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[2,4*], 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
[4]CZ Biohub New York, LLC
*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 that 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—compared 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

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

214 Having observed substantial structural alterations in RNAP upon collision, we next
215 examined whether the structure of the EcoRI* roadblock is also impacted. EcoRI*
216 engages DNA as a homodimer[28,42,45]. The protein's DNA-bound conformation is overall
217 highly similar between the RNAP-EcoRI* collided complex and a previous crystal
218 structure in the absence of RNAP (PDB: 1ERI)[42], with the notable exception for the N-
219 terminal helix (AA 20-31) of the protomer proximal to the collision interface, which flips
220 ~150°, orienting it away from RNAP (**Fig. 3D, Movie S2**). This rearrangement was not
221 observed in the distal EcoRI* protomer, suggesting it is specifically linked to contact with
222 RNAP (**Fig. 3D**). We speculate that it occurs at the collision site to reduce steric clashes
223 with RNAP, becoming kinetically trapped as the RNAP backtracks away. Additionally, this
224 rearrangement repositions EcoRI*'s N-terminal extension (NTE) residues (5K, 9R, 15K)
225 in an orientation where they can access DNA (**Fig. 3D**). These amino acids have been
226 previously shown to enhance DNA binding in biochemical assays, presumably by
227 mediating electrostatic interactions[26,28,42,45].
228 In the EcoRI* crystal structure, the N-terminal helix forms hydrophobic contacts
229 with the protein's structural core[42]. As a protein's potency as a roadblock intuitively
230 depends on its mechanical resilience, we posited that destabilizing the overall folded
231 structure of EcoRI* would compromise its roadblocking efficacy. To test this hypothesis,
232 we introduced alanine substitutions at the positions of hydrophobic residues within the
233 EcoRI* core (**Fig. 3E**), distal from the DNA binding interface, and assessed their impact
234 on E_{RB} using our roadblock assay. We found that all point mutations (with the exception
235 of V166A) led to significantly increased RNAP bypass (decreased E_{RB}) (**Fig. 3F**). Notably,
236 the two mutations that caused the largest magnitude changes, I21A and F252A, are
237 internally facing residues adjacent to the predicted collision interface that are likely to
238 mediate a stabilizing contact between the N-terminal helix and the EcoRI* core (**Fig. 3E,**
239 **F**). These results suggest that the structural stability of a roadblock protein—independent
240 from its affinity to DNA—is linked to its capacity to resist physical impacts from motors
241 like RNAP; correspondingly, mechanical destabilization of roadblocks can enhance
242 bypass.
243
244 *Reconstituting head-on collisions between two RNAPs*
245 To explore the generality of RNAP's collision response, we next investigated
246 another physiological scenario, when an RNAP runs into another active motor complex
247 rather than a stationary roadblock. In previous studies, we demonstrated that head-on
248 collisions between two RNAPs produce long-lived complexes on DNA[16], constituting a
249 prevalent mechanism for transcription termination of convergent genes in *E. coli*[17]. To
250 reconstitute a head-on RNAP-RNAP collision in a format suitable for cryo-EM structural
251 interrogation, we constructed a DNA template with two convergent promoters (**Fig. 4A**).
252 Each promoter was bound by a σ^{70}-RNAP holoenzyme that was walked to the +18/+20
253 position by omitting one of the four nucleotides (CTP), followed by purification via size
254 exclusion chromatography. Transcription was then restarted with addition of 1 mM NTPs
255 to induce a collision (**Fig. 4B**). We also introduced a native sequence from the *E. coli*
256 genomic region between the *YnaJ-UspE* gene pair into the template, which encodes an
257 RNA hairpin containing an 8-bp stem. We have previously shown that this sequence is
258 important for positioning the collision site between a pair of RNAPs, mediating precise
259 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

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

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

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

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

417 This work establishes the feasibility of scrutinizing genomic collisions using cryo-
418 EM, providing a path towards visualizing physiological scenarios of increasing complexity.
419 An important consideration for future studies is the topology of the template DNA, which
420 may be addressable by employing closed templates with varying degrees of
421 supercoiling[48,49]. Additionally, collisions featuring additional machines, such as co-
422 directionally transcribing RNAP and translating ribosomes (both of which can promote
423 roadblock bypass[3,10,15]) or head-on/co-directional DNA replication forks (replication-
424 transcription conflict as a primary source of genome instability[4–6]) are likely accessible by
425 similar approaches. Such studies will be critical for a mechanistic understanding of how
426 mechanical conflicts are managed across the crowded genome and functionally
427 harnessed to enhance the robustness of gene expression.

428 **Methods**

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

474
475 *GreA, GreB, NusA, NusG, σ^{70}*
476 GreA, GreB, NusA, NusG, and σ^{70} were expressed, purified was performed as described
477 previously[16,50,52,53].
478 Endonucleolytic cleavage activity of GreA/B was verified as described previously[16].
479 In brief, a backtracked RNAP EC scaffold (nucleotide sequences in **Table S1**) was formed
480 with a short (2 nt RNA backtracked) and long backtracked EC (6 nt RNA backtracked).
481 Oligonucleotides were high-performance liquid chromatography (HPLC) purified (IDT).
482 Activity assays were performed at 37 °C for 10 minutes by adding GreB or GreA at a final
483 concentration of 600 nM in the absence or presence of NTPs. The reactions were stopped
484 by addition of an equal volume of 2x RNA loading buffer and incubation at 90 °C for 2
485 minutes. Products were separated on a 25% Urea-PAGE gel, followed by fluorescence
486 visualization on a Typhoon FLA 7000 scanner. Full-length RNA products indicative of
487 endonucleolytic cleavage and transcriptional restart were enhanced in the presence of
488 GreA (short backtracked EC) and GreB (long backtracked EC) as expected[33].
489 As reported previously, NusA and NusG were verified to be active by bulk in vitro
490 transcription assays in transcription buffer (25 mM Tris-hydrochloride pH 7.5, 150 mM
491 potassium chloride, 10 mM magnesium chloride, 1mM DTT). In brief, Rho and NusG in
492 combination were shown to enhance Rho mediated termination efficiency on the Rho-
493 dependent λtR1terminator[53]. NusA reduced transcriptional readthrough across the ynaJ-
494 uspE bidirectional terminator[54]. σ^{70} was verified to be active in promoting T7A1 and
495 LambdaPR dependent promoter complex formation and transcription in in vitro
496 transcription assays in a transcription buffer (25 mM Tris-HCl pH 7.5, 150 mM KCl, 10
497 mM $MgCl_2$, 1mM DTT)[55].
498
499 **RNAP-EcoRI* Scaffold Complex Assembly**
500 DNA oligonucleotide sequences (Polyacrylamide gel electrophoresis (PAGE)-purified
501 from IDT) and a 5' Cy3-labeled RNA oligonucleotide (Dharmacon) were designed using
502 a previously used scaffold as a starting point (sequences of template, non-template, and
503 RNA oligos in **Table S1**)[56]. Template DNA and RNA were annealed in a 1:1 ratio in 1x
504 IDT annealing buffer (100 mM potassium acetate (KOAc); 30 mM HEPES, pH 7.5) for 5
505 minutes at 95º C, 5 minutes at 75º C, 5 minutes at 45º C, and 40 cycles of 1º C steps (1
506 minute hold each step, final temperature 4º C) in a thermocycler. RNAP core desalted
507 into scaffolding buffer (20 mM Tris pH 7.5 at RT, 150 mM potassium glutamate (KGlu), 5
508 mM magnesium acetate ($Mg(OAc)_2$), 2.5 mM DTT) was then incorporated and incubated
509 for 15 minutes at RT. 1.5-fold excess of non-template DNA was then added to the
510 complex mixture and incubated for 10 minutes at RT. 5-fold excess of EcoRI* was then
511 added to the mixture (this step was excluded for the -EcoRI* controls) and incubated for
512 10 minutes at RT before addition of 0.1 mg/ml final concentration of heparin sodium salt
513 to remove non-specifically bound RNAP and EcoRI* proteins. The fully assembled
514 complex was purified away from incompletely assembled complexes via gel filtration
515 chromatography using a Superdex 200 Increase column equilibrated with low salt buffer
516 (20 mM Tris-acetate (Tris-OAc) pH 7.8 at RT, 50 mM KOAc, 5 mM $Mg(OAc)_2$, 2.5 mM
517 DTT) containing 0.1 mg/ml heparin (Sigma-Aldrich) to prevent re-binding of non-
518 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.

