Supplemental Information for

RNA polymerase accommodates a pause RNA hairpin by global conformational rearrangements that prolong pausing

Jin Young Kang[1], Tatiana V. Mishanina[2], Michael J. Bellecourt,[2] Rachel Anne Mooney[2], Seth A. Darst[1,4], Robert Landick[2,3,4,5]

[1] The Rockefeller University, 1230 York Avenue, New York, NY 10065, U.S.A.

[2] Department of Biochemistry, University of Wisconsin-Madison, Madison, WI 53706, U.S.A.

[3] Department of Bacteriology, University of Wisconsin-Madison, Madison, WI 53706, U.S.A.

Table S1. Refinement statistics. Related to Figure 1.

	*his*PEC	*his*PEC-minus-PH
Resolution (Å)[a]	3.83	5.49
RMSD		
Bond lengths (Å)	0.009	0.006
Bond angles (°)	1.058	0.983
Ramachandran		
Favored (%)	90.28	92.93
Allowed (%)	0.66	7.01
Outliers (%)	0.06	0.06
Molprobity[b]		
Clash score	4.55	10.03
Rotamer outliers (%)	0.26	0.19
Overall score	1.77	1.97

[a] Gold-standard FSC 0.143 cutoff criteria (Figure S2D) (Rosenthal and Henderson, 2003).

[b] (Richardson et al., 2008)

Table S2. Global conformational change comparing the *his*PEC and EC structures (superimposed via the structural core module). Related to Figure 3.

Structural module	residues	Total # of Cα's	Rmsd (Å)
Entire structure	αI: 7-158; 167-234 αII: 4-159; 169-232 β: 2-890; 913-1,341 β′: 16-933; 946-1,126; 1,135-1,373 ω: 2-74	3,159	2.518
Structural core module	αI: 7-158; 167-234 αII: 4-159; 169-232 β: 2-30; 140-150; 445-455; 513-832; 1,056-1,240 β′: 343-368; 421-786 ω: 2-74	1,451	0.772
Everything outside of the structural core module		1,698	3.359
β1-lobe (protrusion)	β: 31-139; 456-512	166	1.598
β2-lobe:βi4 (lobe:SI1)	β: 151-444	294	1.722
βflap:βi9 (flap-SI2)	β: 833-1,055	199	6.080
Clamp	β: 1,319-1,341 β′: 16-342; 1,318-1,344	377	3.112
β′dock	β′: 369-420	52	1.255
β′shelf/β'jaw	β′: 787-931; 1,135-1,317	328	2.746
β′SI3	β′: 946-1,126	179	4.808
Clamp/β′dock/β′shelf/β′jaw/SI3/β′C	β: 1,241-1,341 β′: 16-347; 369-420; 787-931; 946-1,126; 1,135-1,373	1,050	3.166

Table S3. Conformational changes of *his*PEC-minus-PH compared to EC and *his*PEC. Related to Figure 7.

Structural module	residues	Rmsd (Å)	
		Vs. EC	Vs. *his*PEC
Entire structure	αI: 7-158; 167-234 αII: 4-159; 169-232 β: 3-890; 913-1,341 β′: 16-933; 946-1,126; 1,135-1,373 ω: 2-74	1.403 (3,160 Cα's)	2.512 (3,158 Cα's)
Structural core module	αI: 7-158; 167-234 αII: 4-159; 169-232 β: 2-30; 140-150; 445-455; 513-713; 786-832; 1,056-1,240 β′: 343-368; 421-786 ω: 2-74	0.902 (1,374)	0.809 (1,372)
Everything outside of the structural core module	β: 31-139; 151-444; 456-512; 714-785; 833-1,055; 1,241-1,341 β′: 16-342; 369-420; 787-1,373	1.720 (1,698)	3.337 (1,698)
Everything outside of the structural core module minus SI3	β: 31-139; 151-444; 456-512; 714-785; 833-1,055; 1,240-1,341 β′: 16-342; 369-420; 787-931; 1,135-1,373	1.439 (1,517)	2.991 (1,517)
β1-lobe (protrusion)	β: 31-139; 456-512	1.362 (166)	1.261 (166)
β2-lobe:βi4 (lobe:SI1)	β: 151-444	1.602 (294)	1.550 (294)
βflap:βi9 (flap-SI2)	β: 833-1,055	1.691 (199)	5.006 (199)
Clamp	β: 1,319-1,341 β′: 16-342; 1,318-1,344	1.379 (377)	3.214 (377)
β'dock	β′: 369-420	0.823 (52)	1.208 (52)
β′shelf/β'jaw	β′: 787-931; 1,135-1,317	0.974 (145)	2.463 (145)
β′SI3	β′: 946-1,126	3.242 (179)	5.444 (179)
β′shelf:β′SI3:β'jaw	β′: 787-931; 946-1,126; 1,135-1,317	2.227 (507)	4.178 (507)
Clamp/β′dock/β'shelf/β′jaw/ SI3/β′C	β: 1241-1341 β′: 16-347; 369-420; 787-933; 948-1,126; 1,135-1,373	2.077 (608)	3.866 (608)
Clamp/β′dock/β′shelf/β′jaw/ β′C	β: 1,241-1,341 β′: 16-347; 369-420; 787-933; 1135-1,373	1.313 (871)	2.979 (871)

Pink color indicates that the Rmsd value of *his*PEC-minus-PH versus one structure (EC or *his*PEC) is >1.5 times the Rmsd value of *his*PEC-minus-PH versus the other structure.

