PNAS
www.pnas.org

Supplementary Information for

Structural basis of RNA polymerase I stalling at UV light-induced DNA damage

Marta Sanz-Murillo[#], Jun Xu[#], Georgiy A. Belogurov, Olga Calvo, David Gil-Carton, María Moreno-Morcillo, Dong Wang[*], Carlos Fernández-Tornero[*]

[#]These authors contributed equally to this work

[*]Corresponding authors. Email: dongwang@ucsd.edu; cftornero@cib.csic.es

This file includes:

Supplementary Methods
Supplementary References
Supplementary Table 1
Supplementary Figures 1 to 5

Supplementary Methods

***In vitro* transcription assays using purified RNA polymerases.** The transcription elongation assay with full transcription bubble was performed based on previous reported methods with slight modifications (32). Briefly, an aliquot of 5'- ^{32}P labeled RNA was annealed with a 1.5-fold amount of template DNA strand (TS) and 2.0-fold amount of non-template DNA strand (NTS) to form the RNA/DNA bubble scaffold in a buffer composed of 20 mM Tris-HCl (pH 7.5), 150 mM KCl, 5 mM DTT. An aliquot of annealed scaffold was then incubated with a five excess amount of purified Pol II , Pol I, or *E. coli* RNAP on ice for 5 min, followed by incubation at room temperature (23ºC) for 15 min to ensure the formation of Pol II or Pol I elongation complex. The transcription was chased by adding an equal volume of solution containing 150 mM NaCl, 20 mM Tris-HCl (pH 7.5), 5 mM DTT, 10 mM $MgCl_2$, and 2 mM NTP. Final reaction concentrations after mixing were 30 nM scaffold, 150 nM Pol II or Pol I, 10 mM $MgCl_2$, and 1 mM NTP. Reactions were quenched at various times (0, 0.3, 1, 3, 10, 30, 90 min) by addition of one volume of 0.5 M EDTA (pH 8.0). For the transcript cleavage assay, the scaffold was incubated with the RNA polymerase for 10 min on ice, the cleavage reaction started by adding 5 mM $MgCl_2$, and the temperature switched to room temperature (23ºC) by water bath. In experiments with TFIIS-stimulated cleavage assay, TFIIS was added to the reaction system at time 0 (final concentration of TFIIS: 100 nM). Cleavage reactions were quenched at various time points (0, 1, 3, 10, 30 min). For the chase experiment, 1 mM NTP was added to system with additional 60s incubation before the reaction is quenched. The transcript was analyzed by 12% (wt/vol) denaturing urea/PAGE. The gel was visualized by phosphorimaging and analyzed using Image Laboratory software (BioRad). The DNA and RNA oligo sequences used in this assay were: non-template DNA (5'-GCAGCCTAGTTGATCTCATATTTCATTCCTACTCAGGAGAAGGAGCAGAGC G-3'), CPD-containing template DNA (5'-CGCTCTGCTCCTTCTCCTTTCCTCTCGATGGCTATGAGATCAACTAGGCTGC-3'), where TT represents a thymine dimer; RNA oligos: (8mer for transcription extension: 5'- AUCGAGAG-3'; 11mer for cleavage assay: 5'- AUCGAGAGGAA-3').In all cases, HPLC-purified oligos were obtained from TriLink.

