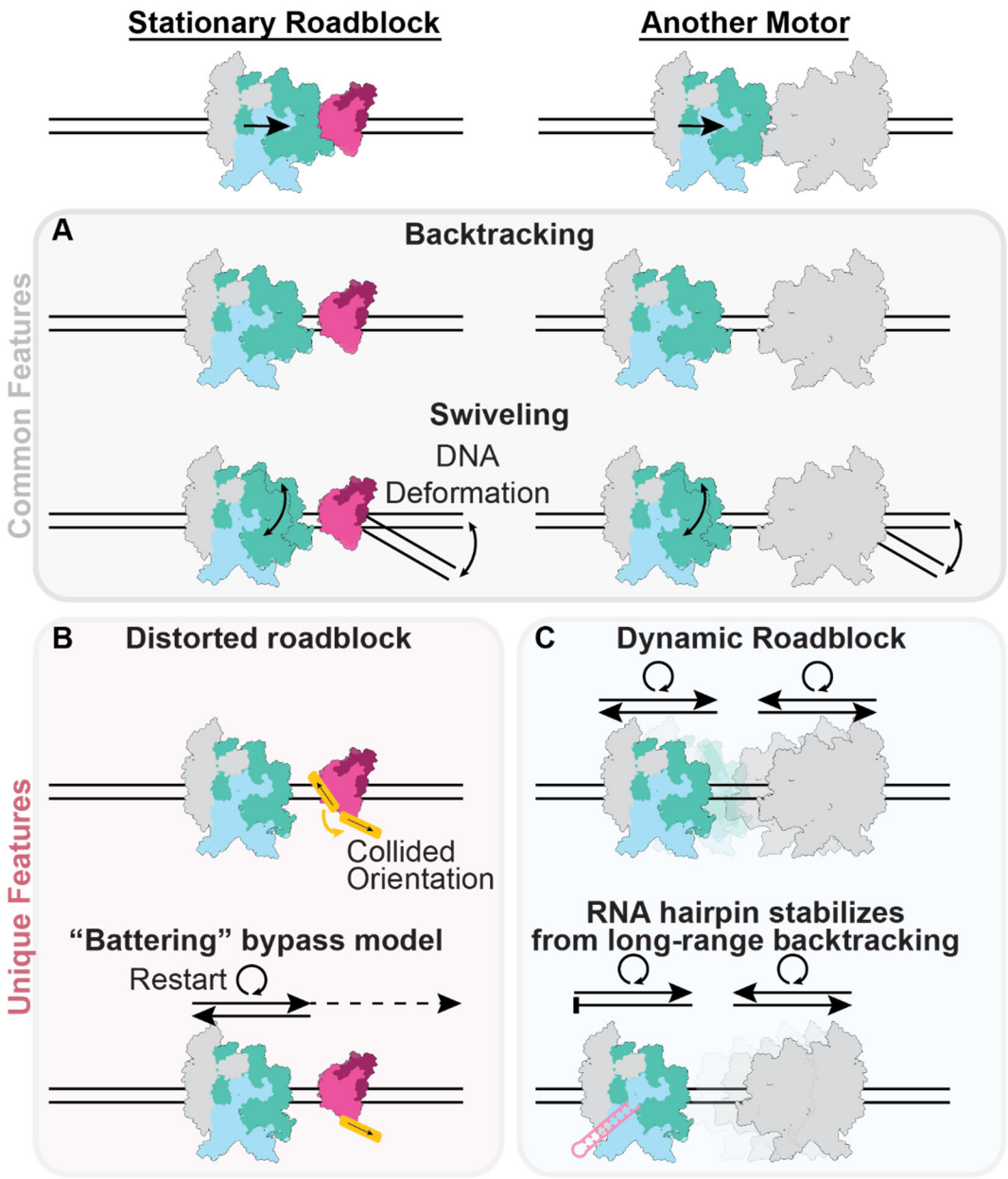

Figure 7. Structural mechanisms of transcriptional collisions.
[A] Common features of all collisions in this study. Regardless of the roadblock identity, RNAP backtracks away from the collision site and swivels into a transcriptionally inactive state, mediated by coupling between template DNA distortion and RNAP swiveling.
[B, C] Unique features of different collisions. **[B]** RNAP collision with a model stationary roadblock (EcoRI*) features roadblock remodeling, a plausible mechanism for how 'battering' (repetitive transcriptional restart and collision) can facilitate roadblock bypass. **[C]** Dynamic RNAP-RNAP collisions feature substantial heterogeneity due to the stochastic transcriptional and backtracking activity of both polymerases. These dynamics can be tempered by an RNA hairpin that constrains the collision positions.