Table S4. Conservation of positively charged residues guiding RNA duplex formation in the bacterial RNAP RNA exit channel. Related to Figure 6.

Eco position	Fraction KR[a]
	bacteria
β	
K890	1
K914	0.95
R919	0.89
R1301	1
K1303	0.87
β′	
R47	1
K50	0.99
K66	0.98
K76	1
R77	0.95
K378	0.92
K395	1
K398	0.98
R403	0.79

[a] Determined from an alignment of > 814 bacterial sequences (Lane and Darst, 2010).

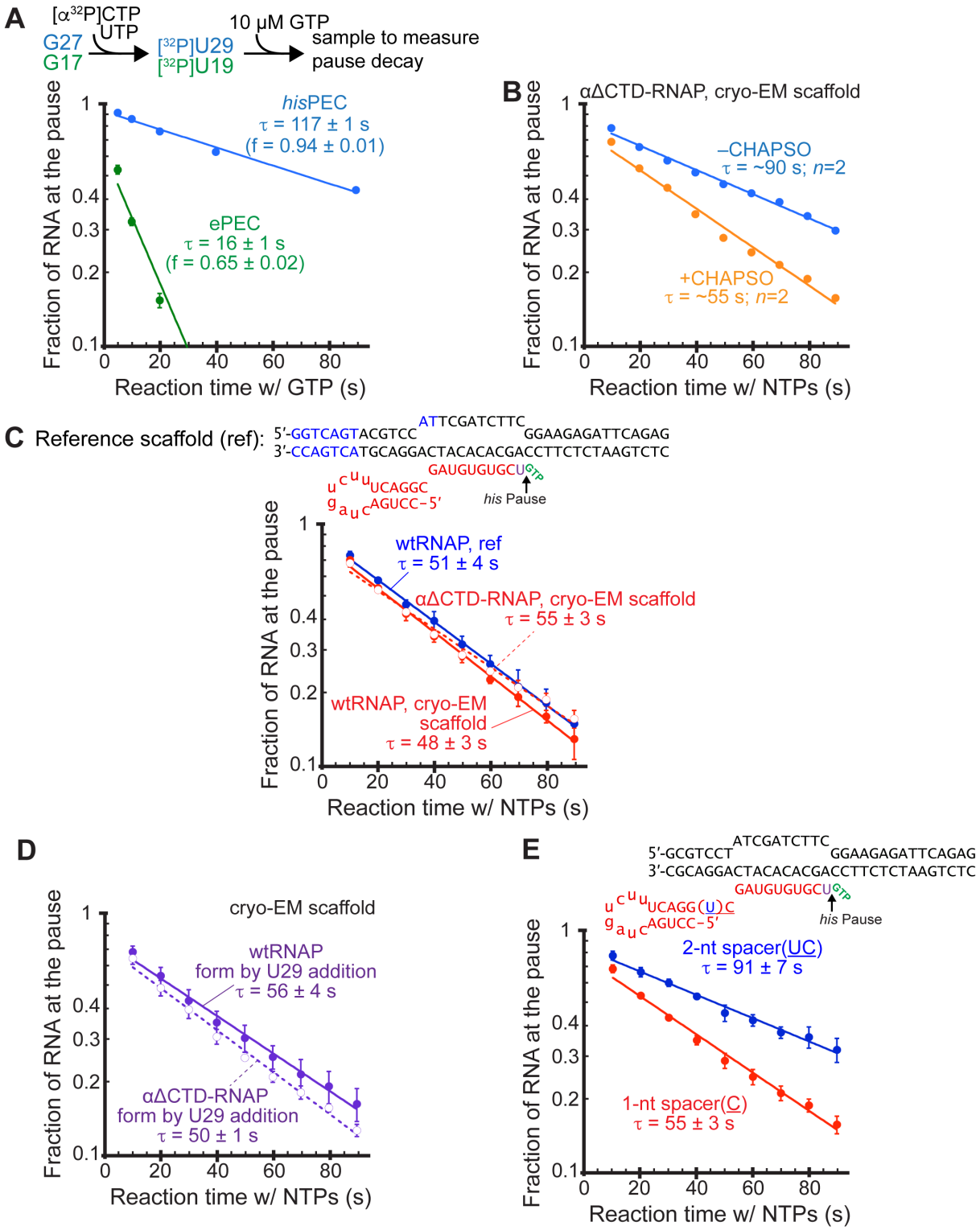

FIGURE S1. *his* pause scaffold verification. Related to Figure 1.

A. Quantitation of triplicate *in vitro* transcription time-course experiments (in the absence of CHAPSO), with representative gel image depicted in Figure 1C. See Methods for experimental details. Error is s.d. of triplicates.

B. CHAPSO (8 mM) has only a modest effect on *his* pause lifetime. ECs formed for direct reconstitution as for cryo-EM sample preparation. Data are averages of duplicates.

C. ΔαCTD-RNAP on the cryo-EM scaffold gives pause behavior indistinguishable from a reference scaffold known to adequately mimic the behavior of the *his* pause sequence (Kyzer et al., 2007). ECs formed for direct reconstitution as for cryo-EM sample preparation with CHAPSO (8 mM) added after reconstitution. Error is s.d. of triplicates.