Protein purification. Purification of yeast Pol II was performed as reported (31). Yeast TFIIS was cloned and expressed essentially as previously described (33). The wild-type and mutants of E.coli RNAP were purified by Ni-NTA, heparin and Q-sepharose chromatography as described previously (34). Purification of yeast Pol I and Pol I-dA12Ct, which lacks the 47 C-terminal residues in subunit A12.2, was performed following described procedures (35) with minor modifications. A *Saccharomyces cerevisiae* strain harboring a tandem affinity-purification (TAP) tag at the C-terminus of subunit A190 was grown in a fermenter containing 50 L of YPDA for 20 h at 30 °C to an OD_{600} of 5.5, harvested by centrifugation and stored at -80°C until use. 500 g of cells were resuspended in Buffer A (250 mM Tris-HCl pH 7.4, 20% glycerol, 250 mM ammonium sulfate, 1 mM EDTA, 10 mM $MgCl_2$, 10 μM $ZnCl_2$, 10 mM beta-mercaptoethanol) supplemented with protease-inhibitors (1 mM phenylmethanesulfonyl fluoride, 2 μg/ml leupeptine, 4 mM benzamidine, 1.4 μg/ml pepstatine A), phosphatase inhibitors (50 mM sodium fluoride, 2 mM sodium pyrophosphate, 5 mM beta-glycerophosphate) and DNAse (Roche) and lysed at 4°C with glass beads using a BeadBeater (Biospec). After centrifugation, the supernatant was incubated with 4 ml of IgG Sepharose (GE Healtcare) for 6 hours, washed with buffer B (50 mM Tris-HCl pH 7.4, 5% glycerol, 200 mM NaCl, 1 mM $MgCl_2$, 10 μM $ZnCl_2$ and 5 mM dithiothreitol)

and incubated overnight with tobacco etch virus (TEV) protease at 4°C. The TEV eluate was further purified using a Mono-Q (GE Healtcare) and a gradient from buffer B to the same buffer containing 1 M NaCl. Pol I-containing fractions were pooled and concentrated to 5 mg/ml, frozen in liquid nitrogen and stored at -80°C for further use.

Preparation of CPD-stalled Pol I for structural studies. Non-template DNA (5'-GCAGCCTAGTTGATCTCATAGCCCATTCCTACTCAGGAGAAGGAGCAGAGC G-3'), CPD-containing template DNA (5'-CGCTCTGCTCCTTCTCCTTTCCTCTCGATGGCTATGAGATCAACTAGGCTGC-3') and RNA (5'- AUCGAGAGGA-3') from Trilink Biotechnology were resuspended in buffer C (10 mM Hepes pH 8.0, 150 mM NaCl). The DNA strands were mixed in equimolar amounts, heated to 95°C and slow-cooled to 4°C. After, the RNA was added in equimolar amounts, the mixture heated to 45°C and slow-cooled to 4°C. The CPD-containing scaffold was then incubated at equimolar amounts with Pol I, at a final concentration of 0.22 mg/ml, for one hour at 20°C in buffer D (10 mM Hepes pH 8, 150 mM NaCl, 5 mM DTT).

Cryo-EM sample preparation and data acquisition. 3.5 µl of CPD-stalled Pol I were applied to glow-discharged copper 400 mesh R2/1 grids (Quantifoil) and incubated in the chamber of a FEI Vitrobot Mark III at 4°C and 95% humidity for 15 sec. The grids were blotted for 3 sec at an offset of –3 mm on both sides and vitrified by plunging into liquid ethane cooled down to liquid nitrogen temperature. Data were collected on a FEI Titan Krios electron microscope operated at 300 kV, using a K2 summit direct electron detector (Gatan) and equipped with the FEI automated single particle acquisition software (EPU). Images were acquired at defocus values varying between 1.5 and 4.0 µm and nominal magnification of 47,170x, yielding a pixel size of 1.06 Å. The camera was operated in dose-fractionation counting mode collecting 32 frames per movie, with a dose rate of 5.9 e– per pixel per sec for 8 sec total exposure.