565 *Cryo-EM data acquisition and processing*
566 For RNAP-EcoRI* complex, three different grids were imaged from two different sample
567 preparations of the complex, frozen on different days. Datasets were collected on the
568 same FEI Titan Krios G3i system operating at 300 kV and equipped with a Gatan K3 direct
569 electron detector, BioQuantum energy filter, and a spherical aberration corrector using
570 super-resolution mode. Frame sequences ("movies") were recorded using the SerialEM
571 software[57] suite at a nominal magnification of ×81,000, corresponding to a calibrated pixel
572 size of 0.86 Å at the specimen level (super-resolution pixel size of 0.43 Å per pixel)[58].
573 Each 1.4 s exposure was dose-fractionated across 35 frames, with a total electron dose
574 of 47.3 e^{-} $Å^{-2}$ (1.35 e^{-} $Å^{-2}$ per frame), with defocus values ranging from −0.8 to −2.0 µm
575 underfocus. Beam-image shift was used to collect 43,378 single exposures (4.6k, 17.8k,
576 23.8k exposures from the different grids) from nine holes in a 3-by-3 grid per each stage
577 translation.
578 The RNAP-RNAP complex dataset was collected similarly, with the following
579 alterations. Data were collected on a single grid on a FEI Titan Krios G2 operating at
580 300 kV and equipped with a Gatan K3 direct electron detector and BioQuantum energy
581 filter. Movies were recorded at a nominal magnification of x105,000, corresponding to a
582 calibrated pixel size of 0.847 Å at the specimen level (super-resolution pixel size of
583 0.4235 Å per pixel)[58]. The total electron dose was 48.79 e^{-} $Å^{-2}$ (1.39 e^{-} $Å^{-2}$ per frame).
584 Beam-image shift was used to collect 20,080 single exposures.
585
586 *Cryo-EM data processing*
587 All processing was performed using cryoSPARC[59]. Movies were aligned with Patch
588 Motion correction (B-factor: 150), and contrast transfer function (CTF) parameters were
589 estimated with Patch CTF using default settings. For each dataset, blob picker was
590 initially used to pick particles from a 256-pixel box, followed by 2D and 3D classification
591 on extracted particles to remove junk particles and RNAPs that did not restart. This initial
592 particle stack was used to generate a preliminary reconstruction that was used to re-pick
593 particles using a template picker approach to yield more particles. All of the picked
594 particles were then re-extracted with a 400-pixel box and separated with multiple rounds
595 of 3D heterogeneous refinement job with candidate 3D volumes as well as junk particle
596 volumes (**Fig. S2, S6**). All models were then further refined via a non-uniform refinement
597 job[60], local CTF refinement, and global CTF refinement. Then low per-particle scale
598 particles were removed, followed by per-particle reference-based motion correction and
599 subsequent local refinement. 3DFlex[30] was then performed (N=5 modes) to improve the
600 density of high flexible regions in the particles and generate variability trajectories for DNA
601 deformation-RNAP swiveling analysis.
602 For the RNAP-RNAP dataset, the following additional steps were performed.
603 During test dataset processing a 400-pixel box size was used and binned by 4 to enhance
604 processing speed during initial processing steps. Particle box size was increased to 700
605 pixels (no binning) to accommodate the additional RNAP. After heterogeneous
606 refinement, an additional round of 3D classification was performed using low-resolution
607 volumes as bait volumes (generated by manual docking to mimic RNAPs with 26-36 bp
608 of separation followed by molmap (resolution 20) volume generation from atomic models)
609 using ChimeraX to see if any additional states could be extracted. This resulted in the
610 detection of the 29 bp inter-RNAP class. After non-uniform refinement, models were re-

611 oriented along their C2 symmetry axis and symmetry expanded before local refinement
612 and further polishing with a one RNAP sized mask.
613 To isolate particles that had clear hairpin density in the exit channel, a new set of
614 particles were extracted from the same micrographs (to remove any potential bias from
615 the C2 symmetric particles that were able to be resolved to high resolution) and subjected
616 to a similar processing pipeline to a resolve a single RNAP class (box size 400 pixels)
617 using a focus mask corresponding to a single RNAP (**Fig. S9**). This class contained
618 downstream DNA and low-resolution signal further downstream that could correspond to
619 collided RNAP density being averaged between different states (since the second RNAP
620 was not included in the focus mask). Taken together, this indicates that these single
621 RNAPs are still in a post-collision state instead of the alternate hypothesis such as run-
622 off RNAPs or non-bound RNAPs. These particles were subjected to polishing and
623 refinement prior to an additional round of focused classification performed with a mask
624 around the RNAP exit channel that showed one class with clear RNA hairpin density.
625 Particles contributing to the hairpin-containing class were subjected to final polishing and
626 refinement. Additionally, to compare the collided states of the hairpin-containing RNAPs,
627 hairpin-containing particles were re-extracted with a larger box size (700 pixels) and
628 subjected to 3D classification with a focus mask excluding the high-resolution hairpin-
629 containing RNAP (N=8 classes). The relative distribution of classes representing either
630 the 26, 29, 31 bp inter-RNAP distances, or other classes consisting of only a single well-
631 refined RNAP was then determined via manual inspection of the density maps and
632 comparison to the 26, 29, and 31 bp maps constructed prior.
633
634 *Model building and refinement*
635 The RNAP-EcoRI* initial model was derived from PDB 6RIP[35]. The model was manually
636 fit into the cryo-EM density maps using ChimeraX[37], modified in Coot/PyMOL/Isolde[61–63],
637 and followed by real-space refinement using PHENIX[64]. The nucleic acid sequence was
638 fixed in the same reference frame as the EcoRI* cognate binding sequence, which
639 matched the cryo-EM map along the rest of the nucleic acid sequence. For real-space
640 refinement, rigid body refinement was followed by all-atom refinement with
641 Ramachandran and secondary structure restraints. For the RNAP-RNAP complexes, the
642 same refinement methodology was applied with the following modifications. The RNAP
643 coordinates from the RNAP-EcoRI* complex were used as the starting model for the
644 RNAP-RNAP complexes. Additionally, since we were unable to determine the nucleic
645 acid sequence by visual inspection in any of the RNAP-RNAP states (likely due to
646 heterogeneity in the positioning of collision sites along the template), poly-G and poly-C
647 sequences were used during modeling building.
648
649 *Computational modeling of collision interface*
650 The nucleic acids of the RNAP-EcoRI* collision were converted to homopolymer models
651 to make modeled transcription steps (stepwise rotation and translation along the DNA
652 axis) easier. The RNAP coordinates were transformed in three-dimensional space by
653 stepwise alignment of the downstream DNA nucleotides to the same sequences plus one
654 step further downstream. SES overlap volume was calculated by summation of the SES
655 volumes of the individual RNAP and EcoRI* proteins and subtracting the SES volumes
656 for the combined RNAP and EcoRI. In the case where the RNAP and EcoRI* volumes

are overlapping, their combined surface volumes will be less than the sum of the individual protein volumes yielding a positive SES overlap volume. SES overlaps were re-scaled such that the 39 nt position equaled zero. Alignment and calculation of SES volumes were performed in ChimeraX[37].

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

The core modules of the different RNAP structures were superimposed in PyMOL by aligning the following the α-carbons: all modeled residues from the α subunit, residues 3:27, 142:152, 445:455, 520:713, 786:828 from the β subunit, residues 343:368, 412:524, 530:552, 569:701, 720:786 from the β' subunit, and all modeled residues from the ω subunit.