D. *his*PECs formed on the cryo-EM scaffold by nucleotide addition of U29 in the presence of CHAPSO (8 mM). Compare to *his*PECs formed by direct reconstitution in panel C. These methods give indistinguishable pause kinetics (see also Kyzer et al., 2007). Error is s.d. of triplicates.

E. Increasing the spacing between the PH and the RNA-DNA hybrid in the *his*PEC by one nucleotide increases pause lifetime. Experiments performed in the presence of CHAPSO (8 mM). Error is s.d. of triplicates.

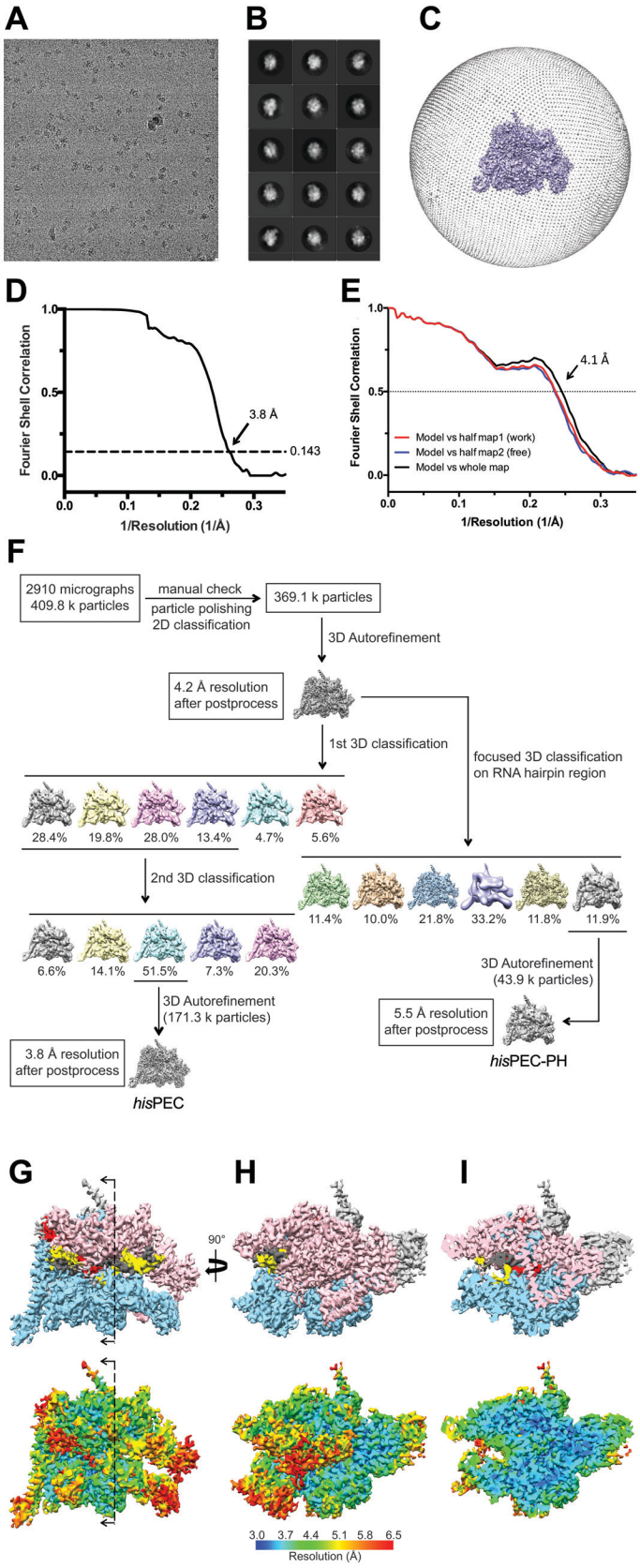

Figure S2. Cryo-EM of the *his*PEC. Related to Figure 1.

A. Representative micrograph of the *his*PEC in vitreous ice.

B. The fifteen most populated classes from 2D classification.

C. Angular distribution for *his*PEC particle projections.

D. Gold-standard FSC of the *his*PEC. The gold-standard FSC was calculated by comparing the two independently determined half-maps from RELION. The dotted line represents the 0.143 FSC cutoff, which indicates a nominal resolution of 3.8 Å.

E. FSC calculated between the refined structure and the half map used for refinement (work), the other half map (free), and the full map.

F. Data processing pipeline for the cryo-EM data. Flowchart showing the image-processing pipeline for the cryo-EM *his*PEC data starting with 2,910 dose-fractionated movies collected on a 300 keV Titan Krios (FEI) equipped with a K2 Summit direct electron detector (Gatan). Movie frames were aligned and summed using Unblur (Grant and Grigorieff, 2015). Particles were autopicked with Gautomatch (http://www.mrc-lmb.cam.ac.uk/kzhang/Gautomatch/) and manually revised from these summed images. The revised particles were polished using direct-detector-align_lmbfgs (Rubinstein and Brubaker, 2015) for subsequent 2D classification using RELION (Scheres, 2012). After 2D classification, the dataset contained 369,100 aligned particles. These particles were auto-refined in RELION using a model of the *Eco* EC (PDB ID 6ALF; [Kang et al., 2017]) as an initial 3D template. 3D classification into six classes (left branch of the flowchart) was performed on the particles using the refined model and alignment angles. Two low-resolution classes were excluded, and the remaining four classes were combined and subjected to a second 3D classification into five classes. After the second 3D classification, one class containing 51.5% of the starting particles (171,300 particles) was autorefined and post-processed in RELION, yielding the final reconstruction at 3.8 Å resolution. For the focused 3D classification (right branch of the flowchart), a soft mask that excluded the PH and nearby protein regions was generated using Chimera (Petterson et al., 2004) and RELION. The mask was used to make a subtracted particle stack in RELION with the filtered map generated in the initial autorefinement. The subtracted particles were 3D classified into six classes without alignment. Among the six classes, one class containing 11.9% of the starting particles (43,900 particles) lacked PH density. The original (unmasked) particles in this class were autorefined and post-processed in RELION, yielding the final reconstruction at 5.5 Å resolution.