Cryo-EM data processing. 2,056 movies were aligned using optical flow correction as implemented in Scipion (36) and their CTF parameters estimated using Ctffind4 (37). Manual inspection of Thong rings allowed discarding 178 micrographs outside of the defocus windows, with drift or astigmatism. In a first step, 1,864 particles were picked manually and reference-free 2D classes were generated, five of which were used for template-based auto-picking after low-pass filtering to 20 Å. Approximately 460,000 particles were automatically selected and extracted with a 288 pixel box using Relion (38), also employed for subsequent processing. Four rounds of reference-free 2D class averaging yielded a stack of 291,639 good-quality particles that were used to generate an initial 3D model employing a reference generated from the PDB entry 4C3I filtered to 60 Å. Four independent runs of 3D classification were performed using masks for the nucleic acid scaffold inside the cleft, A49-Nt/A34.5, the stalk, and upstream DNA plus A49-Ct. For each case, the best classes were selected and subsequently refined, followed by correction of local defocus of the individual particles using GCTF (39) and local motion and radiation damage using particle polishing as implemented in Relion. Final post-processing was performed using automatic masking and B-factor sharpening. FSC and local resolution estimations were performed using the routines implemented in Relion (38). In the highest-resolution map, derived from the first particle subset, ~40% of the particles fell into classes with a similar dominant view. While 3D FSC analysis (40) showed that the effect is not severe (Supplementary Figure 3), we followed described procedures (41) to balance particle orientations by removing three quarters of particles in dominant view classes. The resulting map presents similar 3D FSC statistics and features but slightly worse resolution (Supplementary Figure 3) and superposition

of the models derived from both maps shows a rmsd of 0.4 Å. Subsequent structural analysis was performed with the model derived from the initial map.

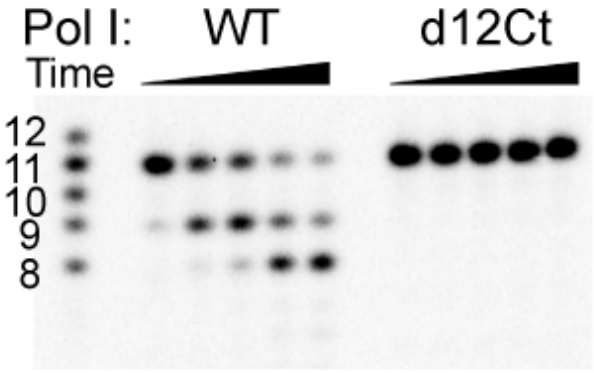
Model building and refinement. Available structures of undamaged Pol I EC (PDBs: 5M3F, 5M5X, 5W66) were fitted into the 3D maps using UCSF Chimera (42) and used as starting point for model building, which was performed in Coot (43). The structures were refined using real space refinement as implemented in Phenix (44). Refinement statistics are summarized in Supplementary Table 1. Figures were prepared using PyMOL (Schrödinger) and USCF Chimera (42).

Supplementary References (not included in the main text)