RNAP EC (PDB: 6RH3)[35] was used as the reference structure (swivel angle = 0°) for swivel angle measurements. To determine the swivel angle for a given RNAP structure, residues comprising the swivel module from the RNAP structure and the reference were selected (β-subunit AA 1241:1341; β'-subunit AA 1:342, 369:420, 787:930, 1135:1375). The rotation axis and rotation angle transforming one swivel module into the other were then computed using the PyMOL script draw_rotation_axis.py (https://pymolwiki.org/index.php/RotationAxis). For DNA deformation-RNAP swiveling coupling analysis, the same method was applied. For symmetry expanded particles in the head-on collision state, the swivel angle of only the RNAP contained in the refinement mask was analyzed.

DNA displacement calculation

Generation of a frame of reference was necessary to orient downstream DNA displacement on different axes. To accomplish this, the center of masses of twelve downstream nucleotides of the active RNAP EC (PDB: 6rh3)[35] were extracted using PyMOL[62]. PCA was performed on their 3D coordinates to generate the PC1 and PC2 orthoplanes along the long axis of the downstream DNA. Since this definition of PC1 and PC2 is due to inherent variation of the center of mass of the nucleotides downstream of the active site in the reference frame RNAP EC (PDB: 6rh3)[35], the plane definition is arbitrary, meaning that the definition of which axis is most responsible for variance (PC1) is descriptive of the RNAP reference structure variability, not descriptive of the variance between the different structures being compared in this study. For each query structure (consensus structures and 3DFlex frames), the structural cores of the RNAPs were aligned to the structural core of the active RNAP EC (PDB: 6rh3)[35], and displacement of the center of mass of each nucleotide was then projected onto PC1 and PC2, summed across the sequence of twelve nucleotides, and averaged.

Biochemical Assays

RNAP-EcoRI Collided Complex*

703 *Backtracking assays*
704 5' Cy3-labeled RNA oligonucleotides corresponding to zero to four nucleotides of
705 backtracking (based on the RNAP starting position along the nucleic acid sequence) were
706 synthesized (Dharmacon) and used as a ladder to verify the RNA length. For the
707 backtracking assay, the scaffold was assembled as described previously except for
708 omitting the EcoRI* binding step before gel filtration purification. Twenty-fold excess of
709 EcoRI* (if applicable) and 1 μM Gre factors (if applicable) were added to the assembled
710 complex and incubated at RT for 20 minutes before addition of 0.5 mM NTPs for 1 minute
711 in a final volume of 15 μL. To process the RNA samples, 1.5 μL of Turbo DNase
712 (ThermoScientific) was added and incubated for 15 minutes at 37º C, 1.5 μL of proteinase
713 K (New England Biolabs) and 50 mM EDTA for 15 minutes at RT, and finally an equal
714 volume of 2x RNA loading dye (1x components - 47.5% formamide, 0.01% SDS, 0.01%
715 bromophenol blue, 0.5 mM EDTA) was added to the samples. Samples were boiled at
716 95º C for 5 minutes and separated on an 18% TBE-Urea PAGE gel and visualized using
717 Cy3-fluorescence on a Typhoon gel imager. Three biological replicates of this assay were
718 performed on different days.
719
720 *EcoRI* bypass assays*
721 RNAP-EcoRI* containing scaffolds were purified as described previously. Samples were
722 then diluted into heparin-free buffer to a final concentration of ~200 nM as quantified by
723 Cy3 RNA fluorescence. Samples were then diluted two-fold into a mixture of low salt
724 transcription buffer supplemented with varying KOAc concentrations and 0.5 mM NTPs.
725 These reactions were incubated at 37º C for 30 minutes and then quenched by addition
726 of 2X RNA loading dye (95% formamide, 0.02% SDS, 0.01% bromophenol blue, 1 mM
727 EDTA) supplemented with KOAc to equalize the final salt concentration between
728 conditions. For analysis, the intensity of the roadblock and scaffold band were
729 background subtracted, measured in FIJI[65], and normalized relative to each other
730 (roadblock intensity divided by scaffold intensity, and additionally rescaled for each salt
731 titration gel so that the 50 mM normalized roadblock intensity was set to E_{RB} = 1, and the
732 background intensity was set to E_{RB} = 0). E_{RB} calculated at 300 mM KOAc (because that
733 salt concentration had the largest dynamic range when all the factor addition experiments
734 were compared) was rescaled so that the average value of the EcoRI* ("normalized E_{RB}")
735 was equivalent to 1 for each sample condition (factor addition experiments in **Fig. 3C** and
736 stability mutants experiments in **Fig. 3F**) to adjust for observed batch effects when
737 different NTP stocks were used.
738
739 *RNAP Head-On Collided Complex*
740 *Complex stability assay*
741 RNAP-RNAP head-on complexes were assembled and purified for DNA templates with
742 and without the YnaJ-UspE hairpin as previously described in low salt transcription buffer
743 supplemented with 0.1 mg/ml heparin. Samples were concentrated using a 100 kDa
744 MWCO concentrator (Amicon). Control samples to measure total DNA amount loaded
745 per well were generated by treating samples with 1 μL proteinase K (New England
746 Biolabs) per 10 μL reaction volume supplemented with 25 mM EDTA (final concentration)
747 for over one hour at room temperature. 1 mM NTPs were added to initiate transcription
748 restart, and aliquots of the reaction were quenched with addition of 25 mM EDTA (final

749 concentration) at the specified time interval. Samples were run on a 2% 0.5X TBE
750 agarose gel after being diluted with a 6x stock of 2M sucrose, and the resulting gel was
751 stained with SYBR safe to visualize the DNA. Lane profiles were extracted using a 0.5x
752 lane width window. Signal was summed in the 400-1000 bp range, then normalized
753 between the -NTP control (corresponding to 100% RNAP-RNAP complexes) and the
754 proteinase K control (corresponding to background signal) to calculate fraction intact
755 complex. Four different experimental replicates were performed. Bi-exponentials were
756 fitted through the data using the scipy python module[66] to approximate the half-life of the
757 collided complex, and 95% confidence intervals are displayed in the plot.
758
759 *SEnd-seq*
760 RNAP-RNAP head-on complexes were assembled and purified as previously described
761 in low salt transcription buffer supplemented with 0.1 mg/ml heparin. For single promoter
762 templates, the same general protocol was followed with a polymerase:DNA ratio of 2:1
763 instead of 4:1 as used for the double promoter templates. Samples were concentrated
764 using a 100 kDa MWCO concentrator (Amicon). +PK labeled samples were treated with
765 proteinase K (New England Biolabs) and 25 mM EDTA for over one hour at room
766 temperature. 1 mM NTPs were added and incubated at RT for 1 minute before quenching
767 with 25 mM EDTA. For complex pulldown, samples were washed and incubated with
768 streptavidin M-280 Dynabeads (Thermo Fisher Scientific) before being pulled down via a
769 magnetic separation rack (Permagen), then washed twice with 200 µL low salt buffer
770 supplemented with 25 mM EDTA and 0.1 mg/mL heparin. The complexes were then
771 treated with proteinase K for 1 hour at 37 °C to elute RNAs off the beads followed by
772 overnight incubation at 4°C. The reaction was purified using 2.2x vol of Agencourt
773 RNAClean XP beads (Beckman Coulter, A63987). After elution, nucleic acids were
774 treated with 0.5 µl TURBO DNase (Life Technologies, AM2238) to remove template DNA.
775 The remaining RNA was further purified via phenol-chloroform-isoamyl alcohol
776 (Invitrogen, 15593031) extraction (25:24:1, v/v/v) and recovered by ethanol precipitation.
777 The RNA pellet was resuspended, mixed with 1 µL of spike-in RNA, and subsequently
778 ligated to a 3' RNA adapter for SEnd-seq library preparation, as previously described[17,46].
779 Sequencing was performed on an Illumina MiSeq instrument in a paired-end mode (150
780 nt x2).
781 Reads were barcode sorted to the correct template and experimental condition.
782 The separate 5' and 3' end reads of each RNA sequence were aligned to both the forward
783 and reverse candidate RNA sequence using a local pairwise aligner in the Biopython
784 module[67] (match_score=2, mismatch_score=-1, open_gap_score=-5,
785 extend_gap_score=-0.5). Alignment scores were calculated by the summation of the two
786 best candidate positions. The higher scoring alignment was taken, as long as the distance
787 between the 5' and 3' ends was greater than 15 nucleotides, as all the RNAs should be
788 at least 17 nucleotides from the A/G/UTP only transcription initiation walkout before
789 subsequent purification. Unmapped reads either had the wrong direction of 5' and 3'
790 alignment or were too short. Furthermore, RNAs that did not initiate from within a 10-
791 nucleotide range near the transcription start site were filtered out. These lost reads upon
792 filtering could come from RNAP initiation from the end of the DNA template, RNA
793 degradation of the 5' end, or non-templated addition at the 3' end[68].
794