G. (top) The 3.8-Å resolution cryo-EM density map of the *his*PEC is colored as follows: αI, αII, ω subunits, gray; β, cyan; β′, pink; t-strand DNA, dark gray; nt-strand DNA, yellow; RNA, red.

(bottom) Same view as (top) but colored by local resolution (Cardone et al., 2013).

H. (top) The view of (A) is rotated by 90° as shown.

(bottom) Same view colored by local resolution.

I. (top) Cross-section of the view in (B), sliced at the dotted line shown in (A).
(bottom) Same view colored by local resolution.

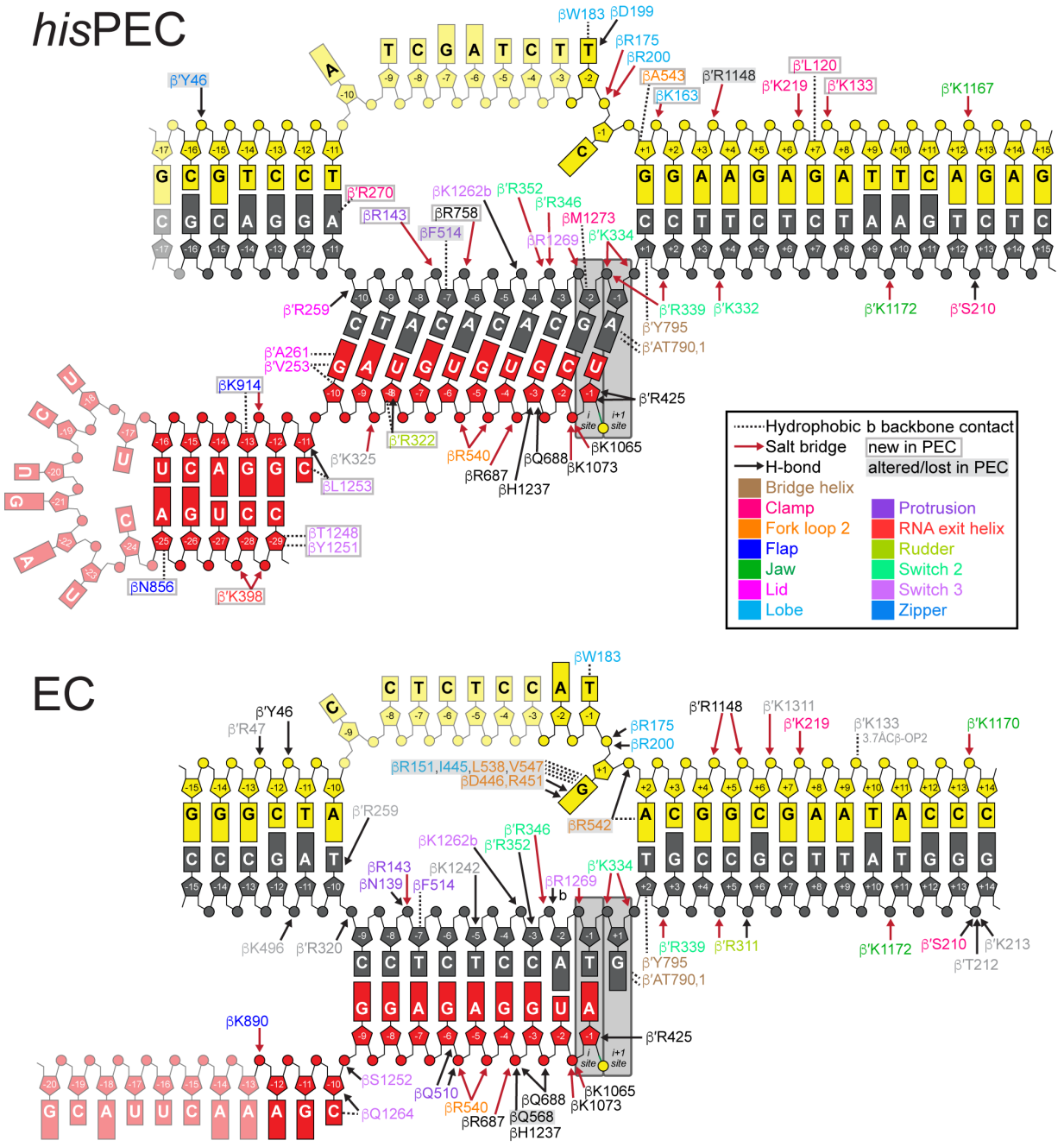

Figure S3. Schematics of RNAP-nucleic acid contacts. Related to Figure 2.
Contacts observed between RNAP and DNA or RNA observed in the *his*PEC (top) and EC (bottom) cryo-EM structures. Contacts are color-coded based on the module making the contact and whether the contact changes in the *his*PEC relative to the EC using the scheme shown in the legend in the figure. Nucleotides disordered in the structures are faded.