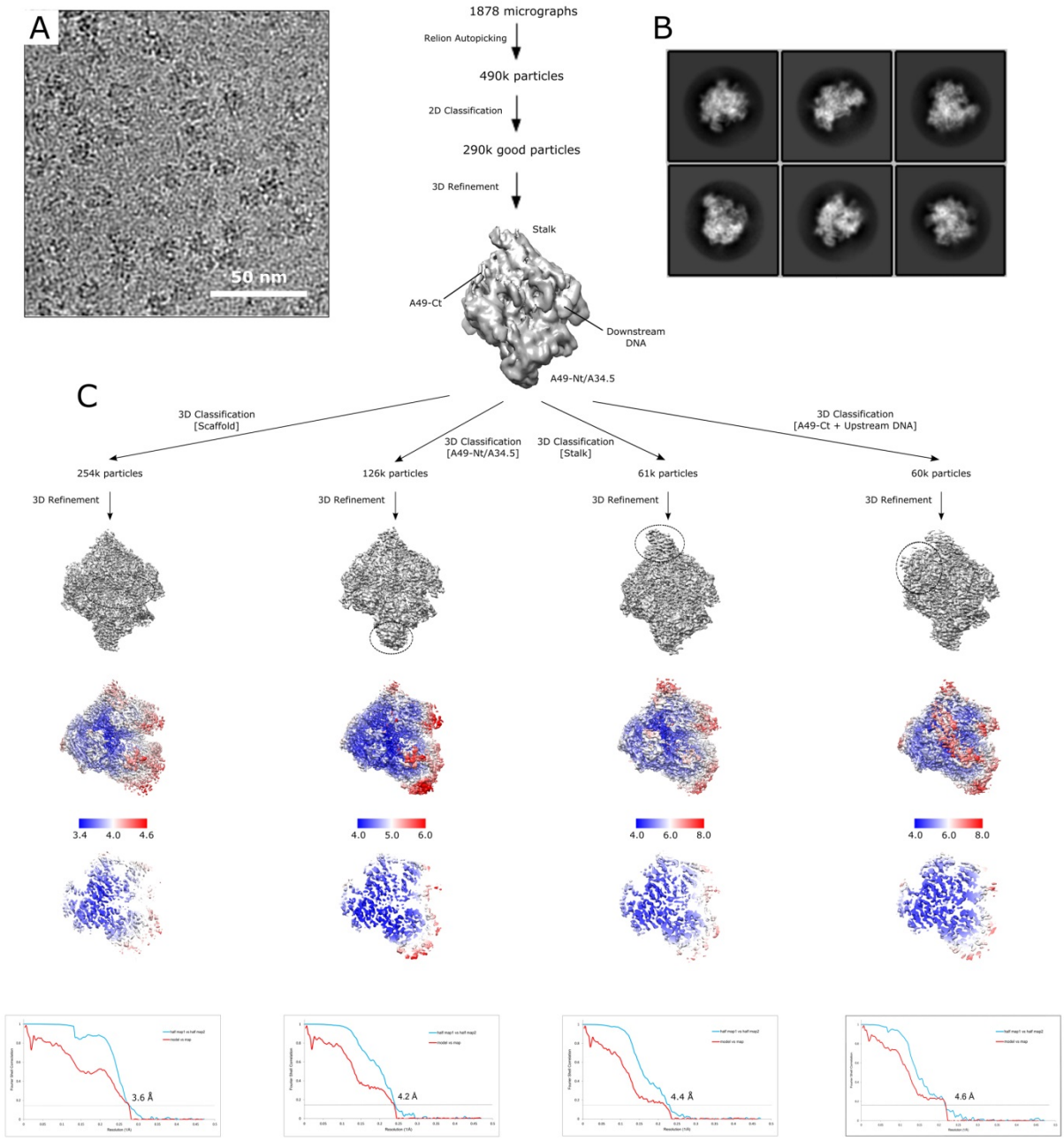
36. Abrishami V*, et al.* (2015) Alignment of direct detection device micrographs using a robust Optical Flow approach. *J Struct Biol* 189:163-176.
37. Mindell JA & Grigorieff N (2003) Accurate determination of local defocus and specimen tilt in electron microscopy. *J Struct Biol* 142:334-347.
38. Kimanius D, Forsberg BO, Scheres SH, & Lindahl E (2016) Accelerated cryo-EM structure determination with parallelisation using GPUs in RELION-2. *Elife* 5:e18722.
39. Zhang K (2016) Gctf: Real-time CTF determination and correction. *J Struct Biol* 193:1-12.
40. Tan YZ*, et al.* (2017) Addressing preferred specimen orientation in single-particle cryo-EM through tilting. *Nat Methods* 14:793-796.
41. Urnavicius L*, et al.* (2015) The structure of the dynactin complex and its interaction with dynein. *Science* 347:1441-1446.
42. Pettersen EF*, et al.* (2004) UCSF Chimera--a visualization system for exploratory research and analysis. *J Comput Chem* 25:1605-1612.
43. Emsley P, Lohkamp B, Scott WG, & Cowtan K (2010) Features and development of Coot. *Acta Cryst D*66:486-501.
44. Adams PD*, et al.* (2010) PHENIX: a comprehensive Python-based system for macromolecular structure solution. *Acta Cryst D*66:213-221.

Supplementary Table 1. Cryo-EM data collection and refinement statistics.