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). J.W.W. acknowledges support from National Institutes of Health NRSA training grant T32GM066699. S.L. acknowledges support from National Institutes of Health grant R01GM149862, and G.M.A acknowledges support from grant R35GM161251. 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). G.M.A is a Biohub investigator.

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.

845 **References**

846 1. Le, T.T., and Wang, M.D. (2018). Molecular Highways—Navigating Collisions of
847 DNA Motor Proteins. J. Mol. Biol. *430*, 4513–4524.
848 https://doi.org/10.1016/j.jmb.2018.08.006.
849 2. Sneppen, K., Dodd, I.B., Shearwin, K.E., Palmer, A.C., Schubert, R.A., Callen,
850 B.P., and Egan, J.B. (2005). A Mathematical Model for Transcriptional
851 Interference by RNA Polymerase Traffic in Escherichia coli. J. Mol. Biol. *346*,
852 399–409. https://doi.org/10.1016/j.jmb.2004.11.075.
853 3. Klumpp, S., and Hwa, T. (2008). Stochasticity and traffic jams in the transcription
854 of ribosomal RNA: Intriguing role of termination and antitermination. Proc. Natl.
855 Acad. Sci. U. S. A. *105*, 18159–18164. https://doi.org/10.1073/pnas.0806084105.
856 4. Helmrich, A., Ballarino, M., and Tora, L. (2011). Collisions between Replication
857 and Transcription Complexes Cause Common Fragile Site Instability at the
858 Longest Human Genes. Mol. Cell *44*, 966–977.
859 https://doi.org/10.1016/j.molcel.2011.10.013.
860 5. Sankar, T.S., Wastuwidyaningtyas, B.D., Dong, Y., Lewis, S.A., and Wang, J.D.
861 (2016). The nature of mutations induced by replication–transcription collisions.
862 Nature 2016 535:7610 *535*, 178–181. https://doi.org/10.1038/nature18316.
863 6. Prado, F., and Aguilera, A. (2005). Impairment of replication fork progression
864 mediates RNA polII transcription-associated recombination. EMBO Journal *24*,
865 1267–1276. https://doi.org/10.1038/sj.emboj.7600602.
866 7. Pomerantz, R.T., and O'Donnell, M. (2008). The replisome uses mRNA as a
867 primer after colliding with RNA polymerase. Nature *456*, 762.
868 https://doi.org/10.1038/nature07527.
869 8. Zhang, H., Shi, Z., Banigan, E.J., Kim, Y., Yu, H., Bai, X.-C., and Finkelstein
870 Correspondence, I.J. (2023). CTCF and R-loops are boundaries of cohesin-
871 mediated DNA looping. Mol. Cell *83*, 2856-2871.e8.
872 https://doi.org/10.1016/j.molcel.2023.07.006.
873 9. Epshtein, V., and Nudler, E. (2003). Cooperation between RNA polymerase
874 molecules in transcription elongation. Science (1979). *300*, 801–805.
875 https://doi.org/10.1126/science.1083219.
876 10. Wee, L.M., Tong, A.B., Jose Florez Ariza, A., Cañ ari-Chumpitaz, C., Grob, P.,
877 Nogales, E., and Bustamante, C.J. (2023). A trailing ribosome speeds up RNA
878 polymerase at the expense of transcript fidelity via force and allostery. Cell *186*,
879 1244-1262.e34. https://doi.org/10.1016/j.cell.2023.02.008.
880 11. Wang, C., Molodtsov, V., Firlar, E., Kaelber, J.T., Blaha, G., Su, M., and Ebright,
881 R.H. (2020). Structural basis of transcription-translation coupling. Science (1979).
882 *369*, 1359–1365. https://doi.org/10.1126/science.abb5317.
883 12. Webster, M.W., Takacs, M., Zhu, C., Vidmar, V., Eduljee, A., Abdelkareem, M.,
884 and Weixlbaumer, A. (2020). Structural basis of transcription-translation coupling
885 and collision in bacteria. Science (1979). *369*, 1355–1359.
886 https://doi.org/10.1126/science.abb5036.
887 13. Zhang, J., Lu, G., Zhou, W., Molodtsov, V., Cheng, Z., Zhang, Y., Li, R., Jiang, Z.,
888 Yang, M., Li, L., et al. (2026). Structural basis of translation in transcription-
889 translation coupling. bioRxiv, 2024.03.13.584796.
890 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.

936 27. Finkelstein, I.J., Visnapuu, M.-L., and Greene, E.C. (2010). Single-molecule
937 imaging reveals mechanisms of protein disruption by a DNA translocase. Nature
938 *468*, 983–987. https://doi.org/10.1038/nature09561.
939 28. Frederick, C.A., Grable, J., Melia, M., Samudzi, C., Jen-Jacobson, L., Wang, B.C.,
940 Greene, P., Boyer, H.W., and Rosenberg, J.M. (1984). Kinked DNA in crystalline
941 complex with EcoRI endonuclease. Nature 1984 309:5966 *309*, 327–331.
942 https://doi.org/10.1038/309327a0.
943 29. Doughty, B., Kazer, S.W., and Eisenthal, K.B. (2011). Binding and cleavage of
944 DNA with the restriction enzyme EcoR1 using time-resolved second harmonic
945 generation. Proc. Natl. Acad. Sci. U. S. A. *108*, 19979–19984.
946 https://doi.org/10.1073/PNAS.1115498108.
947 30. Punjani, A., and Fleet, D.J. (2023). 3DFlex: determining structure and motion of
948 flexible proteins from cryo-EM. Nat. Methods *20*, 860.
949 https://doi.org/10.1038/S41592-023-01853-8.
950 31. Komissarova, N., and Kashlev, M. (1997). RNA polymerase switches between
951 inactivated and activated states by translocating back and forth along the DNA
952 and the RNA. Journal of Biological Chemistry *272*, 15329–15338.
953 https://doi.org/10.1074/jbc.272.24.15329.
954 32. Nudler, E., Kashlev, M., Nikiforov, V., and Goldfarb, A. (1995). Coupling between
955 transcription termination and RNA polymerase inchworming. Cell *81*, 351–357.
956 https://doi.org/10.1016/0092-8674(95)90388-7.
957 33. Borukhov, S., Sagitov, V., and Goldfarb, A. (1993). Transcript cleavage factors
958 from E. coli. Cell *72*, 459–466. https://doi.org/10.1016/0092-8674(93)90121-6.
959 34. Zhang, J., and Landick, R. (2009). Substrate Loading, Nucleotide Addition, and
960 Translocation by RNA Polymerase. RNA Polymerases as Molecular Motors, 206–
961 235. https://doi.org/10.1039/9781847559982-00206.
962 35. Abdelkareem, M., Saint-André, C., Takacs, M., Papai, G., Crucifix, C., Guo, X.,
963 Ortiz, J., and Weixlbaumer, A. (2019). Structural Basis of Transcription: RNA
964 Polymerase Backtracking and Its Reactivation. Mol. Cell *75*, 298-309.e4.
965 https://doi.org/10.1016/j.molcel.2019.04.029.
966 36. Guo, X., Myasnikov, A.G., Chen, J., Takacs, M., Schultz, P., and Weixlbaumer, A.
967 (2018). Structural Basis for NusA Stabilized Transcriptional Pausing. Mol. Cell *69*,
968 816–827. https://doi.org/10.1016/j.molcel.2018.02.008.
969 37. Meng, E.C., Goddard, T.D., Pettersen, E.F., Couch, G.S., Pearson, Z.J., Morris,
970 J.H., and Ferrin, T.E. (2023). UCSF ChimeraX: Tools for structure building and
971 analysis. Protein Science *32*, e4792. https://doi.org/10.1002/pro.4792.
972 38. Kang, J.Y., Mishanina, T. V., Landick, R., and Darst, S.A. (2019). Mechanisms of
973 transcriptional pausing in bacteria. J. Mol. Biol. *431*, 4007.
974 https://doi.org/10.1016/J.JMB.2019.07.017.
975 39. Kang, J.Y., Mishanina, T. V., Bellecourt, M.J., Mooney, R.A., Darst, S.A., and
976 Landick, R. (2018). RNA polymerase accommodates a pause RNA hairpin by
977 global conformational rearrangements that prolong pausing. Mol. Cell *69*, 802.
978 https://doi.org/10.1016/j.molcel.2018.01.018.
979 40. Delbeau, M., Omollo, E.O., Froom, R., Koh, S., Mooney, R.A., Lilic, M., Brewer,
980 J.J., Rock, J., Darst, S.A., Campbell, E.A., et al. (2023). Structural and functional
981 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. 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.
50. 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.
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.