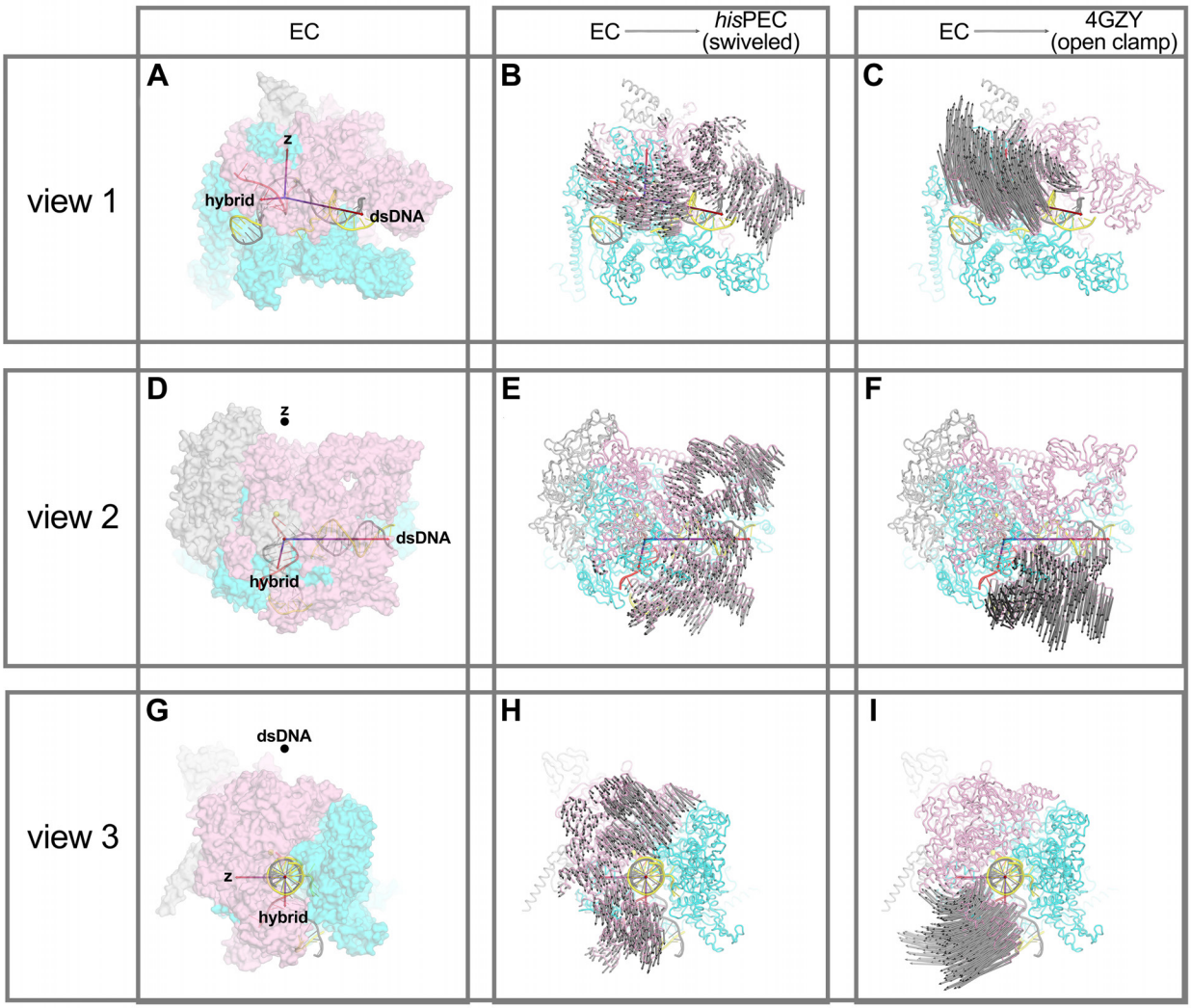

Figure S4. Comparison of EC → *his*PEC (swiveled RNAP) and EC → 4GZY (open clamp) conformational changes. Related to Figure 3.

Each panel shows a view of the *Eco* EC structure (6ALF; Kang et al, 2017). The top row (A-C) shows view1. The leftmost column (A, D, G) shows the EC structure as a transparent molecular surface (αI, αII, ω subunits, gray; β, cyan; β′, pink) with the nucleic acid scaffold shown as a cartoon (t-strand DNA, dark gray; nt-strand DNA, yellow; RNA, red). The reference axes shown correspond to the helical axis of the downstream duplex DNA (dsDNA), the helical axis of the RNA-DNA hybrid (hybrid), and 'z' orthogonal to the plane defined by the dsDNA and hybrid axes. The middle column (B, E, H) shows the EC structure as a Cα-backbone worm. The gray arrows denote the direction and distance (multiplied by a factor of 2) of the Cα(EC) → Cα(*his*PEC) conformational change (swivel module rotation). The rightmost column (C, F, I) shows the EC structure as a Cα-backbone worm, with the gray arrows denoting the direction and distance (multiplied by a factor of 2) of the Cα(EC) → Cα(4GZY) conformational change (clamp opening). The second row (D, E, F) shows view2 (looking down the 'z' axis). The third row (G, H, I) shows view3 (looking down the dsDNA axis). Panel (E) illustrates that the swivel module rotation of the *his*PEC is about an axis approximately parallel with 'z'. Panel (I) illustrates that the clamp opening of 4GZY is a rotation about an axis approximately parallel with the dsDNA axis.