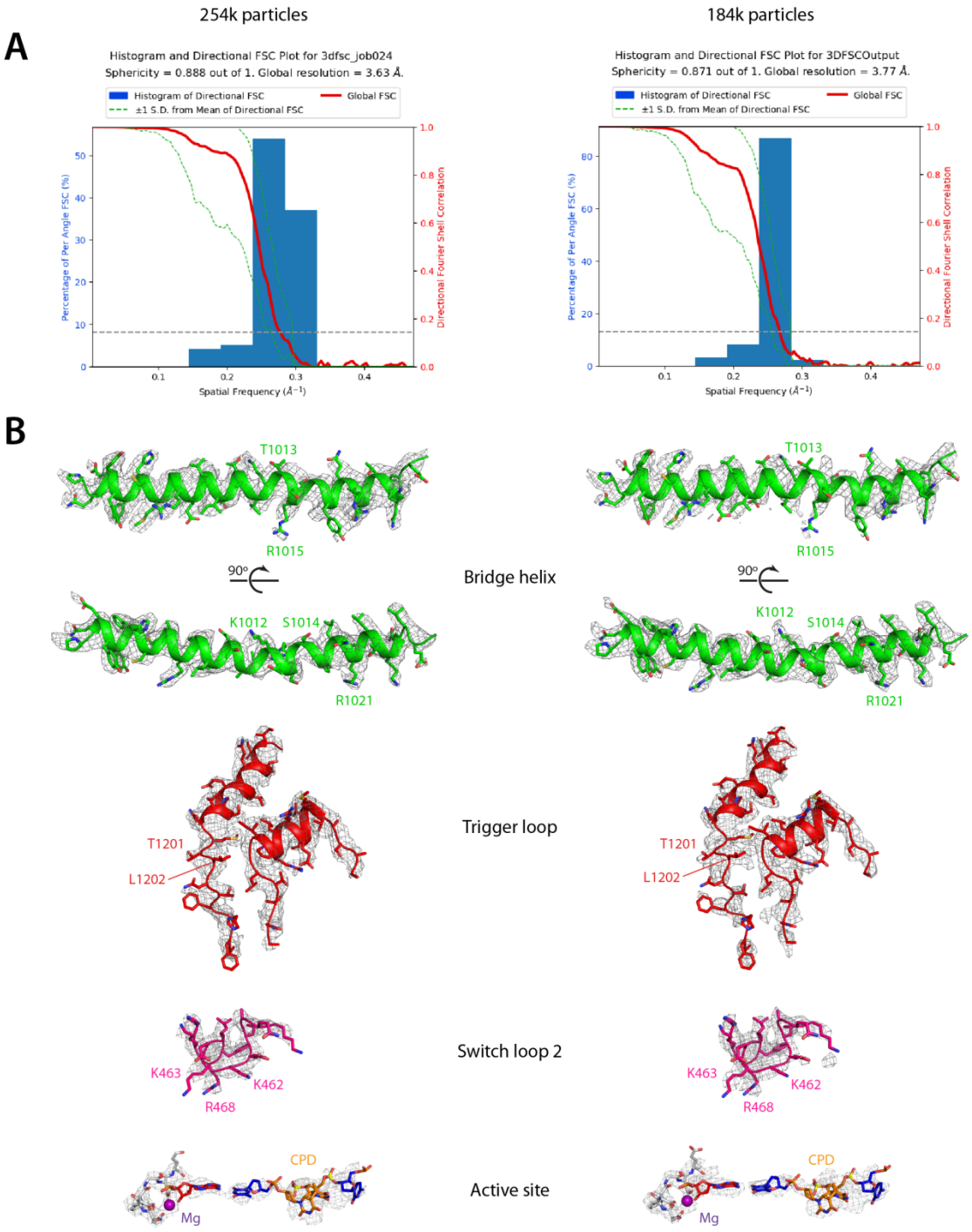
	Pol I-CPD	Pol I-CPD + uptream-DNA
Pixel size (Å/pixel)	1.06	
Number of grids	1	
Days of data collection	2	
All micrographs	2,056	
Selected micrographs	1,878	
Particles after 2D-class	291,639	
Final number of particles	254,079	60,297
Resolution (Å)	3.6	4.6
AccuracyRotations (º)	1.46	1.78
AccuracyTranslations (pixel)	0.80	1.05
Map sharpening B-factor (Å^2)	-94	-133
Ramachandran plot		
Outliers (%)	0%	0%
Allowed (%)	7.5%	8.5%
Favoured (%)	92.5%	91.5%
Map CC (around atoms)	0.711	0.711
RMSD bond lengths (Å)	0.06	0.05
RMSD bond angles (°)	0.87	0.82
All-atom clashscore	5.46%	6.70%
Rotamer outliers (%)	0.05%	0.05%
C-beta deviations	0	0
EMDB code	EMD-0146	EMD-0147
PDB code	6H67	6H68