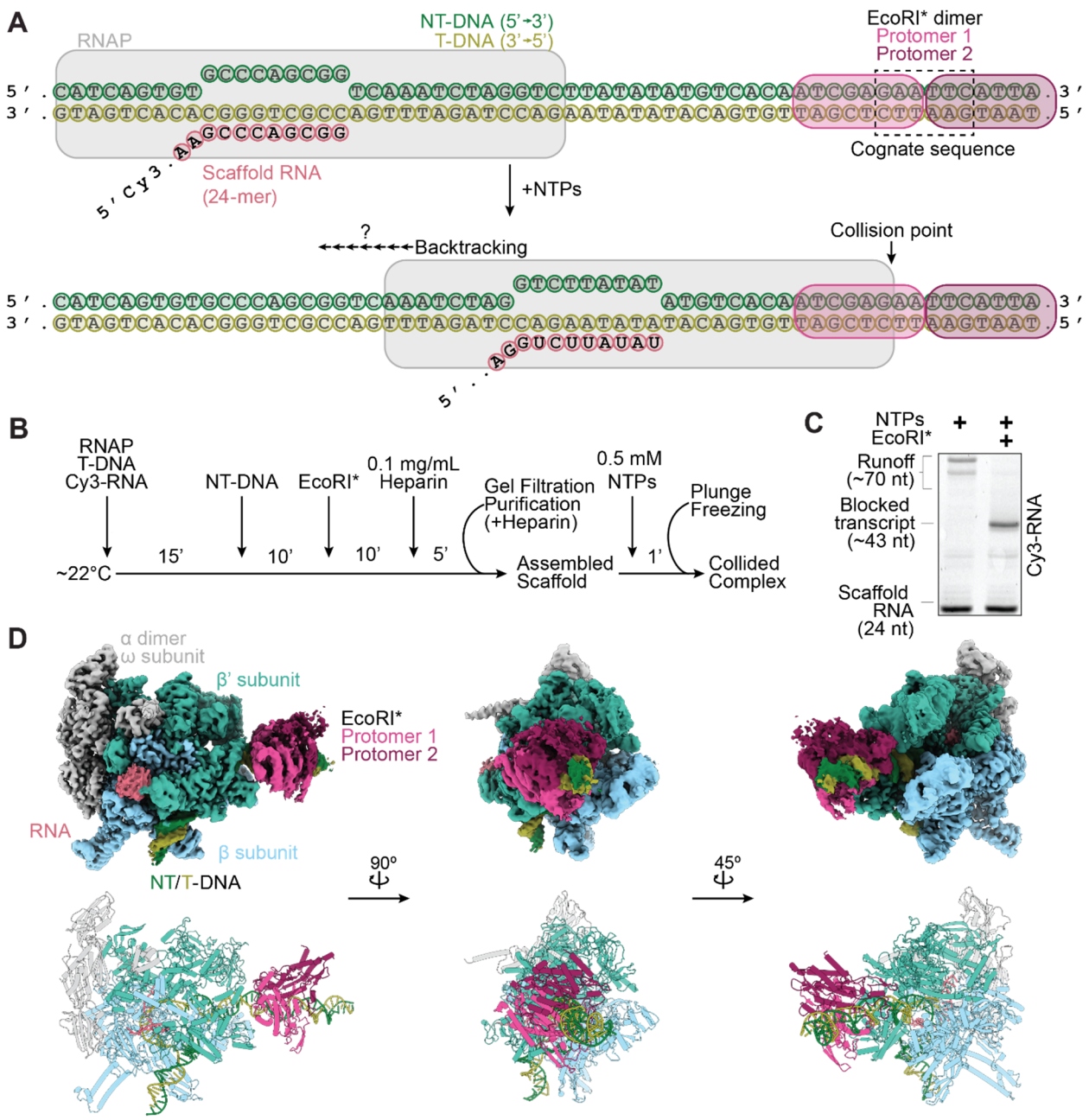

Figure 1. Reconstituting transcriptional collision with a stationary roadblock.
[A] Synthetic nucleic acid scaffold design to assemble E. coli RNAP and EcoRI* onto a DNA template. NT-DNA, non-template DNA strand; T-DNA, template DNA strand.
[B] Scaffold assembly and purification scheme.
[C] Representative 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.