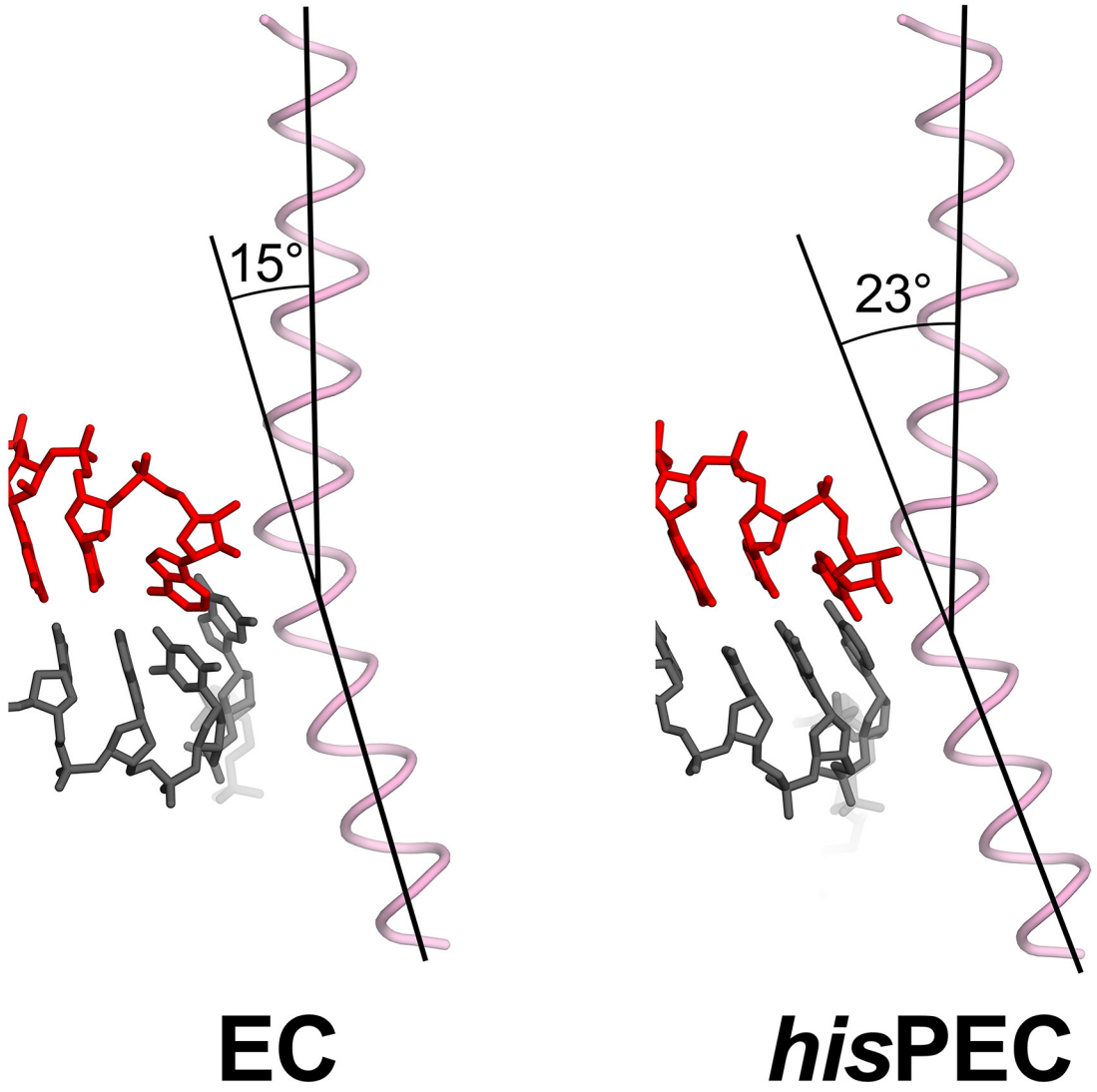

Figure S5. BH bending in the EC and *his*PEC. Related to Figure 3.

Shown is the BH (pink Cα worm) and downstream RNA-DNA hybrid (RNA, red; t-strand DNA, dark gray) for the EC (left) and *his*PEC (right). The bend or kink angle of each BH is denoted.