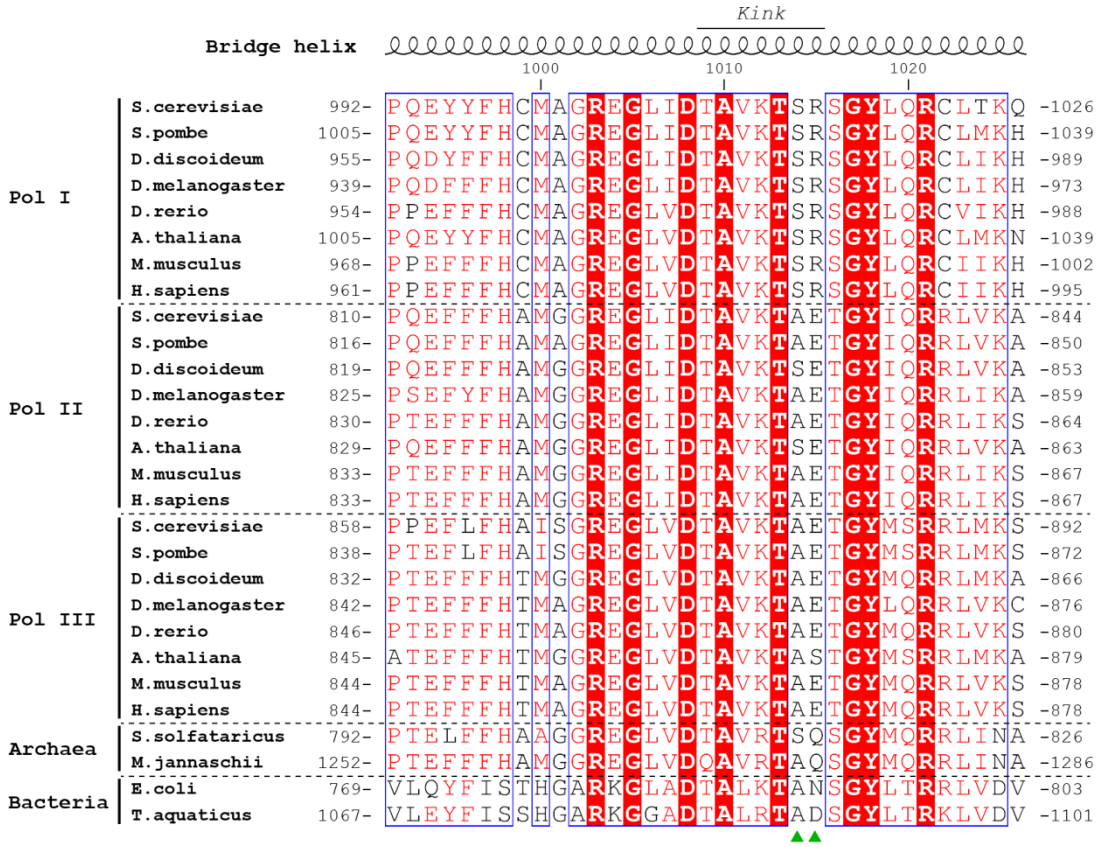
Supplementary Figure 1. Inactivation of Pol I intrinsic cleavage in Pol I-dA12Ct. Deletion of A12-Ct results in the inactivation of the intrinsic cleavage activity of Pol I. The methods and scaffold used in this figure are the same as in Figure 1D.