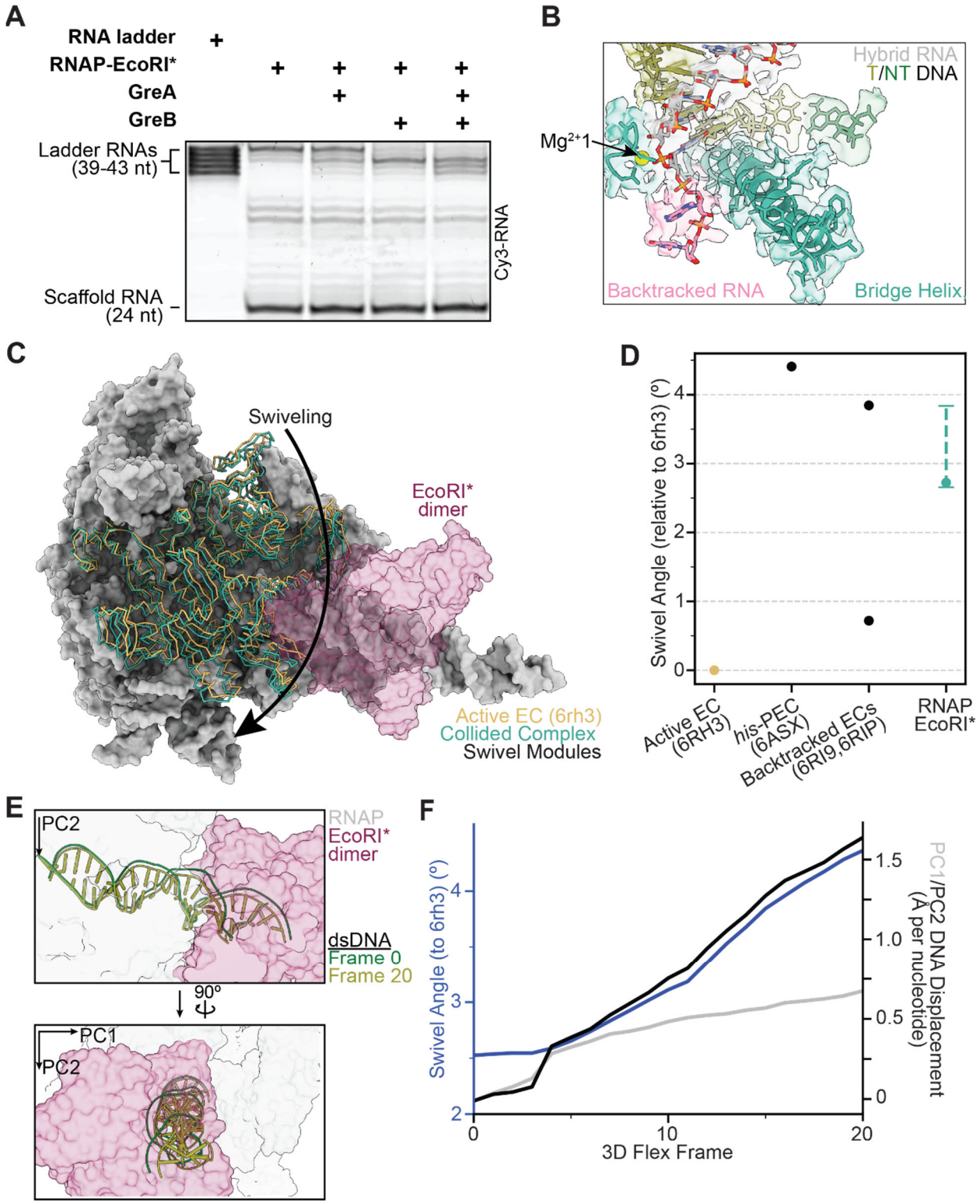

Figure 2. EcoRI* collided complex is backtracked and moderately swiveled. **[A]** Representative gel of in vitro transcription assay in the presence and absence of Gre factors, with an RNA ladder corresponding to potential transcript lengths within the collided complex.

[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.

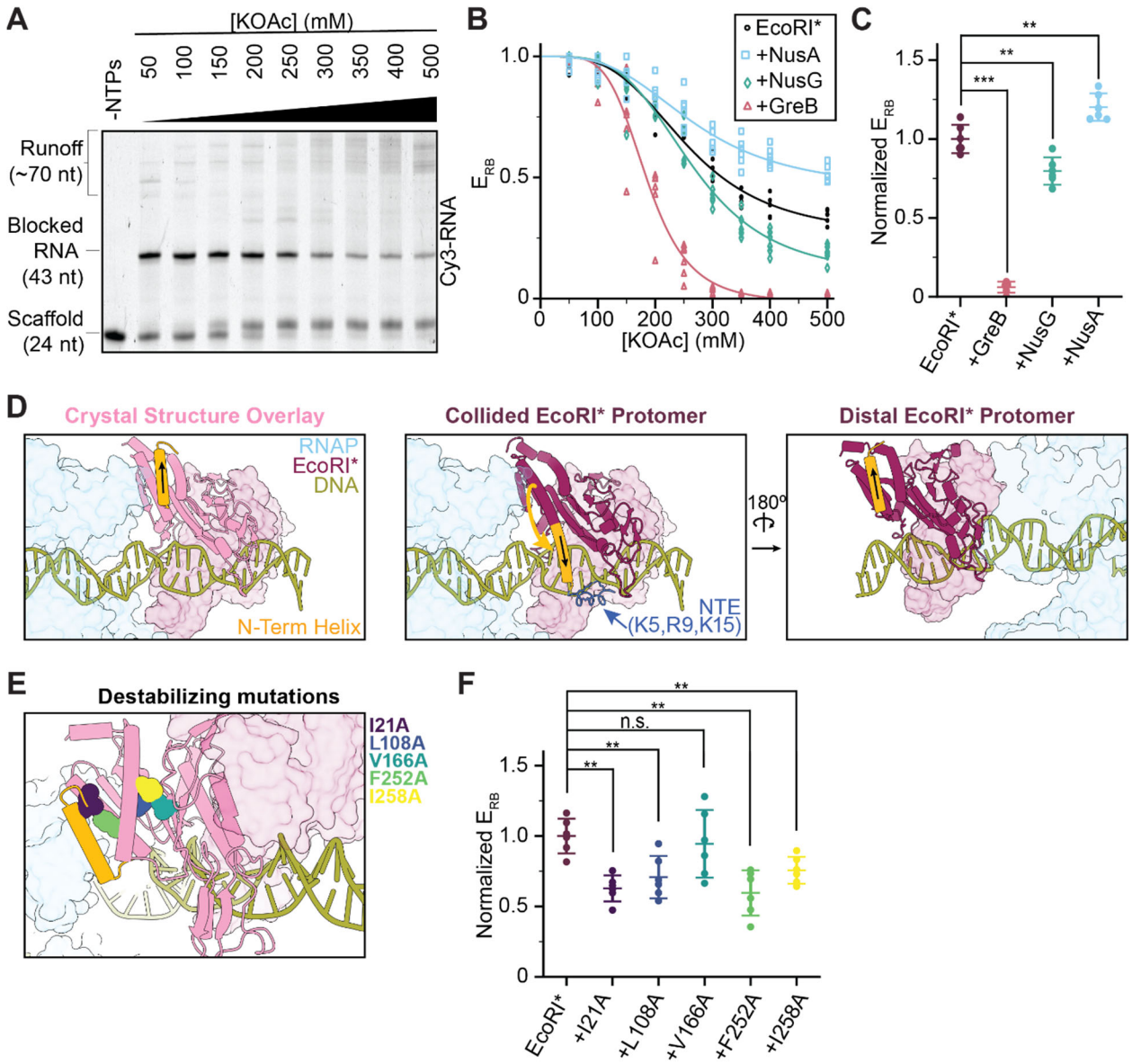

Figure 3. Factors that affect bypass efficiency.
[A] Representative EcoRI* bypass assay. The RNA lengths under different conditions are visualized via Cy3 fluorescence.
[B] Fraction of blocked RNAP in the absence and presence of factors (GreB, NusA, NusG). Raw measurements and their sigmoidal fit for 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.

1135 **[F]** Normalized roadblock efficiencies for the destabilizing EcoRI* mutations at 300 mM
1136 KOAc. Data are shown as mean ± S.D. Individual points are also displayed (N = 6).
1137 Conditions were compared via two-tailed, unpaired t test including Welch's correction and
1138 Benjamini-Hochberg correction for multiple testing. n.s., not significant; *$p < 0.05$; **$p <$
1139 0.01; ***$p < 0.001$.