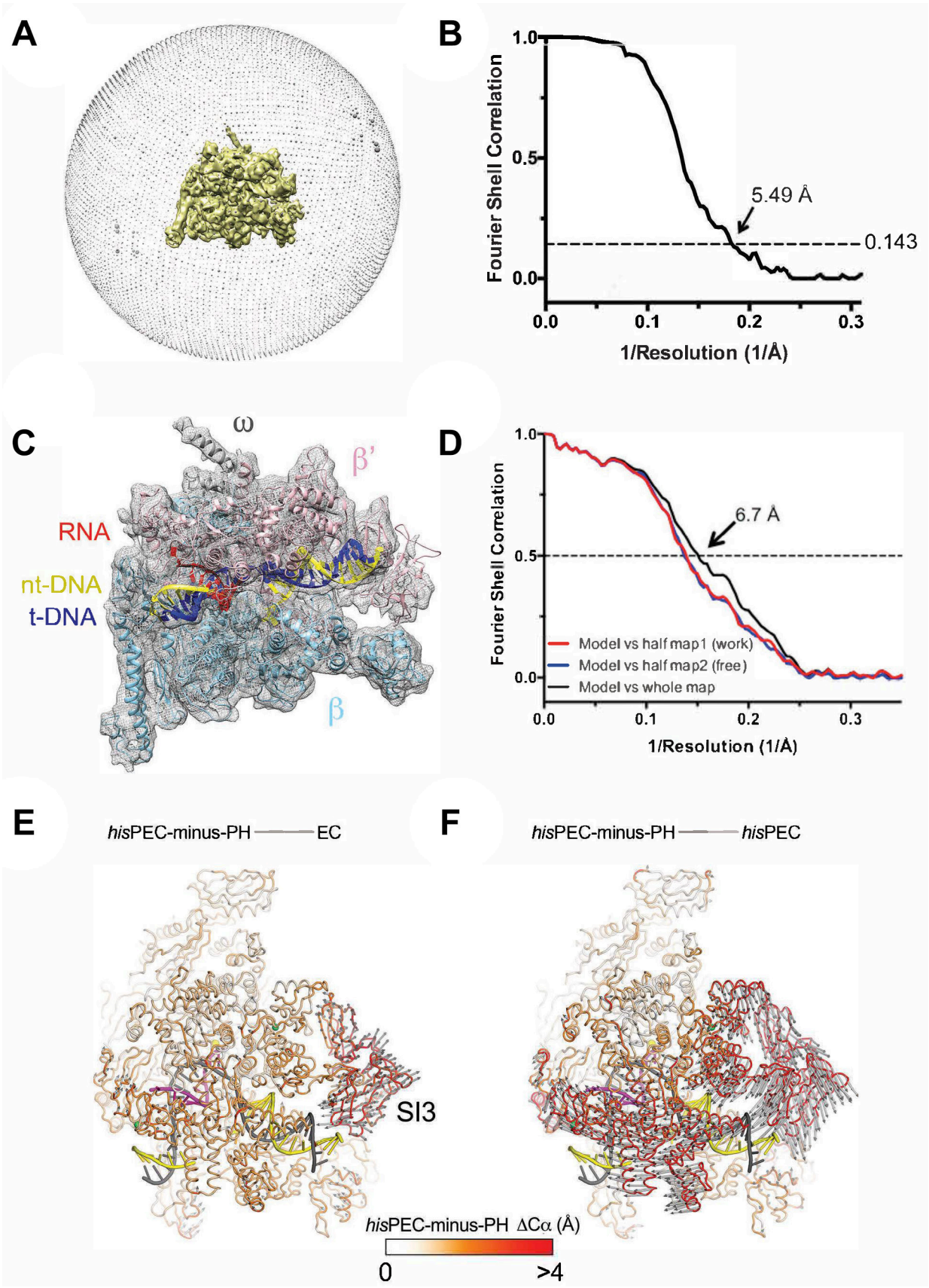

Figure S6. Cryo-EM of the *his*PEC-minus-PH. Related to Figure 7.
A. Angular distribution for *his*PEC-minus-PH particle projections.

B. Gold-standard FSC of the *his*PEC-minus-PH. The gold-standard FSC was calculated by comparing the two independently determined half-maps from RELION. The dotted line represents the 0.143 FSC cutoff, which indicates a nominal resolution of 5.5 Å. Local resolution calculations (Cardone et al., 2013) indicated that the inner core of the structure (around the RNAP active site and RNA-DNA hybrid) is determined to about 4.3 Å resolution.

C. The 5.5 Å resolution cryo-EM density map of the *his*PEC-minus-PH is shown as a gray mesh. Superimposed is the final refined model (Table S1); the RNAP and nucleic acids are shown as backbone ribbons, color-coded as labeled.

D. FSC calculated between the refined structure and the half map used for refinement (work), the other half map (free), and the full map.

E. The *his*PEC-minus-PH structure is shown as a Cα backbone worm. The nucleic acids are shown in cartoon format (t-strand DNA, dark gray; nt-strand DNA, yellow; RNA, magenta). The protein is color-coded as a ramp (color-key shown below) according to the Cα(*his*PEC-minus-PH)-Cα(EC) distance, where the two structures were superimposed via the structural core module (see Table S2). The gray arrows denote the direction and distance (multiplied by a factor of 2) for the Cα(*his*PEC-minus-PH)-Cα(EC) changes that are > 2 Å. The average ΔCα is 1.2 Å (± 0.7 Å standard deviation). The minimum and maximum ΔCα values are 0.079 and 6.2 Å, and the largest ΔCα values are exclusively in SI3.