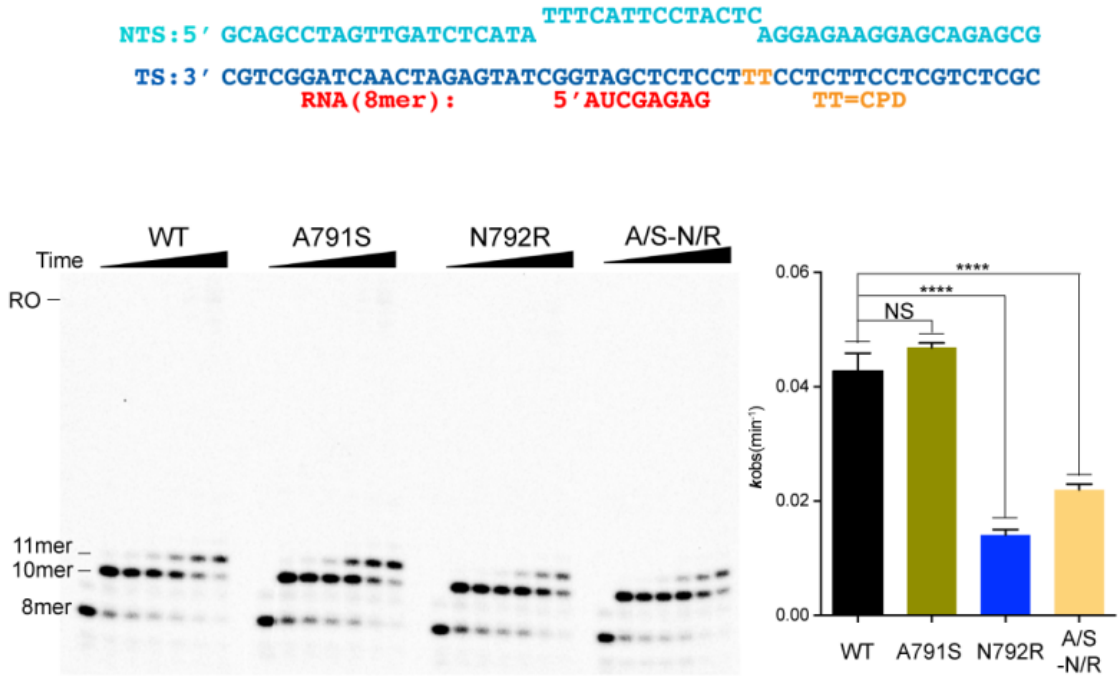
Supplementary Figure 2. Cryo-EM structure of CPD-stalled Pol I. (A) Typical field of the cryo-EM grids. The scale bar represents 50 nm. (B) Initial reference-free 2D averages showing a significant level of detail. (C) Data processing strategy, showing the first 3D-refine with good-quality particles and the final maps resulting from the three focused classifications. Below, local resolution estimation according to the legend. For each map, an overall view (above legend) and a central slice (below legend) are presented. At the bottom row, FSC curves between half maps according to gold-standard (blue) and between the cryo-EM map and derived atomic model (red).



Supplementary Figure 3. Quality analysis of the highest resolution cryo-EM map. The map derived from all particles in this subset includes is on the left, while the map obtained after removal of particles in preferential views is on the right. (A) 3D-FSC analysis including global half map FSC (red line), spread of directional resolution values defined by ±1σ from the mean (green dotted lines) and histogram of values evenly sampled over the 3D FSC (blue bars). Sphericity and global resolution values are indicated above the graph. (B) Representative areas of the maps with fitted atomic models.



Supplementary Figure 4. Sequence alignment of RNA polymerase bridge helix. Fully and partially-conserved residues are boxed in red and white, respectively. Top numbering is for S. cerevisiae Pol I. Green triangles indicate Pol I-specific residues mutated in this work. Figure prepared with ESPript (http://espript.ibcp.fr).



Supplementary Figure 5. *In vitro* transcription assay of WT and mutants of *E. coli* RNAP. The reaction was chased by adding 1 mM of NTPs and stopped at identical time points as in Figure 1B. Quantification of the 10-mer RNA extension is shown on the right. Data are mean and SD (n = 3). ****P < 0.0001; NS, not significant, two-tailed Student's *t*-test. A/S-N/R designates the double mutant A791S-N792R in *E. coli* RNAP subunit ß'.