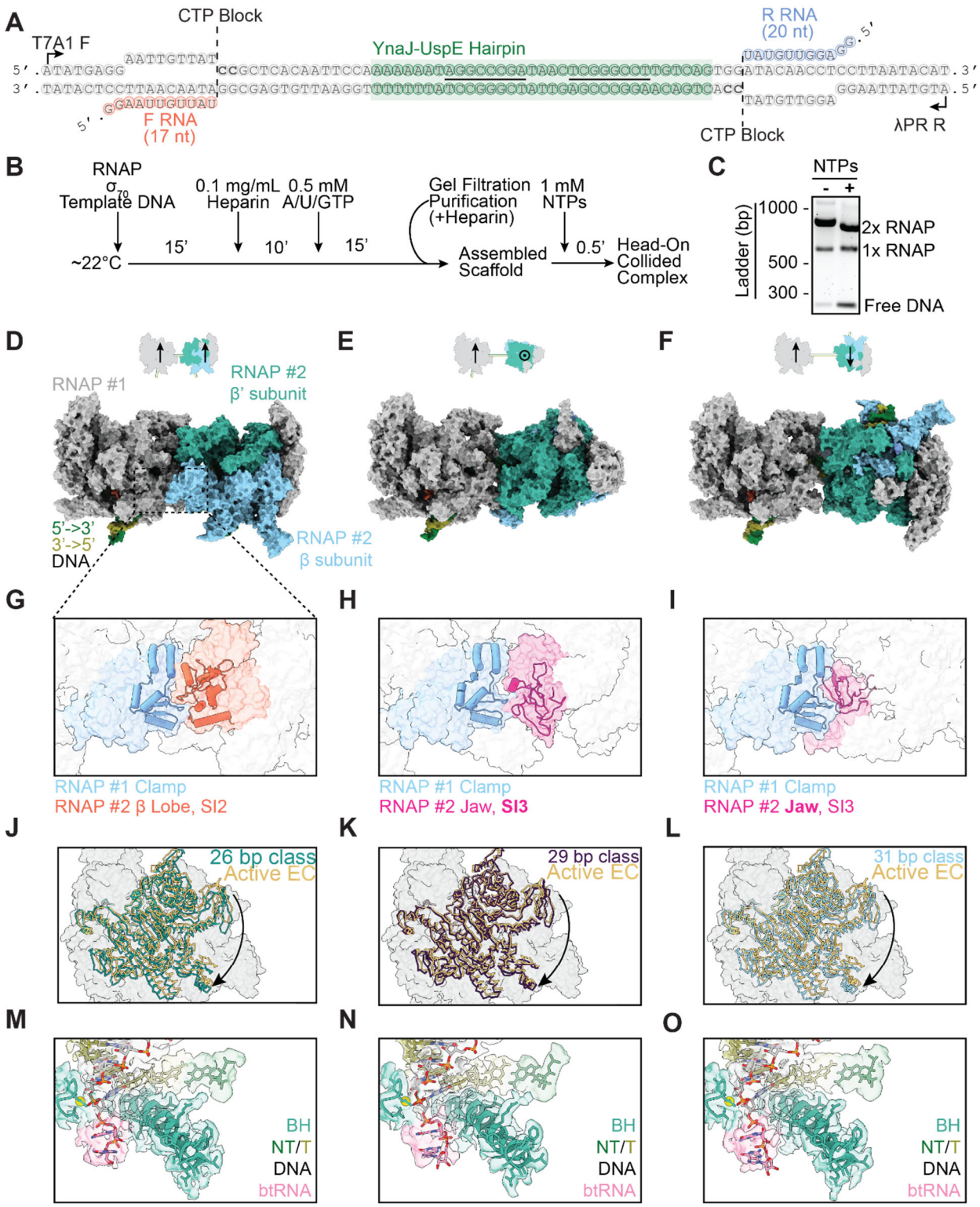

Figure 4. Reconstituting head-on RNAP-RNAP collisions.
[A] DNA template featuring two promoters, designed to stall RNAP at cytosine sites. F RNA, forward RNA; R RNA, reverse RNA. Underlined segments of the YnaJ-UspE hairpin indicate complementary regions.

[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.

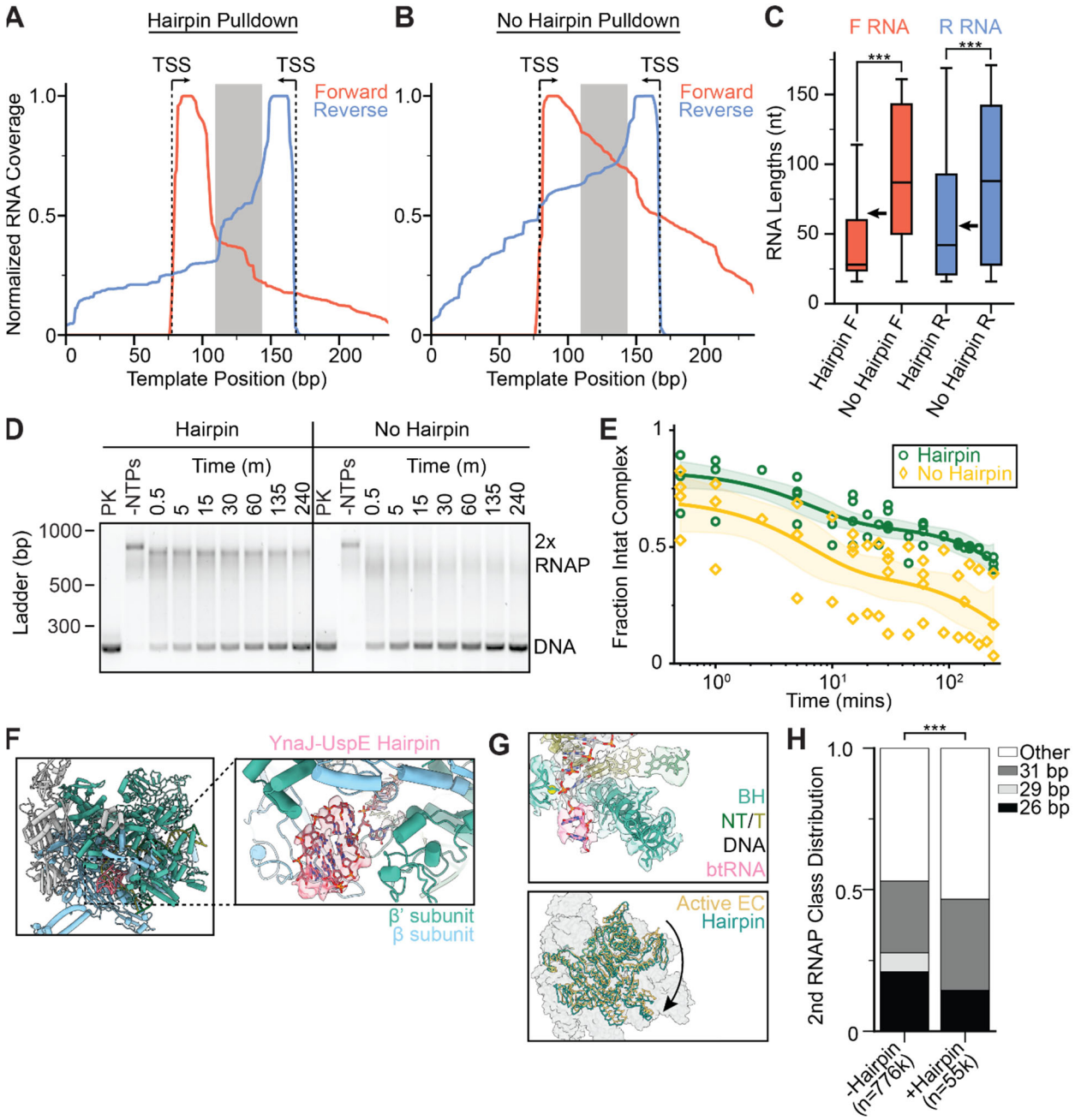

Figure 5. RNA hairpin impacts collision positioning and stability.
[A, B] SEnd-seq profiles of forward and reverse RNAs in head-on RNAP-RNAP collided complexes on a template encoding the YnaJ-UspE hairpin (**A**) or a no-hairpin template (**B**). The grey box defines the hairpin sequence region or the 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.

1167 **[D]** Native agarose gels for evaluating collided complex disassembly. PK: Proteinase K
1168 treated. Timepoints indicate time elapsed after NTP addition, followed by EDTA quench.
1169 Left: hairpin-containing template DNA; right, no-hairpin template DNA. DNA was
1170 visualized with SYBR safe.
1171 **[E]** Quantification of the fraction of intact collided complex under indicated conditions.
1172 Nucleic acid intensity values were normalized for each of four experimental replicates to
1173 the -NTPs and +Proteinase K conditions. Shaded areas indicate 95% confidence interval
1174 of the fits.
1175 **[F]** View of the hairpin-containing head-on collided RNAP structure obtained through
1176 symmetry expansion.
1177 **[G]** Views of the hairpin-containing RNAP (cryo-EM map and atomic model) highlighting
1178 the backtracked RNA (top) and rotation of the swivel module relative to the active
1179 elongation complex.
1180 **[H]** Class distributions of no-hairpin and hairpin-containing collided RNAP complexes.
1181 "Other" is defined as particles that show well-defined density for one RNAP but the density
1182 downstream of the well-defined RNAP is too low of a resolution to confidently model an
1183 RNAP. Significance was calculated via a chi-squared test. ***$p < 0.001$.