F. The *his*PEC-minus-PH structure is shown as an Cα backbone worm. The nucleic acids are shown in cartoon format (t-strand DNA, dark gray; nt-strand DNA, yellow; RNA, magenta). The protein is color-coded as a ramp (color-key shown below) according to the Cα(*his*PEC-minus-PH)-Cα(*his*PEC) distance, after the two structures were superimposed via the structural core module (see Table S2). The gray arrows denote the direction and distance (multiplied by a factor of 2) for the Cα(*his*PEC-minus-PH)-Cα(*his*PEC) changes that are > 2 Å. The average ΔCα is 1.8 Å (± 1.7 Å standard deviation). The minimum and maximum ΔCα values are 0.047 and 13 Å, with the largest ΔCα values located in SI2.

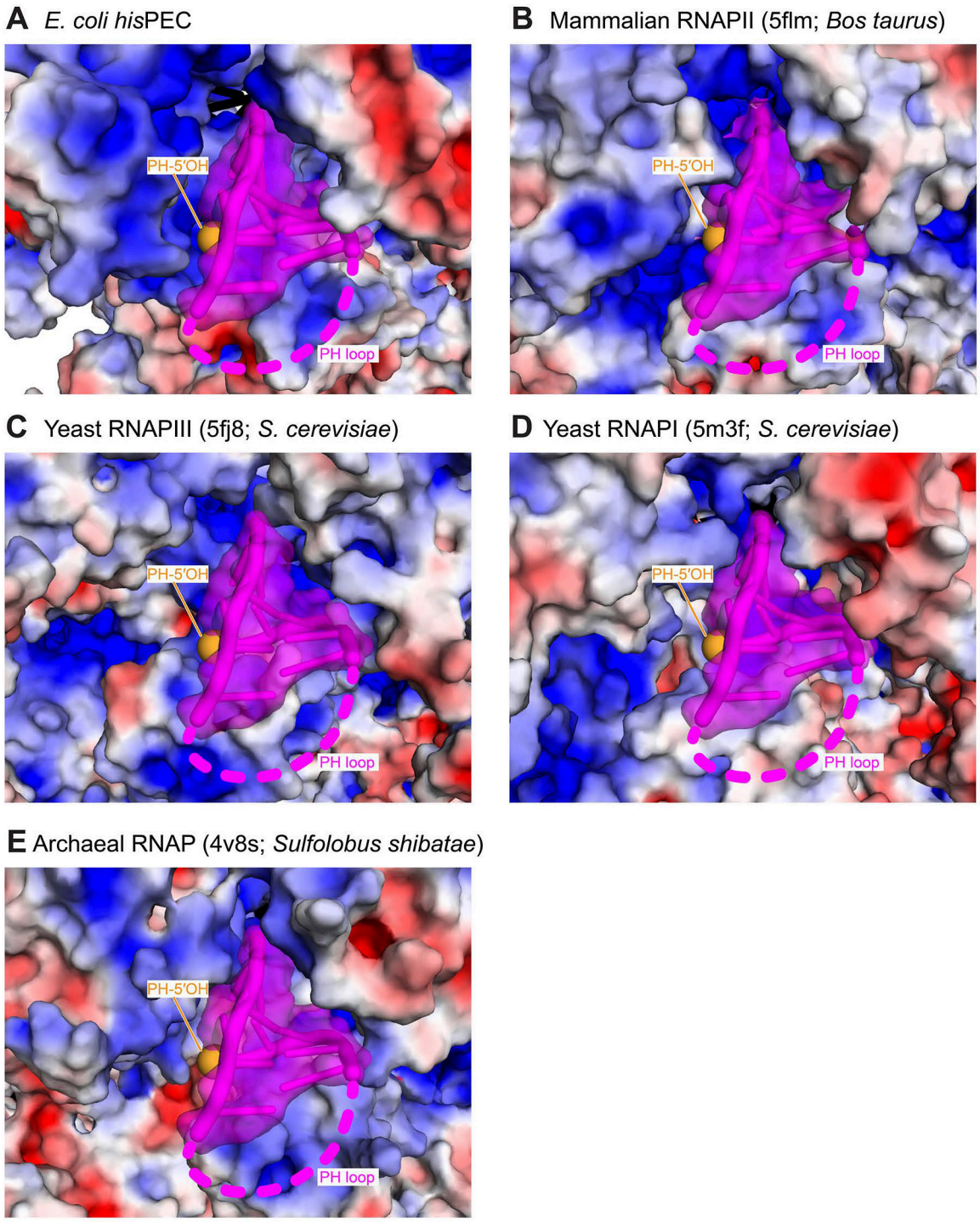

Figure S7. The *his* PH fits into positively charged RNA exit channels in diverse RNAPs. Related to Figure 6.

Views into the RNA exit channel for RNAPs containing a modeled *his* PH are shown. Although only the *his*PEC is in hairpin-stabilized pause configuration, the charge distributions would be unlikely to change dramatically if the exit channels of these RNAPs shifted to accommodate a PH (see Figure 6C).

A. View of the *his* PH in the *his*PEC (6ASX) from outside the RNA exit channel looking inward. The orange sphere shows the location of the 5′-OH at the 5′ end of the PH RNA.

B. The *his* PH modeled into mammalian RNAPII (5FLM; Bernecky et al., 2016).

C. The *his* PH modeled into yeast (*S. cerevisiae*) RNAPIII (5FJ8; Hoffmann et al., 2015).

D. The *his* PH modeled into yeast (*S. cerevisiae*) RNAPI (5M3F; Neyer et al., 2016).

E. The *his* PH modeled into archaeal (*S. shibatae*) RNAP (4V8S; Wojtas et al., 2012).