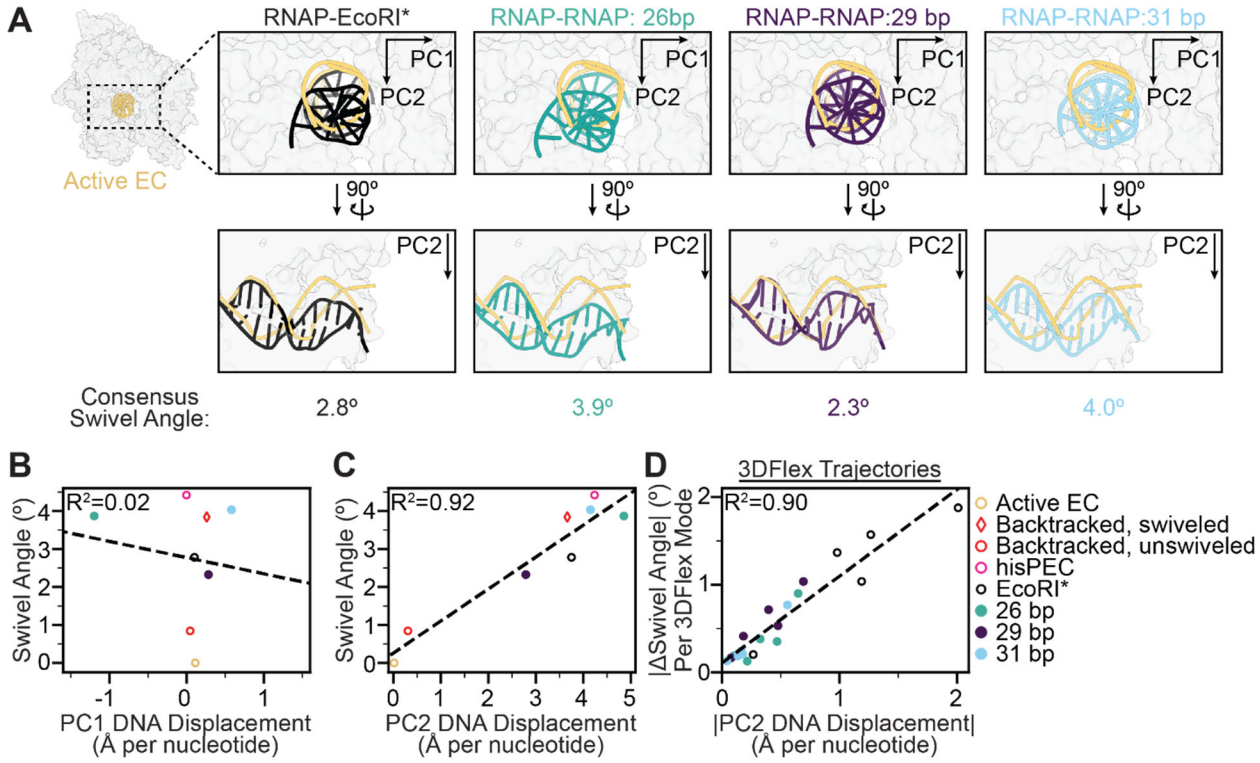

Figure 6. Inter-RNAP distance modulates deformation-swivel coupling.
[A] Downstream DNA deformation of each collision scenario compared against active EC (PDB 6RH3).
[B-C] Plots of swiveling angle versus DNA displacement along **[B]** PC1 and **[C]** PC2 of consensus collision structures (this study) 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).

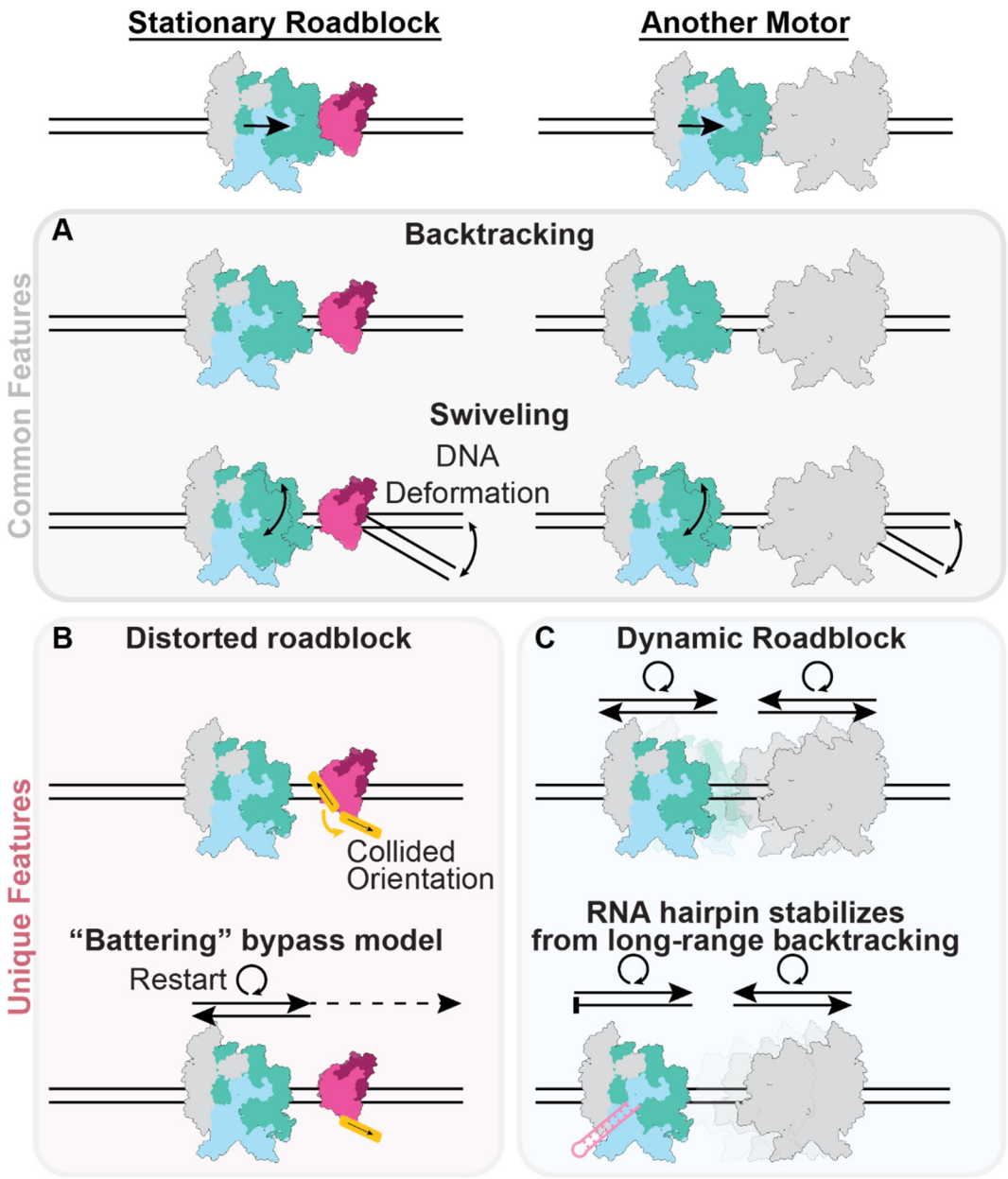

Figure 7. Structural mechanisms of transcriptional collisions.
[A] Common features of all collisions in this study. Regardless of the roadblock identity, RNAP backtracks away from the collision site and swivels into a transcriptionally inactive state, mediated by coupling between template DNA distortion 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.